

BIOGRAF
**Reference
Manual**



**Molecular
Simulations**

BIOGRAF

Reference Manual

Version 3.0

February 17, 1992

**Molecular Simulations, Inc.
200 Fifth Ave.
Waltham, MA 02154**

© 1992 Molecular Simulations, Inc. All rights reserved.

Note: The words BioDesign, BIOGRAF, POLYGRAF, NMRgraf, Polaris, CERIUS, COMSTRUCT, and Molecular Simulation System are trademarks of Molecular Simulations, Inc. DEC 5000 is a trademark of Digital Equipment Corp. Macintosh is a trademark of Apple Computer Inc. AIX, RISC System/6000, PC, PC/XT, and PC/AT are trademarks of IBM Corp. ESV is a trademark of Evans and Sutherland Corp. UNIX is a trademark of AT&T Bell Laboratories. IRIS, Personal IRIS, Power IRIS, and Indigo are trademarks of Silicon Graphics, Inc. Sun Workstation, Sun-4, and Network File System are trademarks of Sun Microsystems. Alliant, FX/Series, FX/Fortran, and Concentrix are trademarks of Alliant Computer Systems Corp. CONVEX is a trademark of Convex Computer Corp. Stardent, Doré, and Titan are trademarks of Stardent Computer Corp. Cray is a trademark of Cray Research, Inc., PostScript is a trademark of Adobe Systems Inc. Chem-X is a trademark of Chemical Design Ltd.

Contents

1	Introduction	1-1
1.1	Starting BIOGRAF.....	1-2
1.1.1	Resetting the Program.....	1-3
1.2	Manipulating the Graphics Display Image.....	1-4
1.2.1	The Dials Unit.....	1-4
1.2.2	The Mouse.....	1-6
1.2.3	The Function Keys	1-8
1.3	Viewing the Image in Three Dimensions.....	1-10
1.3.1	Titan's Stereo Monitor.....	1-10
1.3.2	The Personal Stereo Viewer for the IRIS.....	1-10
1.4	The Graphical User Interface (GUI)	1-11
1.4.1	The Graphics Window	1-12
1.4.2	The Modules/Messages Window	1-12
1.4.3	The Menus.....	1-12
1.4.4	The DISPLAY, TOP MENU, and COMMAND Buttons.....	1-13
1.4.5	The GRAPHICS, TEXT, and PLOT Buttons	1-13
1.5	Communicating with the User.....	1-14
1.5.1	Selecting Menu Boxes and Buttons and Picking Atoms on the Display	1-14
1.5.2	Dialog and Alert Boxes.....	1-15
1.6	Entering Filenames.....	1-16
1.6.1	The Browser Box	1-16
1.6.2	The New Dir and New File Buttons.....	1-17
1.7	Using the Menus.....	1-18
1.7.1	The Functions Menus	1-18
1.7.2	Basic Conventions for Using Menus.....	1-19
1.8	Running the Program in the Background.....	1-21
1.9	Basic Operations	1-22
2	In-Out Operations	2-1
2.1	READ	2-2
2.1.1	BIODESIGN	2-3
2.1.2	BROOKHAVEN.....	2-3
2.1.3	CAMBRIDGE.....	2-4
2.1.4	CHEMLAB CHM	2-4
2.1.5	MACROMODEL.....	2-4
2.1.6	MOLEDIT.....	2-4
2.1.7	MOLFILE.....	2-4
2.1.8	MOPAC.....	2-4
2.1.9	USER.....	2-5
2.1.10	FF PARAMETERS	2-5

2.1.11	GROUP	2-5
2.1.12	ROTATE/TRANSLATE.....	2-5
2.1.13	DOT.....	2-6
2.1.14	VECTOR.....	2-6
2.2	WRITE.....	2-7
2.2.1	BIODESIGN	2-7
2.2.2	BROOKHAVEN.....	2-8
2.2.3	CSSR.....	2-8
2.2.4	CHEMLAB CHM.....	2-9
2.2.5	MACROMODEL.....	2-9
2.2.6	MOLEDIT.....	2-9
2.2.7	MOLFILE.....	2-9
2.2.8	MOPAC.....	2-9
2.2.9	BIG STRAIN 2.....	2-10
2.2.10	USER.....	2-10
2.2.11	FF PARAMETERS	2-10
2.2.12	GROUP	2-10
2.2.13	ROTATE/TRANSLATE.....	2-11
2.3	COPY FILE.....	2-12
2.4	CAPTION.....	2-13
2.4.1	SELECT OBJECTS	2-14
2.4.2	DEFINE.....	2-14
2.4.3	DISPLAY	2-15
2.4.4	COLOR	2-15
2.4.5	ORIENT	2-16
2.4.6	CHANGE FONT.....	2-17
2.4.7	DEFAULT FONT	2-17
2.4.8	DEFAULT COLOR.....	2-17
2.5	PRINT SCREEN.....	2-19
2.5.1	GRAY/BLACK&WHITE/COLOR	2-19
2.5.2	NEGATIVE/POSITIVE.....	2-19
2.5.3	CREATE PS FILE	2-19
2.6	AUTO GROUP	2-20
2.7	HALF BOND	2-20
2.8	CENTER NEXT/CENTER ALL/NO CENTER.....	2-21
2.9	BY FILE.....	2-21
3	Build Operations	3-1
3.1	ORGANIC.....	3-3
3.1.1	SETUP.....	3-4
3.1.2	NEW ATOM IDS.....	3-5
3.1.3	READ LIBRARY	3-5
3.1.4	BASE FRAGMENT, FUNCTIONAL GRP, USER BASE FRAG, USER FUNC GRP	3-6
3.1.5	READ FRAGMENT	3-7
3.1.6	COPY	3-8
3.1.7	CONDENSE.....	3-8
3.1.8	REPLACE.....	3-8
3.1.9	CONNECT FRAG	3-9
3.1.10	DRAW.....	3-9

3.1.11	DEFAULT ATOM.....	3-11
3.1.12	EDIT.....	3-12
3.1.13	EDIT ID.....	3-20
3.1.14	MECHANICS	3-27
3.1.15	SIMULATE.....	3-27
3.1.16	ANALYZE.....	3-27
3.2	PEPTIDE.....	3-28
3.2.1	SETUP.....	3-28
3.2.2	L_LIBRARY, D_LIBRARY	3-29
3.2.3	USER LIBRARY	3-29
3.2.4	READ LIBRARY	3-30
3.2.5	DEFINE ANGLE	3-31
3.2.6	MODIFY	3-32
3.2.7	REPLACE	3-33
3.2.8	INSERT.....	3-35
3.2.9	DELETE.....	3-39
3.2.10	EDIT, EDIT ID	3-40
3.2.11	MECHANICS, SIMULATE	3-40
3.2.12	ANALYZE.....	3-40
3.3	RNA-DNA	3-41
3.3.1	SETUP.....	3-42
3.3.2	READ LIBRARY	3-42
3.3.3	SINGLE, DOUBLE, TRIPLE.....	3-42
3.3.4	USER.....	3-44
3.3.5	BAP.....	3-44
3.3.6	MODIFY SUGAR.....	3-45
3.3.7	EDIT, EDIT ID	3-45
3.3.8	MECHANICS, SIMULATE	3-45
3.3.9	ANALYZE.....	3-45
3.4	DNA-OLD.....	3-46
3.4.1	SETUP.....	3-46
3.4.2	EDIT, EDIT ID	3-46
3.4.3	MECHANICS, SIMULATE	3-47
3.4.4	ANALYZE.....	3-47
3.5	LIPID.....	3-48
3.5.1	SETUP.....	3-48
3.5.2	READ LIBRARY	3-48
3.5.3	BUILD.....	3-49
3.5.4	EXTEND.....	3-50
3.5.5	EDIT, EDIT ID	3-51
3.5.6	MECHANICS, SIMULATE	3-51
3.5.7	ANALYZE.....	3-51
3.6	CARBOHYDRATE.....	3-52
3.6.1	SETUP.....	3-52
3.6.2	READ LIBRARY	3-53
3.6.3	LINKAGE.....	3-53
3.6.4	DEFINE ANGLE	3-54
3.6.5	MODIFY CARBO	3-54
3.6.6	CARBO LIBRARY.....	3-55
3.6.7	USER LIBRARY	3-56
3.6.8	DEFINE.....	3-56

	3.6.9	REPEAT.....	3-56
	3.6.10	EXTEND.....	3-57
	3.6.11	EDIT, EDIT ID	3-58
	3.6.12	MECHANICS, SIMULATE	3-58
	3.6.13	ANALYZE.....	3-58
3.7		CRYSTAL.....	3-59
	3.7.1	SETUP.....	3-60
	3.7.2	INPUT FILE.....	3-60
	3.7.3	INPUT COORD	3-61
	3.7.4	EXTEND CELL.....	3-64
	3.7.5	UNEXTEND CELL.....	3-64
	3.7.6	CONNECT	3-65
	3.7.7	EXTRACT MOLEC	3-65
	3.7.8	DEFAULTS	3-66
	3.7.9	CHARGES	3-68
	3.7.10	SPACE GROUP.....	3-68
	3.7.11	EDIT, EDIT ID	3-68
	3.7.12	ANALYZE.....	3-68
3.8		SOLVATE.....	3-69
	3.8.1	SETUP.....	3-70
	3.8.2	SOLVENT.....	3-71
	3.8.3	GRID TYPE.....	3-72
	3.8.4	GRID SPACE.....	3-72
	3.8.5	INNER CUTOFF	3-72
	3.8.6	OUTER CUTOFF.....	3-72
	3.8.7	BOX SOLVENT	3-73
	3.8.8	AUTO GROUP	3-73
	3.8.9	HALF BOND	3-73
	3.8.10	EXCLUDE SELF.....	3-73
	3.8.11	EXECUTE.....	3-74
3.9		ROTATE BOND.....	3-75
	3.9.1	SETUP.....	3-76
	3.9.2	NBONDS	3-77
	3.9.3	NBRANCHES.....	3-77
	3.9.4	EXECUTE.....	3-77
	3.9.5	UPDATE	3-78
	3.9.6	CLEAR.....	3-78
	3.9.7	FORWARD, REVERSE	3-78
	3.9.8	FILE ORIGIN	3-78
	3.9.9	GROUP ORIGIN	3-79
	3.9.10	PICK ORIGIN.....	3-79
	3.9.11	ZERO ORIGIN	3-79
3.10		DOCK.....	3-80
	3.10.1	EXECUTE.....	3-81
	3.10.2	MODIFY COORDS.....	3-81
	3.10.3	ROTATE ALL	3-82
	3.10.4	SET ORIGIN.....	3-82
	3.10.5	ENERGY DOCK (Interactive Docking).....	3-83
3.11		MODIFY H	3-88
	3.11.1	ADD ALL H, ADD HET H	3-88
	3.11.2	DELETE ALL H.....	3-89

3.11.3	H MASS=H, H MASS=D, H MASS=T.....	3-89
3.12	MOMENT	3-90
3.13	CONVERT	3-91
3.13.1	READ CONV FILE	3-91
3.13.2	CONVERT	3-91
3.13.3	EDIT RESIDUE	3-91
3.14	CONNECT	3-93
3.14.1	SETUP	3-93
3.14.2	CONNECT ALL/NOT CONNECTED	3-93
3.14.3	ALL CELLS/UNIT CELL ONLY	3-94
3.14.4	MIN DISTANCE, MAX DISTANCE	3-94
3.14.5	CRITERIA	3-94
3.14.6	SCALE	3-94
3.14.7	EXECUTE	3-95
3.14.8	DELETE BONDS	3-95
4	Visualize Operations	4-1
4.1	MAKE GROUPS	4-2
4.1.1	SETUP	4-3
4.1.2	FILES	4-3
4.1.3	FULL FILE	4-3
4.1.4	CHAINS	4-4
4.1.5	RESIDUES	4-4
4.1.6	LIMITS	4-5
4.1.7	CREATE	4-9
4.1.8	RENAME FILE	4-9
4.1.9	RENAME GROUP	4-9
4.1.10	H BONDS	4-10
4.1.11	RIBBONS	4-10
4.2	RENDER GROUPS	4-12
4.2.1	VECTORS/CYLINDERS/SPHERES/BALL & STICK	4-12
4.2.2	HALF BOND	4-13
4.2.3	VECTOR WIDTH	4-13
4.2.4	CYLINDER SCALE, BALL SCALE	4-13
4.2.5	SHADING TYPE	4-14
4.2.6	DIVISIONS	4-14
4.2.7	SPECULAR INTEN	4-14
4.3	SURFACE	4-15
4.3.1	CALCULATE	4-16
4.3.2	READ	4-17
4.3.3	READ PART	4-17
4.3.4	SETUP	4-18
4.3.5	GROUPS	4-18
4.3.6	PROP DEFAULTS	4-19
4.3.7	PROPERTY	4-20
4.3.8	EXECUTE	4-21
4.3.9	WRITE	4-22
4.3.10	WRITE PART	4-22
4.3.11	Surface File Format	4-22
4.4	SET ORIGIN	4-24

4.4.1	FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN	4-24
4.4.2	CENTER	4-24
4.4.3	RESET DIALS	4-24
4.4.4	RESET ALL	4-24
4.5	CHANGE COLORS	4-25
4.6	STEREO	4-27
4.6.1	ON/OFF	4-28
4.6.2	MODE	4-28
4.6.3	ANGLE	4-29
4.6.4	SEPARATION	4-29
4.6.5	OFFSET	4-29
4.7	ROCK	4-31
4.8	AUTO ROTATE	4-32
5	Simulate Operations	5-1
5.1	EEX OPTIONS	5-3
5.1.1	FAST SETUP	5-3
5.1.2	RIGID MOLECULE	5-3
5.2	SETUP EEX	5-4
5.3	MODIFY EEX	5-5
5.3.1	MOVE ALL, FIX ALL	5-5
5.3.2	MOVE RANGE, FIX RANGE, FIX SINGLE	5-5
5.3.3	MOVE SINGLE, FIX SINGLE, EXCLUDE SINGLE	5-5
5.3.4	FIX GROUPS	5-6
5.4	SHOW EEX	5-7
5.5	CONSTRAINTS	5-8
5.5.1	DISTANCE	5-8
5.5.2	ANGLE	5-9
5.5.3	TORSION	5-9
5.5.4	INVERSION	5-9
5.5.5	K(DISTANCE), K(ANGLE), K(TORSION), K(INVERSION)	5-10
5.5.6	RESET	5-10
5.5.7	READ CONSTRAINT	5-10
5.5.8	WRITE CONSTRAINT	5-10
5.5.9	SHOW ALL/SHOW CURRENT	5-11
5.5.10	PROTEIN CONSTRNT	5-11
5.6	CHARGES	5-13
5.7	MECHANICS	5-14
5.7.1	ONE ENERGY	5-15
5.7.2	MINIMIZE	5-15
5.7.3	SIMPLE DYNAMICS	5-15
5.7.4	# OF STEPS	5-15
5.7.5	RMS FORCE CVRG	5-16
5.7.6	DEFAULTS	5-16
5.7.7	GEOMETRY PROP	5-16
5.7.8	XYZ TEMP SAVE	5-18
5.8	DYNAMICS	5-19
5.8.1	Functions on All Dynamics Menus	5-21
5.8.2	ADIABATIC Dynamics	5-26
5.8.3	QUENCHED Dynamics	5-27

5.8.4	ANNEALED Dynamics.....	5-28
5.8.5	IMPULSE Dynamics.....	5-29
5.8.6	CANONICAL (TVN) Dynamics	5-30
5.8.7	TEMPDAMP (TVN) Dynamics.....	5-31
5.9	CONSTRAINED MIN.....	5-32
5.9.1	DIRECTION START.....	5-32
5.9.2	DIRECTION STOP.....	5-32
5.9.3	PLANE/LINE/POINT.....	5-32
5.9.4	INCREMENT.....	5-32
5.9.5	UPDATE FREQ.....	5-33
5.9.6	POINTS.....	5-33
5.9.7	E CONVERGE.....	5-33
5.9.8	# OF STEPS	5-33
5.9.9	EXECUTE.....	5-34
5.10	GRID SEARCH	5-35
5.10.1	SETUP.....	5-36
5.10.2	NTORSIONS	5-36
5.10.3	EXAMINE	5-36
5.10.4	# OF STEPS	5-37
5.10.5	CONSTRAINTS ON/OFF	5-37
5.10.6	DRAW ON/OFF	5-37
5.10.7	NBOND ON/OFF	5-37
5.10.8	H BOND ON/OFF	5-37
5.10.9	DEFAULTS	5-38
5.10.10	EXECUTE.....	5-38
5.10.11	SET BEST, RESET BEST, BEST	5-38
5.10.12	READ PLOT.....	5-39
5.11	M-CARLO SEARCH.....	5-40
5.11.1	DEFAULTS	5-40
5.11.2	RINGS.....	5-41
5.11.3	SP2-SP2.....	5-41
5.11.4	CENTER	5-41
5.11.5	NBOND UPDATE.....	5-41
5.11.6	H BOND UPDATE.....	5-42
5.11.7	VDW SCALE.....	5-42
5.11.8	# OF STEPS	5-42
5.11.9	WRITE TRAJECT	5-42
5.11.10	EXECUTE.....	5-43
5.11.11	RESET.....	5-43
5.12	DEFAULTS	5-44
5.12.1	ENERGY VAR.....	5-46
5.12.2	MINIMIZE VAR.....	5-47
5.12.3	TEMPERATUR VAR.....	5-49
5.12.4	DYNAMICS VAR.....	5-51
5.12.5	NONBOND VAR	5-54
5.12.6	H BOND VAR	5-58
5.12.7	MISC VAR.....	5-59
5.13	MODIFY PARAMS.....	5-61
5.13.1	PARAM INOUT	5-62
5.13.2	MISSING PARAM	5-63
5.13.3	CHECK/MODIFY	5-63

5.13.4	BOND, ANGLE, INVERSION, ANG-ANG INVERS, BOND-X-ANGLE, TORSION, UREY-BRADLEY	5-64
5.13.5	NEW ATOM, VAN DER WAALS, OFFDIAG NONBND, CHARGE, HYDROGEN BOND	5-64
5.14	ANIMATE	5-65
5.14.1	DEFINE TRAJECT	5-66
5.14.2	START, STOP	5-66
5.14.3	LIST	5-67
5.14.4	READ	5-69
5.14.5	EXTRACT	5-71
5.14.6	AUTO GROUP	5-71
5.14.7	DISPLAY	5-72
5.14.8	AVERAGE COORDS	5-75
5.14.9	RMS COORDS	5-76
5.14.10	USER TRAJECT	5-76
6	Analyze Operations	6-1
6.1	IDENTIFY	6-2
6.1.1	SINGLE ATOM	6-2
6.1.2	GROUP	6-3
6.1.3	ATOM TYPE	6-3
6.1.4	ATOM CHARGE	6-3
6.1.5	ATOM MASS	6-4
6.1.6	ATOM LABEL	6-4
6.1.7	ATOM RADIUS	6-4
6.1.8	ATOM NUMBER	6-4
6.1.9	RESIDUE NAME	6-4
6.1.10	CHAIN	6-4
6.1.11	CLEAR	6-5
6.1.12	NEW GROUP	6-5
6.2	GEOMETRY	6-6
6.2.1	SINGLE	6-6
6.2.2	GROUP	6-7
6.2.3	DISTANCE	6-8
6.2.4	ANGLE	6-8
6.2.5	DIHEDRAL	6-9
6.2.6	INVERSION	6-9
6.2.7	PLANE	6-9
6.2.8	LINKAGE	6-10
6.2.9	DIFFERENCE	6-12
6.2.10	WRITE FILE	6-13
6.2.11	CLEAR, NEW GROUP	6-13
6.2.12	GROUP ORIGIN	6-14
6.2.13	PICK ORIGIN	6-15
6.3	ONE ENERGY	6-16
6.4	ENERGY-FORCE	6-17
6.4.1	SINGLE ENERGY	6-18
6.4.2	GROUP ENERGY	6-19
6.4.3	SNGL ENERGY/ATOM	6-20
6.4.4	GRP ENERGY/ATOM	6-20

6.4.5	SNGL FORCE/ATOM.....	6-21
6.4.6	GRP FORCE/ATOM	6-21
6.4.7	DEFAULTS	6-22
6.4.8	GROUP ORIGIN, PICK ORIGIN	6-24
6.4.9	CLEAR, NEW GROUP	6-24
6.5	CONTACTS.....	6-25
6.5.1	1-4 NONBOND.....	6-25
6.5.2	DISPLAY DISTANCE	6-26
6.5.3	DISPLAY ENERGY.....	6-26
6.5.4	VDW ENERGY CUT	6-26
6.5.5	EXECUTE.....	6-26
6.5.6	CLEAR, NEW GROUP	6-26
6.6	MATCH	6-28
6.6.1	GROUPS	6-28
6.6.2	PAIRS	6-28
6.6.3	CENTROIDS	6-29
6.6.4	EXECUTE.....	6-29
6.6.5	WRITE	6-29
6.7	TRAJECTORY	6-31
6.7.1	DEFINE.....	6-31
6.7.2	TRAJ DEFAULTS	6-32
6.7.3	PLOT DEFAULTS	6-33
6.7.4	READ PLOT	6-33
6.7.5	MAKE PLOT	6-33
6.7.6	REPLOT.....	6-35
6.7.7	IDENTIFY	6-36
6.7.8	EXTRACT	6-36
6.7.9	WRITE PLOT	6-36
6.7.10	WRITE BGF	6-37
6.7.11	FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN	6-37
6.8	PLOT.....	6-38
6.8.1	DEFINE.....	6-38
6.8.2	NEW GROUP	6-38
6.8.3	PLOT DEFAULTS	6-39
6.8.4	REPLOT.....	6-45
6.8.5	IDENTIFY	6-46
6.8.6	MERGE.....	6-46
6.8.7	WRITE PLOT	6-47
6.8.8	FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN	6-47
6.9	PROTEIN.....	6-48
6.9.1	STEREOCHEMISTRY.....	6-48
6.9.2	PEPTIDE LINKAGE	6-48
6.9.3	PHI-PSI ANGLES.....	6-48
6.9.4	DISULFIDE BONDS.....	6-49
6.10	VOLUMES.....	6-50
6.10.1	CALCULATE GRID	6-51
6.10.2	DEFAULTS	6-52
6.10.3	UNION, INTERSECTION, RESIDUAL, INVERSE.....	6-53
6.10.4	SURFACE ONLY	6-53
6.10.5	DISPLAY	6-54
6.10.6	CUBIC VOLUME.....	6-54

6.10.7	SURFACE AREA	6-54
6.10.8	READ	6-55
6.10.9	WRITE	6-55
7	Utilities.....	7-1
7.1	USER.....	7-1
7.2	PROGRAM RESET	7-2
7.3	LABEL COLOR	7-2
7.4	BROWSER	7-2
8	Exit Operations.....	8-1
A	Standard Group Elements.....	A-1
A.1	Standard Residues	A-1
A.2	Standard Structural Elements (Primitives).....	A-2
A.3	Standard Structural Descriptors (Selectors).....	A-3
B	Fragment Libraries.....	B-1
B.1	BG_ORG_FRG	B-2
B.2	BG_ORG_FUN.....	B-2
B.3	BG_SOLVENT.....	B-3
B.4	BG_PEP_L, BG_PEP_D.....	B-3
B.5	BG_LIPID	B-5
B.6	BG_CARBO	B-6
B.7	BG_DNA_SINGLE	B-7
B.8	BG_DNA_DOUBLE	B-7
B.9	BG_DNA_TRIPLE	B-8
C	Supplementary Files.....	C-1
C.1	Data Files.....	C-1
C.2	Macro Files.....	C-3
C.3	User Files.....	C-4
D	File Formats.....	D-1
D.1	Brookhaven	D-1
D.2	BioDesign General Format.....	D-3
D.3	Atom Labels	D-8
D.4	Standard File Name Extensions	D-9
D.5	Trajectory Files	D-10
D.6	Constraints Files.....	D-19
D.7	Group Files.....	D-21
D.8	Plot Files.....	D-30
D.9	Defaults File	D-31
D.9.1	Graphics Parameters.....	D-31
D.9.2	General Defaults.....	D-32

D.9.3	Energy Parameters.....	D-32
D.9.4	Energy Switches.....	D-33
D.9.5	Dynamics Switches.....	D-33
D.9.6	Nonbond Switches.....	D-33
D.9.7	Nonbond Variables.....	D-33
D.9.8	Hydrogen Bond Variables.....	D-33
D.9.9	Miscellaneous Energy Variables.....	D-34
D.9.10	Energy and Force Analysis Variables.....	D-35
D.9.11	Contact Check Analysis Variables and Switches.....	D-36
D.10	Brookhaven Connect Files.....	D-37
D.11	Atom and Residue Synonyms File.....	D-38
E	Units.....	E-1
F	Rotations.....	F-1
G	Force Fields.....	G-1
G.1	Topology and Automatic Energy Expressions.....	G-2
G.1.1	Nonperiodic.....	G-2
G.1.2	Periodic.....	G-3
G.2	Bond Stretch.....	G-3
G.3	Bond Angle Bend.....	G-6
G.3.1	Cosine Expansions.....	G-6
G.3.2	Theta Expansions.....	G-8
G.3.3	Polyene.....	G-10
G.4	Dihedral Angle Torsions.....	G-10
G.5	Inversions.....	G-11
G.6	Electrostatic Interactions.....	G-13
G.6.1	Charges.....	G-13
G.6.2	Coulomb Summations.....	G-14
G.6.3	Nonbond Exclusions.....	G-15
G.6.4	Dielectric Constant.....	G-15
G.6.5	Nonbond Lists.....	G-16
G.7	van der Waals Interactions.....	G-16
G.8	Combination Rules (Off-Diagonal van der Waals).....	G-19
G.8.1	Geometric Mean.....	G-19
G.8.2	Arithmetic Mean.....	G-20
G.8.3	Conventions.....	G-20
G.8.4	Nonbond Off-Diagonal.....	G-20
G.9	Hydrogen Bond.....	G-21
G.10	Urey-Bradley.....	G-22
G.11	Default Potential Terms for the Standard Force Fields.....	G-23
H	Long-Range Cutoffs.....	H-1

I	Parameters	I-1
I.1	Parameter File Format	I-2
I.1.1	Parameter File Header	I-3
I.1.2	Defaults Parameters (optional)	I-3
I.1.3	Force Field Type Labels Parameters	I-9
I.1.4	Implicit Hydrogen Relationships Parameters	I-10
I.1.5	Lone Pairs Parameters	I-10
I.1.6	Charge Equilibration Parameters	I-10
I.1.7	Del Re Charge by Element Parameters	I-11
I.1.8	Del Re Charge by Force Field Parameters	I-11
I.1.9	Gasteiger Charge Parameters	I-12
I.1.10	Atomic Parameters	I-12
I.1.11	Auto Type Parameters	I-13
I.1.12	Off-diagonal Nonbond Parameters	I-14
I.1.13	Bond Stretching Parameters	I-14
I.1.14	Bond Angle Bending Parameters	I-15
I.1.15	Dihedral Angle Torsion Parameters	I-17
I.1.16	Inversion (Improper) Torsion Parameters	I-17
I.1.17	Hydrogen Bond Parameters	I-18
I.1.18	User-supplied Parameters	I-18
I.2	Conversion File Format	I-18
I.3	The Dreiding-I and Dreiding-II Force Fields	I-19
I.3.1	Atom Types	I-19
J	Dynamics Algorithms	J-1
J.1	Unit of Time	J-1
J.2	Standard Verlet	J-1
J.2.1	2 Point Verlet	J-2
J.2.2	1 Point Verlet	J-2
J.3	Summed Verlet	J-3
K	Command and Macro Modes	K-1
K.1	Command Mode	K-1
K.1.1	Menu Commands	K-1
K.1.2	Multiple Commands	K-2
K.1.3	Continuation Characters	K-2
K.1.4	Abbreviations and Wild Cards	K-2
K.1.5	Ambiguous Commands	K-2
K.1.6	Two-State and Cyclic Toggles	K-3
K.1.7	No-opt Commands	K-3
K.1.8	Atom Referencing Commands	K-3
K.1.9	Dial Commands	K-4
K.1.10	Special Commands Summary	K-5
K.2	Macros	K-7
K.2.1	Variables	K-8

K.2.2	Inquire and Set Functions.....	K-10
K.2.3	Math Functions.....	K-11
K.2.4	Logical Expressions	K-12
K.2.5	If..then..else Blocks	K-12
K.2.6	Branching and Statement Labels.....	K-13
K.2.7	Loop Constructs	K-13
K.2.8	Input/output Statements.....	K-14
K.2.9	Other Macro Functions.....	K-15
K.2.10	User Routines	K-16
K.2.11	Debugger	K-16
K.3	Command Summary.....	K-17
K.4	Sample Macros.....	K-18
L	Dreiding: A Generic Force Field for Molecular Simulations.....	L-1
M	Standard Space Group Names	M-1
N	Auxiliary Programs.....	N-1
N.1	SHTRJ	N-2
N.2	RDTRJ.....	N-2
N.3	WRTRJ.....	N-3
N.4	Transferring Trajectory Files Between Computers with Incompatible Trajectory File Formats	N-3
O	Reading and Building Structures with the AMBER Force Field in BIOGRAF	O-1
O.1	Building RNA or DNA	O-1
O.2	Reading DNA BioDesign General Format or Brookhaven Files.....	O-1
O.3	Building Peptides	O-2
O.4	Reading Peptide BioDesign General Format Files (created using the Dreiding force fields)	O-2
O.4.1	Files Without Hydrogen Atoms	O-2
O.4.2	Files With Hetero Hydrogens (but no hydrogens on carbons).....	O-2
O.4.3	Files With All Their Hydrogens (no implicit hydrogens).....	O-2
O.5	Reading Peptide Brookhaven Files	O-3
P	Equilibrium Molecular Dynamics	P-1
P.1	Constant Energy Constant Volume Dynamics.....	P-2
P.2	Constant Enthalpy Constant Pressure Dynamics	P-2
P.3	Constant Temperature Constant Volume Dynamics	P-3
P.4	Constant Temperature Constant Pressure Dynamics	P-3
P.5	Non-Hamiltonian Forms of Constant Temperature Dynamics	P-3
P.6	Nosé-Hoover Canonical Dynamics	P-4
Index	Index-1	

Chapter 1

Introduction

BIOGRAF is a menu-driven program for biological and chemical simulations. The program builds molecular structures, optimizes conformational energies, calculates the dynamics of motion for molecular systems, and analyzes molecular structure and dynamics. BIOGRAF can be run interactively or as a batch job (in the background). It employs an X-Windows based graphical user interface (GUI), making it highly portable and capable of running on a wide range of platforms. The interface is designed to be both easy to learn and easy to use. The program has the power to handle structures with more than ten thousand atoms, yet is simple enough to use on diatomic molecules.

BIOGRAF is driven by selecting operations from the Functions menus at the left of the Menu Pad. Other menus are also used to display information or choices which are needed to execute a particular function. The program may also be driven by entering commands from the keyboard. This command mode feature is particularly useful in that it allows the user to create macros so that a series of commands can be automatically executed. A description of the use of command mode and macros is given in Appendix K.

This chapter describes the human interface. It includes information on how to use the menus, dialog boxes, and buttons. It also provides a general introduction to the program, including information on basic operations such as starting and resetting the program, manipulating the graphics display, viewing the image in three dimensions, using the menus, inputting data, and selecting atoms. A brief description of the top level functions is also included at the end of the chapter. It is assumed that the user is familiar with basic windowing operations including how to open, close, push down, pop up, reposition, and resize windows, as well as how to use scroll bars and slider bars.

This manual follows the logic of the Functions menus. After this introductory chapter there are seven chapters which correspond to the seven operations displayed on the top level menu. These are the IN-OUT, BUILD, VISUALIZE, SIMULATE, ANALYZE, UTILITIES, and EXIT operations. Sections correspond to operations on submenus displayed when one of these top level functions has been selected. Similarly, subsections correspond to subsubmenus. Throughout this manual, all menu items are shown in uppercase (*e.g.*, BUILD).

1.1 Starting BIOGRAF

BIOGRAF is X-Windows-based and runs on systems using the UNIX operating system. It is therefore highly portable and presents a common interface on a wide range of platforms. If you have not yet installed BIOGRAF, go through the procedures on how to use the COMSTRUCT program described in the *Molecular Simulations Installation Guide*. COMSTRUCT is used to create the command procedures and defaults files that assign all the directories, input and output files and devices that BIOGRAF needs to operate and then run the program.

To start up BIOGRAF, log on and type in the start up command. (*biograf* is commonly used, but this command may be different at your site). Note that UNIX is case sensitive. For example, the name "biograf" is different than "BIOGRAF" or "Biograf". All standard BIOGRAF commands and file names are in lower case with UNIX.

A start up window appears displaying the logo and copyright notice. The defaults, parameter conversion files, and other pertinent data are read in and the final start up display appears containing the five windows as shown in Figure 1.1. If you want to read in and display a molecule on start

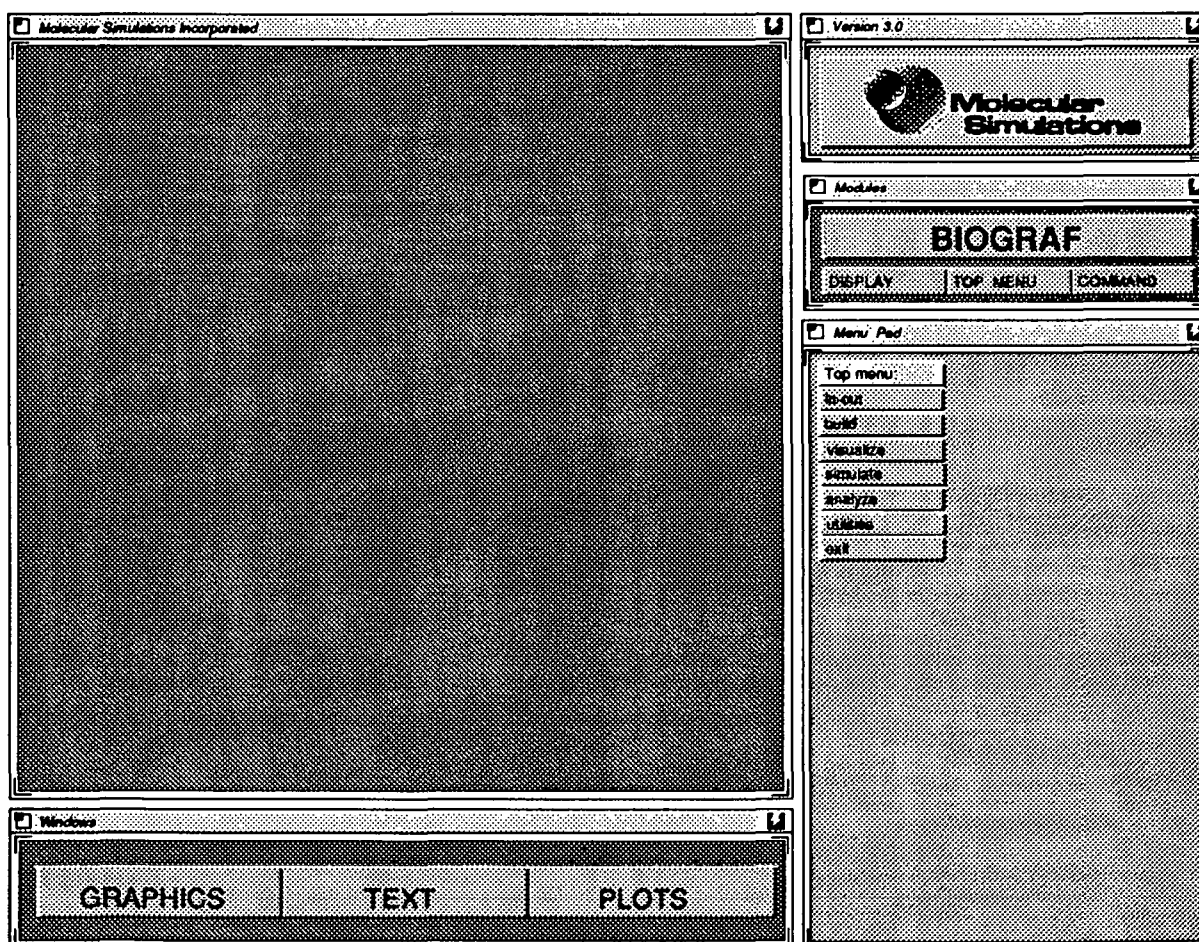


Figure 1.1: BIOGRAF start up display

up, this can be done by issuing a macro written to do this instead of entering the start up command or if the molecule is in a BioDesign General Format (*bgf*) file it can be read in by including the file name at the end of the start up command (*e.g., biograf filename.bgf*).

1.1.1 Resetting the Program

There may be times when you would like to work on a different molecule and would like to clear out all the atoms, groups, and files left over from the current session so that you can "start from a clean slate". This can be done using the program reset function. This allows you to start over without taking the extra time to quit BIOGRAF and restart it. The program reset option is found on the Utilities menu.

1.2 Manipulating the Graphics Display Image

Most of the graphics systems that run BIOGRAF provide a mouse, special function keys and dials which allow you to manipulate the image shown in the graphics window. These differ somewhat depending on the graphics system.

1.2.1 The Dials Unit

A dials unit is optional on Silicon Graphics IRIS 4-D workstations and Stardent's Titan Graphics Supercomputer. A dials unit is also available for the IBM RISC System/6000, but is currently not supported. (Henceforth, these systems will be referred to respectively as the IRIS, the Titan, and the RS6000. The Evans & Sutherland ESV and Digital Equipment's DEC 5000 will be referred to respectively as the ESV and the DEC 5000). The dials unit consists of eight dials which allow you to move an image within the molecular display area. These include six dials which are used to either rotate or translate the image along the x, y, or z axis, a zoom dial which expands or shrinks the image, and a depth cueing dial.

The orientation and position of the dials differs depending on the graphics system; diagrams of the different types of dials units are shown in Figure 1.2 below.

x rotate 1	x translate 5
y rotate 2	y translate 6
z rotate 3	z translate 7
zoom 4	depth 8

A) Dials unit for Silicon Graphics IRIS 4-D Workstations

x translate 5	y translate 6	z translate 7	depth 8
1 x rotate	2 y rotate	3 z rotate	4 zoom

B) Dials unit for Stardent's Titan Graphics Supercomputer

Figure 1.2: Dials units for different graphics systems

The dials are labeled and numbered. Note the position of dial 8 for your system. This dial has multiple functions. It is normally used for depth cueing. However, when animating a trajectory, dial 8 is used to cycle through the different conformers being displayed and on some systems it controls the scan speed during auto scan. It is also used to rotate a portion of a molecule about a designated bond. Dials 1 through 7 may also take on this function when multiple bonds are being rotated simultaneously.

The x, y, and z axes are defined as shown in Table 1.1.

x	⇔	horizontal	positive to right
y	⇔	vertical	positive up
z	⇔	out of the screen	positive toward you

Table 1.1: Definition of x, y, and z axes

The edges of the graphics display define the limits of x and y. The x, y, and z axes are fixed in their own frame of reference instead of on the image: *e.g.*, turning the y rotate dial always makes the image pirouette around the vertical axis of the graphics display regardless of its previous orientation.

When an image is depth cued, its intensity decreases with its apparent distance from the viewer, creating a three-dimensional effect. The parts of the molecule "closest" to you look the brightest and the parts farthest away look the faintest. On depth cued systems there are two *clipping planes* parallel to the display that define the allowed range of z. The front clipping plane is usually fixed, while the rear clipping plane can move. The intensity of the distant atoms is proportional to the separation between the rear and front clipping planes. The depth cueing dial controls the position of the rear clipping plane. Moving the rear clipping plane away from you causes distant parts of the molecule to become brighter; moving it toward you progressively dims the back of the molecule until only a thin slab can be seen.

Warning (Titan users only): We have observed that rotating the dials on the Titan during an operation that updates the display (*e.g.*, energy minimization, dynamics calculations, auto scan) sometimes causes the program to crash. It would be prudent to minimize the user of the dials while these operations are being performed.

1.2.2 The Mouse

The mouse may also be used to manipulate the image in the graphics window. The cursor must be in the graphics window, however. Table 1.2 summarizes the button and key presses and mouse movements required for each function.

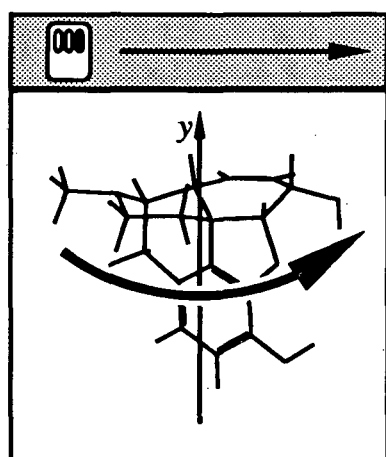
Depressing the middle mouse button and moving the mouse allows you to drag the image around the screen (this corresponds to the x and y translate functions). To translate the image along the z axis, the shift key and the middle mouse button must both be depressed while moving the mouse; movement up or to the right translates the image in the +z direction (toward the viewer), while moving the mouse down or to the left translates it in the -z direction (away from the viewer).

Similarly, depth cueing may be altered by depressing the control key and the middle mouse button; moving the mouse up or to the right dims the back portions of the image, while moving it down or to the left intensifies them. (When animating a dynamics trajectory, depressing the middle mouse button and control key and moving the mouse allows you to display earlier or later conformations, while when auto scanning, it allows you to change the scan speed.)

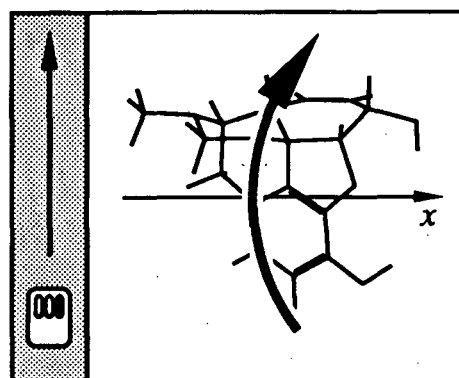
Function	Mouse Button	Key	Mouse Motion
x translate	middle	-	right /left
y translate	middle	-	up/down
z translate	middle	shift	up or right (toward viewer), down or left (away from viewer)
depth cueing	middle	control	up or right (dims distant objects), down or left (intensifies distant objects)
x rotate	right	-	up/down
y rotate	right	-	right /left
z rotate	right	shift	up or right (clockwise), down or left (counter-clockwise)
zoom	right	control	up or right (enlarges), down or left (reduces)

Table 1.2: Using the mouse to manipulate the image in the graphics window

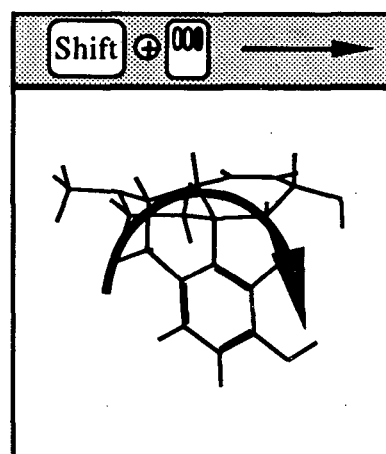
The right mouse button is used for the rotate and zoom functions (see Figure 1.3). Depressing the right mouse button and moving the mouse to the right or left rotates the image around the y axis, while moving the mouse up or down rotates the image around the x axis. For z rotations, the shift key must also be depressed. In this case, moving the mouse up or to the right rotates the image clockwise around the z axis, while moving the mouse down or to the left rotates it in the counter-clockwise direction. To resize the image, the control key must be depressed in conjunction with the right mouse button. Moving the mouse up or to the right enlarges the image, while movement down or to the left makes it smaller.



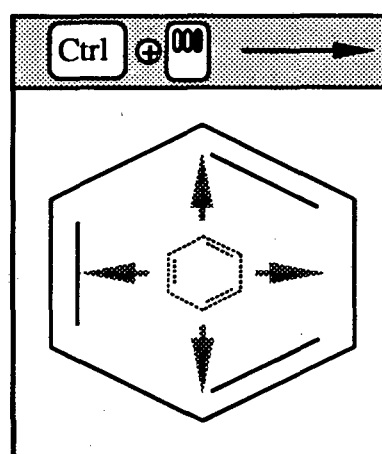
Y Rotate



X Rotate



Z Rotate



Zoom

Figure 1.3: Use of mouse for rotate and zoom functions

1.2.3 The Function Keys

Several function keys are provided which can be used to modify the display. These are function keys F1 - F5, and F11. They are described in Table 1.3 below. In addition, the program makes use of an abort key. This is used to stop any interactive molecular mechanics calculation after completing the current step or to abort charge equilibration calculations. On the Titan, F10 is the abort key, while on all the other machines, the F12 key is used. The user must have the cursor in the graphics display window to use the function keys.

Key	Description	Iris	Titan	RS6000	ESV	DEC 5000
F1	Removes menus, enlarges graphics window	Yes	Yes	Yes	Yes	Yes
F2	Displays two images for split screen stereo viewing	Yes	Yes	Yes	Yes	Yes
F3	Rotates right image displayed by F2 key by 90°	Yes	Yes	Yes	Yes	Yes
F4	Same as F1 but shows double image for 3-D viewing	Yes	Yes	Yes	Yes	Yes
F5	Activates hardware for stereo viewing (glasses or monitor)	Yes	Yes	No	No	No
F11	Brings text window forward	Yes	---	No	No	No
F10 or F12	Aborts molecular mechanics and charge equilibration calculations	Yes	Yes	Yes	Yes	Yes

Table 1.3: Function keys

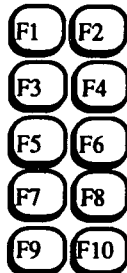
The position of the function keys on the keyboard differs depending on the graphics system. On the IRIS, RS6000, and ESV there are 12 function keys along the top of the keyboard, while the DEC 5000 has 14 function keys at the top of the keyboard. The Titan has 10 function keys in a keypad at the left of the keyboard (see diagrams A - C in Figure 1.4).



A) Function keys on Silicon Graphics IRIS 4-D, IBM RISC System/6000, and Evans & Sutherland ESV Workstations (top of keyboard)



B) Function keys on DEC 5000 Workstation (top of keyboard)



C) Function keys on Stardent's Titan Graphics Supercomputer (left side of keyboard)

Figure 1.4: Function keys for the different graphics systems

1.3 Viewing the Image in Three Dimensions

The program provides a STEREO function which renders the image in the molecular display area as a stereo pair so that it can be seen in three-dimensions. A trained user can easily see the three-dimensional image (he must either cross his eyes (proximal mode) or rotate them outward (distal mode)). The STEREO function is described in Section 4.6. Alternatively, special equipment may be purchased for some systems which allows the user to view images in three dimensions (without crossing his eyes). On the Titan, a special stereo monitor may be purchased, while on the IRIS, special glasses (Personal Stereo Viewer) may be used. On both of these systems, this stereo hardware is activated by pressing the F5 key. On the Titan, however, a special alias command must be set up first. This is described below.

1.3.1 Titan's Stereo Monitor

In order to enable the stereo monitor on the Titan, an alias command which will execute the stereo version of the program must be set up. This is done by editing a start up macro to include the stereo option, running COMSTRUCT with the desired alias command, and editing the defaults and .cshrc files. This procedure is described in the *Molecular Simulations Installation Guide*. This procedure needs to be performed only once. If you have a Titan with the stereo monitor, make sure this alias has been set up before proceeding.

A few more steps must also be done *after every log in* to get into stereo mode.

1. Enter the X-Windows environment by typing `xstart -stereo`
2. Type `xterm -display unix:0.1&`
3. Hold down ALT and press the F2 key to go into the stereo screen environment.
4. Type `setenv DISPLAY unix:0.1
awm&`
5. Run the program in stereo mode by typing the alias command set up previously (e.g., `biostereo`).
6. Use the F5 key to activate the stereo mode (F5 toggles you back and forth between stereo and normal modes). It may be necessary to reposition the window using Move Opaque in the X-Windows menu.

1.3.2 The Personal Stereo Viewer for the IRIS

The stereo glasses (Personal Stereo Viewer) for the IRIS are enabled by pressing the F5 key. Switches are provided on the glasses which turn them on and off, switch between stereo and normal modes, and change the phase. Consult the appropriate SGI manuals for further details. Please note to turn the glasses off when finished as they are battery-powered.

1.4 The Graphical User Interface (GUI)

The program employs an X-Windows-based graphical user interface (GUI), making it highly portable and capable of running on a wide range of platforms. Standard X-Windows type widgets are used including file browsers, dialog boxes, alert boxes, slider bars, and buttons. The program is basically menu-driven, easy to learn, and easy to use. The interface is designed to allow faster menu drawing and picking, less cursor motion, simpler menus, automatic file and group selection, and more visible messages. Commonly used functions are readily accessible via constantly displayed buttons. Buttons have also been provided which make it easy for the user to switch between displaying the text window and the graphics window. A file browser feature speeds up file input and makes it easy to change your default directory.

The screen is divided into five sections as shown in Figure 1.5:

- 1) A graphics window for displaying molecules (upper left)
- 2) A logo window (upper right)
- 3) A Modules/Messages window directly beneath the logo (middle right)
- 4) A Menu Pad window which contains the menus (lower right)
- 5) A window containing the GRAPHICS, TEXT, and PLOT buttons (lower left)

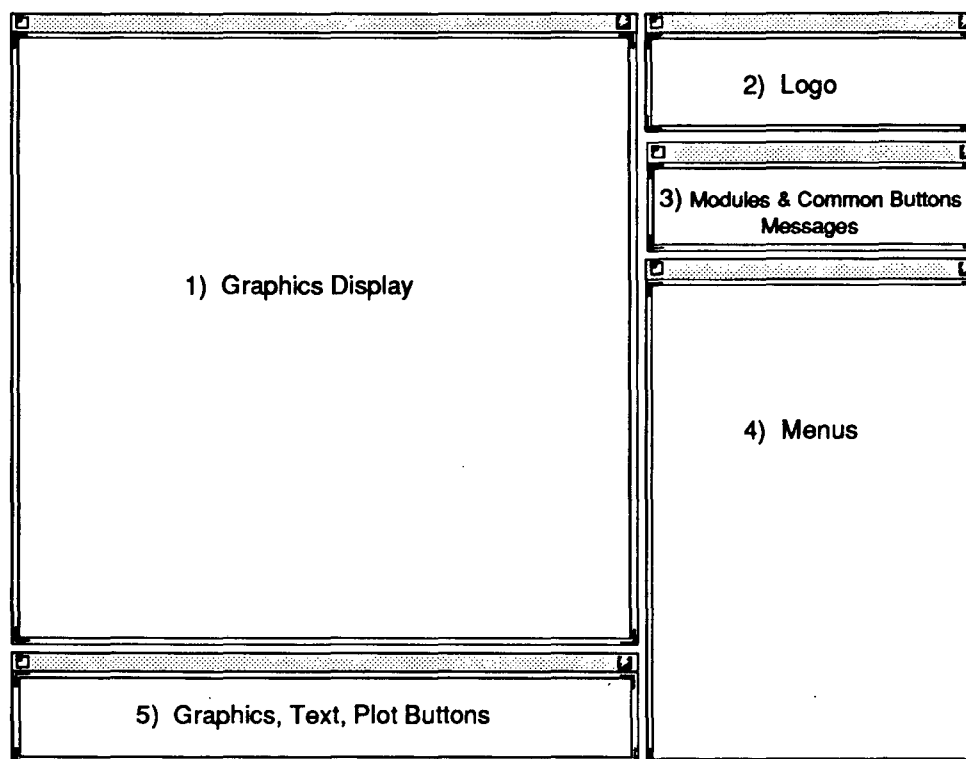


Figure 1.5: BIOGRAF screen layout

Each of these windows can be relocated, resized, or closed if desired. The system window (used to issue the start up command) in addition to displaying system messages also displays results and information for a number of program functions. It is referred to here as the text window and can be readily viewed by clicking the TEXT button.

1.4.1 The Graphics Window

The graphics window is used for the display and visualization of molecules. The cursor must be in the graphics window in order to manipulate molecules with the mouse and to activate the function keys.

1.4.2 The Modules/Messages Window

The Modules/Messages window includes a modules button (top) which is used to switch between the different Molecular Simulations products or modules that have been installed on the system (*e.g.*, BIOGRAF, POLYGRAF, NMRgraf, and Polaris). The name of the module currently being used is displayed. The user must be at the top level menu in order to switch (see TOP MENU below). Beneath the modules button are three buttons used to implement commonly used program functions. These are the DISPLAY, TOP MENU, and COMMAND buttons, described below. This window is also used to display prompts and messages. When this is required, the other buttons are overwritten with the message or with a dialog box requesting input from the user (dialog boxes are described in Section 1.5.2 below). Prompts and messages are used extensively by the program and are the primary means of communication with the user. The user should therefore check this window regularly.

1.4.3 The Menus

The menus are the means by which BIOGRAF accepts commands. The Functions menus are used to drive the program. They are always located at the left of the Menu Pad. Additional menus appear adjacent to the Functions menus which display additional operations or choices which are needed to execute a particular function. These additional menus appear only when input is needed from the user. They include the Files menu, which is used to display the names of the files which are currently in the system, the Groups menu, which displays the groups currently in the system, and other Selections menus displaying the choices available for the currently active function. The Simulation results menu is also displayed here during minimization and dynamics calculations; it lists several key energy-related values.

Operations are initiated by picking one of the functions from the top level Functions menu which appears at start up. When one of these functions is selected, the top level menu is replaced by a submenu displaying additional functions. These functions may also have their own submenus, *etc.* Thus, we have a Functions menu tree which consists of several levels. The top box of every Functions menu is always highlighted and contains the current menu's name. Every Functions menu also has a RETURN function at the bottom. Toggling the RETURN box restores the previous Functions menu, taking you up one level in the menu tree. A TOP MENU button is also readily accessible in the Modules/Messages window so that the user can immediately return to the top level menu if desired (see Section 1.4.4 below).

The use of the menus is described in Section 1.7 below.

1.4.4 The DISPLAY, TOP MENU, and COMMAND Buttons

These three buttons are frequently used and have been placed in the Modules/Messages window so that they are readily accessible. The DISPLAY button brings up the Display menu which lists the groups that are currently in the system. Groups are collections of atoms and bonds chosen in such a manner as to reduce a molecular structure into functional parts. Program functions can then be applied to selected groups of atoms. For example, groups can be rendered differently (*e.g.*, displayed in different colors) enhancing visualization. Groups are created using the MAKE GROUPS function described in Chapter 4. This ability to create groups is one of the most powerful features of the program. The user can turn a group "on" or "off" (have it displayed or not) by toggling on the group name in the Display menu. Throughout the figures in this manual, menu items that are "on" appear in **bold format**.

The TOP MENU button returns you immediately to the top level menu no matter where you are in the Functions menu tree. There may be instances when you have finished using an option on one of the submenus and want to return to the top level menu. Toggling the RETURN box in any Functions menu always restores the previous Functions menu and takes you up one level in the menu tree. If you are several submenus down in the tree, this must be done repeatedly in order to get back to the top level menu. Toggling TOP MENU speeds this process up and returns you to the top level immediately.

In addition to using the Functions menus to initiate operations, the program provides another mode of operation where the user may issue commands by entering text in the text window. This is called command mode. Command mode allows the use of some very powerful special commands and macros. Toggling the COMMAND button puts you in command mode. A command prompt will appear in the text window, indicating that you are in command mode. The use of command mode and macros is described in Appendix K.

1.4.5 The GRAPHICS, TEXT, and PLOT Buttons

The GRAPHICS, TEXT, and PLOT Buttons allow the user to easily switch between viewing the graphics window, text window, and plot windows. For example, when running minimizations or molecular dynamics calculations, information related to the energy calculations appears in the text window. This simulation information consists of several lines of text and is often more fruitfully viewed in a *large* text window. However, a large-sized text window might obscure visualization of the molecules in the graphics window. One would probably like to display a full-sized graphics window when building and visualizing molecules, but bring up a full-sized text window when running energy calculations. The GRAPHICS and TEXT buttons allow the user to quickly switch between these two alternatives. The PLOT button will similarly be used to bring up a window displaying plots of various functions. (The PLOT button feature is not available in this release.)

1.5 Communicating with the User

The program communicates with the user by printing messages in the Modules/Messages window and by bringing up dialog and alert boxes. Also for certain operations, results and other informative data are displayed in the text window. (This is the systems window that was used to start the program and where commands are entered when in command mode.)

The user communicates with the program in three ways: by clicking in menu boxes and buttons, by typing information or making selections in dialog boxes, and by picking atoms in the graphics window. When in command mode, the user communicates by entering commands in the text window.

1.5.1 Selecting Menu Boxes and Buttons and Picking Atoms on the Display

The user interacts with the program primarily by selecting menu boxes and buttons and by picking atoms on the graphics display with the mouse. All three of the mouse buttons may be used when selecting menu items and buttons, but only the left button is operative when picking atoms. It is important to note that the position of the mouse on the mousepad is irrelevant; it is the x and y movement of the mouse that determines the position of the cursor. When using the menus, each menu label appears bright when the cursor is positioned over it. The menu item is not selected however until the user clicks in the box with the mouse, causing it to be highlighted (*i.e.*, the background becomes bright). The user can "deselect" the menu item by clicking in the box again.

When the user is required to pick an atom, the program indicates this by surrounding the graphics window with a red border. Each picked atom is marked with a colored cross. When picking an atom, if two atoms are displayed in the same area (within the tolerance of the small box of the cursor), the atom appearing closer to the user may not be the one that is picked. This problem can usually be resolved by changing the orientation or the size of the object on the display so that only the desired atom lies within the area of the cursor box. If you are not able to successfully pick the desired atom, it may be due to another atom which is displayed in the same area, but which is not visible to the user because depth cueing causes it to be displayed with little or no intensity. Temporarily reduce the degree of depth cueing to see if additional atoms lie within your picking area.

Picking a Range of Atoms

There are instances when the user is required to pick a "range" of atoms. A range of atoms is defined as a subset of the total atoms displayed which are all connected to one another through a series of covalent bonds. For example, a molecule of n-butane displayed without hydrogens has a maximum range of four atoms. To pick a range of atoms from the display, the user must first pick an atom called the base atom. This is indicated with a red cross. At this point, the range comprises the base atom and all the atoms connected to it (these are designated with yellow crosses). Stopping atoms may then be picked to delimit the range (the atoms no longer included in the range will no longer be marked with yellow crosses). When the range is defined as desired, the user picks a non-atom. This causes the red border surrounding the graphics window to disappear. A menu then appears and the user has the option to continue picking atoms by selecting CONTINUE from the menu, or he can complete the operation by picking EXECUTE. Picking RETURN aborts the

operation. To specify the entire molecule as the range, the user simply picks the base atom (any atom), a non-atom, and EXECUTE. Stopping atoms may be "unpicked" to restore atoms to the range; this is done by simply picking the stopping atom again. Similarly, the user may start over completely by picking the base atom again.

Picking a List of Atoms

When picking a list of atoms, one picks the atoms individually (each one selected is marked with a colored cross). As in picking a range of atoms above, an atom may be "unpicked" (picked again) to remove it from the list. When the list is defined as desired, the user picks a non-atom, then picks EXECUTE to complete the operation (he can instead choose CONTINUE to continue picking atoms or pick RETURN to abort the procedure).

1.5.2 Dialog and Alert Boxes

Several program functions require the user to enter information via dialog boxes. An example is shown in Figure 1.6 below. The dialog box includes a prompt box at the top and a text entry box (highlighted) directly beneath it. The cursor is automatically positioned appropriately so the user doesn't need to move it (text can be input as long as the cursor is located anywhere within the dialog box). The user just types in the required information, then clicks the OK button. The OK button can be "clicked" using the mouse or by simply pressing the <Enter> key. The Cancel button is used to abort the function.

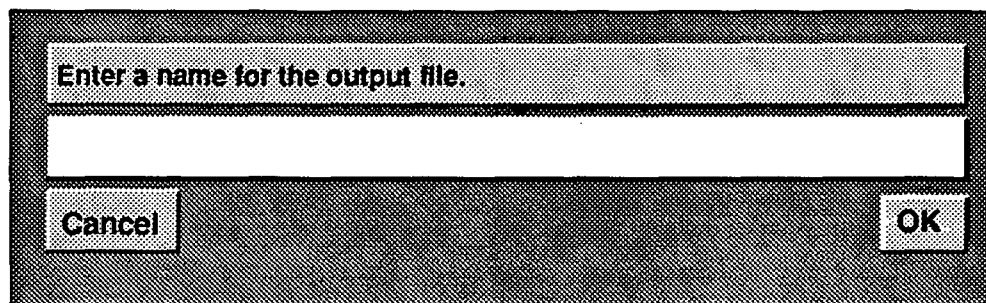


Figure 1.6: Dialog box example

The Alert box is similar to the dialog box in that it also contains a prompt or message and OK and Cancel buttons. It appears in the graphics window area, however, and is used for critical functions such as resetting or exiting the program.

1.6 Entering File Names

The user is required to specify file names when reading in and writing out files. When writing files, the user enters the file name directly in the dialog box which appears. If the file being saved is not in the current working directory, the full pathname including disk and directory specifications must be included (*"../"* may be used to indicate the previous directory). Logical names cannot be used. The program automatically adds on the appropriate three- or four-letter extension if not already included by the user.

When reading in files, a special file name entry dialog box appears (see Figure 1.7). The user may enter the file name directly (using the New Dir and New File buttons) or he may select the file from listings provided in the browser box. These are described below.

1.6.1 The Browser Box

A special file name entry dialog box is used for file input (see Figure 1.7). This dialog box is displayed when reading in files (*e.g.*, when using the READ option on the IN-OUT menu (see Section 2.1) and when using the BROWSER function to search through files or change the default directory (see Section 7.4 under UTILITIES). The full pathname for the current working directory is given at the top next to "Dir:", the full pathname for the currently selected file is shown next to "File:", and the subdirectories and files within the current working directory are listed in the browser box. The directories are always shown in red, while the files are shown in black and have a 3- or 4- letter extension. The previous directory (up one level) is always shown at the top of the list in parentheses (*"../"* indicates the previous level). The user can select this to move up in the directory path or he can select one of the subdirectories to move down. Selection is done by clicking on the desired directory or file name with the mouse (file names become highlighted when selected). Selections made in the browser box are also reflected by changes in the directory and file names shown at the top (the filename is not shown until a file has been selected). When reading in a file, the listings only include files of the filetype specified (*e.g.*, if Mopac was chosen as the filetype, only *.mpc* files are listed). If using the browser function on the Utilities menu, however, all filetypes are shown. The user can thus browse through the listings, changing directories if desired, until he finds the appropriate file. A final selection is not made however, until the user clicks the OK button.

The scroll bar at the right of the browser box can be used to scroll through the listings. The user must position the cursor within the scroll bar box, then click one of the mouse buttons. The left button scrolls the list up, the right button scrolls it down, and the middle button scrolls it to a particular position within the list (*e.g.*, to move to a position 1/2 way through the list, put the cursor in the middle of the scroll box and click the middle mouse button, or press the middle mouse button and drag the scroll bar to this spot). These features are particularly useful when working with long lists.

Note: The use of exceedingly long lists may cause delays (it takes longer for the browser box to display them). This may be avoided by keeping no more than 50 files in any directory.

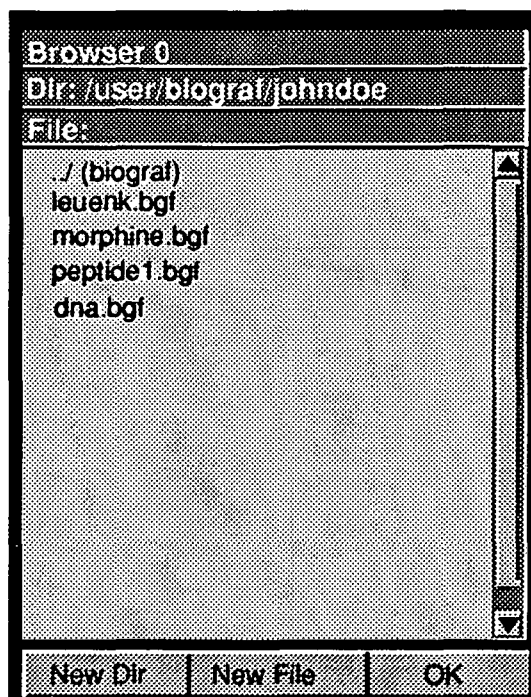


Figure 1.7: Filename Entry and Browser Box

1.6.2 The New Dir and New File Buttons

The user can also change the default directory or select a new file using the New Dir or New File buttons. Clicking these buttons with the mouse brings up a dialog box requesting the user to enter the name of the new directory or file desired (the New Dir dialog box also lists the current directory). If the new file or directory is not in your current working directory, the full pathname including disk and directory specifications must be included ("../" may be used to indicate the previous directory). The entry is verified by clicking the OK button (Cancel aborts the entry). The user must then click a second OK button (in the first dialog box shown in Figure 1.7) to implement the entry.

1.7 Using the Menus

The Functions menus located at the left of the Menu Pad are used to initiate operations. Other menus appearing adjacent to the Functions menus complete these functions by providing selections for the user. A number of conventions apply to the use of the menus. These are described below.

1.7.1 The Functions Menus

The items on each of the Functions menus do one of four things:

- 1) bring up a submenu displaying additional functions.
- 2) execute the function indicated by the menu item's name. This may be done immediately or after additional information is obtained from the user.
- 3) indicate a parameter value (usually in the box directly beneath the item). The parameter values are non-pickable menu items and are displayed in a different color so they may be easily recognized. Generally, toggling the menu box containing the parameter name brings up prompts which allow the user to change the associated value.
- 4) indicate the current state or mode which will be applied when another operation is performed. This includes items which are either in the "on" or "off" state where "on" is indicated by highlighting the menu item, and items which cycle through different possibilities where the one currently displayed is the one which is applied.

The IN-OUT menu provides some examples of the different types of menu items (see Figure 1.8).

Menu titles always appear at the top, end in a colon, and are partially highlighted (a light gray background is used). Picking them has no effect. In addition to the menu title at the top, we have two subtitles on this menu, READ DEFAULTS and WRITE DEFAULTS, indicating that the items beneath them are the read defaults and the write defaults. Subtitles are distinguishable from regular menu items since they are displayed in a different color (the default label color is blue) and always end with a colon. Picking subtitles also has no effect.

The CAPTION and PRINT SCREEN menu items are examples of options which bring up their own submenus. These submenus display additional options used in implementing the caption and printing functions.

The IN-OUT menu also has examples of on/off and cyclic menu items. Two of the read defaults (AUTO GROUP and HALF BOND) are on/off menu items that are currently in the "on" state, indicating that these options will be applied when the operation they refer to (the READ operation) is performed. This means that a group will automatically be created for the file to be read in and that the new molecule will be displayed with the half-bond colors (rather than in monochrome). Toggling these items turns them off (they no longer appear highlighted).

The third read default (**CENTER NEXT**) is an example of a cyclic menu item. It controls the centering of molecules on the display. Toggling this box cycles through three different centering methods: **CENTER NEXT**, **CENTER ALL**, and **NO CENTER**. The one which is currently shown in the menu box is the one which is applied when the read operation is executed. The default (**CENTER NEXT**) centers the next molecule (the one being read in) in the graphics window.

This menu also has an example of two mutually exclusive menu items. These are the write defaults, **BY FILE** and **BY GROUP**. Only one of these may be on (highlighted) at a time. Toggling the item which is off turns it on and at the same time turns the other item off.

The rest of the items on the IN-OUT menu are all functional, that is, picking them causes an operation to be performed. Most of them, however, require further input from the user before the operation may be completed. The program employs certain conventions in using the menus which make this process easier. These are described in the next section.

Throughout this manual, all menu titles, subtitles, and menu items in the "on" state appear on the menu in **bold**. Also, uppercase is used when referring to a menu item (*e.g.*, To abort the function, pick **RETURN**).

In-Out:
read
write
copy file
caption
print screen
Read defaults:
auto group
half bond
center next
Write defaults:
by file
by group
return

Figure 1.8:
IN-OUT menu

1.7.2 Basic Conventions for Using Menus

The program employs the following conventions for using the menus:

- 1) Once a function has been selected from the Functions menu, additional information may be needed in order to complete the operation. The program requests input in one of three ways:
 - (a) It displays another menu with a list of selections. A prompt may appear in the Modules/Messages window. If still visible, the original menu is dimmed indicating that it is currently inactive (picking menu items has no effect).
 - (b) It surrounds the graphics display window with a red border and prints a message in the Modules/Messages window asking you to perform an operation within the graphics window.
 - (c) It displays a dialog box requesting you to make a selection or enter information. The cursor is automatically positioned in the dialog box.

- 2) If the program requires you to make only one selection from a menu, then the operation continues as soon as you pick the item. When it is possible to make more than one selection, however, you can select and deselect items until you have picked the set of items you want. The operation is then completed by picking RETURN.
- 3) Most item selection operations can be aborted by picking RETURN or ABORT.
- 4) The program can indicate completion of an operation in several ways: by restoring the Modules and Common Buttons (if there was a message in the Modules/Messages Window), by removing the red border surrounding the graphics window (if picking atoms), and by restoring the menus and/or dialog boxes to their previous condition.

1.8 Running the Program in the Background

In order to run the program in the background, the user must 1) specify that graphics are not used and 2) create a macro to drive the program. The user specifies that the program be run without graphics by including the line `L_BATCH YES` in the defaults file (e.g., *biomacro.def*). This can be done by either editing an existing *.def* file or by creating a new one using COMSTRUCT. If running COMSTRUCT, NONE is chosen for the graphics option. The default alias produced by COMSTRUCT for BIOGRAF will be *biomacro*. For more details consult the COMSTRUCT chapter in the *Molecular Simulations Installation Guide*.

The user typically runs the program interactively to generate the macro that will be used to drive the program in background mode. By default, the program creates a listing of all user input (menu picks, atom picks, text entry *etc.*) in a log file called *logfile.macro*. This macro file should be appropriately renamed (or it will be overwritten). It may also need to be edited. For example, certain parameters such as the number of steps used for energy calculations might be changed or the macro may be embellished to meet the user's needs. See Appendix K for more details on macros.

A UNIX command is then issued to run the program driven by the macro. For example the command *biomacro name.macro >& project.out&* might be used where *biomacro* is the alias, *name.macro* is the name of the macro that was generated, and *project.out* is the name of the output file used to contain the results of the program, if any.

1.9 Basic Operations

The basic operations displayed on the top level Functions menu are listed below with a very brief description of their use. Detailed descriptions may be found in the subsequent chapters.

In-Out	Transfer files to and from disk, copy files, add captions, arrows, borders <i>etc.</i> to the graphics display, generate PostScript files from the images on the screen for output to a printer.
Build	Sketch new structures and build polypeptides, lipids, DNA, RNA, and carbohydrates from monomeric units and organic molecules from organic fragments, build crystals from crystallographic data, and solvate molecules. Additional functions are used to rotate bonds, dock structures and interactively monitor interatomic distances, force vectors, and energy, add or delete hydrogens or modify their mass, calculate moments for a structure, convert coordinate files, and generate connectivity for input files which don't have bonding information, .
Visualize	Partition molecules into functionally useful subsets of atoms (make groups), change rendering mode (vector, solid, cylinder, ball and stick), create ribbons, calculate and display dotted surfaces that show the structure's contact surface, electrical charge, or electrostatic potential, change the center of rotation of the display, change the color of structures, display images in stereo, cause objects to rock or rotate about the y axis.
Simulate	Perform energy minimizations, conformational searches and molecular dynamics, animate the sequences generated, perform statistical analysis on them, specify the parameters that govern the energy calculations, and modify the force field parameters.
Analyze	Identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, check for "bad" van der Waals contacts, measure molecular volumes, check proteins for irregular structural features, match structures by least-squares fits, analyze and plot trajectory file data, display, merge, and change the appearance of plot files.
Utilities	Reset the program, call up user-defined functions, browse through files in the system, change the default directory, and change the color of the labels.
Exit	Exit the program.

Chapter 2

In-Out Operations

The operations on the IN-OUT menu shown in Figure 2.1 allow the user to READ and WRITE various files from and to the computer's file system, to copy files (COPY FILE), to add captions, borders, boxes, *etc.* to a displayed structure (CAPTION) and to create a PostScript file of the images on the screen for output to a printer (PRINT SCREEN). Files that may be read in and/or written out include coordinate files in BioDesign General Format as well as in Brookhaven, Cambridge, CSSR, Big Strain 2, CHEMLAB CHM, CHEMLAB MOL, Macromodel, Moledit, and MOPAC formats. Other file types include force field parameter files, group files, coordinate transformation files (*i.e.*, rotation and translation matrices), dot files (*e.g.*, solvent-accessible surfaces), and vector files (*e.g.*, electron density contours). Appendix D.4 lists all of the standard three-and four-character extensions used in the file names that accompany the program and appear in the documentation. Section 1.6 explains how to enter file names in a format that can be read by the program.

Three read defaults are provided which allow the user to automatically create a group for the structure being read in (AUTO GROUP), display it in the half-bond colors (HALF BOND), and center it on the screen (CENTER NEXT). Write defaults are also provided for saving information BY FILE or BY GROUP (only BioDesign General Format files may be saved BY GROUP).

In-Out:
read
write
copy file
caption
print screen
Read defaults:
auto group
half bond
center next
Write defaults:
by file
by group
return

Figure 2.1:
IN-OUT menu

2.1 READ

The READ operation allows the user to read in various files. When reading in coordinate files, the user should first set the READ defaults to the desired settings. As shown in Figure 2.1, options are available to automatically create a group for the file that is being read in (AUTO GROUP), to display the structure in a single color or with the half-bond colors (HALF BOND), and to center it on the screen (CENTER NEXT). The default settings have the AUTO GROUP and HALF BOND toggles on and the CENTER NEXT option chosen. These options are described in Sections 2.6 - 2.8.

After picking READ, the menu shown in Figure 2.2 appears displaying the types of files that may be input. The coordinate file types are listed first followed by five other file types (these are described in more detail below). The user selects the type of file to be input from the menu and then enters the name of the file in the file name entry dialog box. This dialog box includes a browser box which provides listings of the current directories and files. The user may either select the desired file from the listings or he may type in the file name directly via the New Dir and New File buttons. (See Section 1.6 for details on entering file names).

The message box will say *Working...* while the file is being read. Upon completion, the molecule is displayed in the graphics window. If the error message *"error opening input file..."* appears, the file name, its directory location, or its syntax was incorrect.

Note: When entering file names directly, if the three- or four-letter extension is not included in the file name, the program first searches for the file name alone, then it searches for the file name including the appropriate extension. For example, if the user chooses the BioDesign General Format file type and just enters *test* for the file name, the program will find the file *test.bgf*.

User Input:

1. Coordinate file types only: set the READ defaults to the desired settings (see AUTO GROUP, HALF BOND, and CENTER NEXT in Sections 2.6 - 2.8 below).
2. Pick READ from the IN-OUT menu.
3. Pick the type of file to be read from the File types menu.
4. Select the desired directory and file name from the listings in the file name entry dialog box or use the New Dir and New File buttons and type in the name directly (see Section 1.6 for details).
5. Verify the selection by pressing <Enter> or by clicking the OK button.

File types:
BioDesign
Brookhaven
Cambridge
Chemlab Chm
Macromodel
Moledit
Molfile
Mopac
user
FF parameters
group
rotate/translate
dot
vector
return

Figure 2.2:
File types menu
(READ)

2.1.1 BIODESIGN

BioDesign General Format files contain atomic coordinate and connectivity information as well as force field information. The format is similar to that of Brookhaven Protein Data Bank files. Appendix D describes and contrasts the BioDesign General Format with the Brookhaven file format. The BioDesign General Format is recommended for internal use with the program since it contains all pertinent information.

To read in a BioDesign General Format file, follow the procedure described in Section 2.1 above, picking BIODESIGN for the file type. The standard three-letter extension used for BioDesign General Format file names is *bgf*. Periodic files created by the program and saved using the *xtl* extension may also be read in using BIODESIGN for the file type.

Note: If you are planning to run energy calculations on RNA, DNA or peptide structures using the Molecular Simulations implementation of the AMBER force field, then additional steps must be taken after reading in the BioDesign General Format file (see Appendix O). These procedures produce structures with atom types and charges that allow usable energy expressions to be set up.

2.1.2 BROOKHAVEN

To read a Brookhaven Protein Data Bank file stored on disk, follow the procedure described in Section 2.1 above, picking BROOKHAVEN for the file type. The standard three-letter extension used for Brookhaven file names is *bkv*. Files ending with the *pdb* extension (Protein Data Bank) may also be read in.

The error message "*no implicit connectivity generated for residue xxx*" indicates that no implicit topology has been defined for residue xxx. This will generally appear for nonstandard residue types. If the input file does not contain explicit connectivity for such residues, then the topology must be defined before display and energy optimization of this residue. This can be done using the CONNECT functions described in Section 3.14 and the EDIT operations on any of the Builder menus (see Section 3.1.12).

The error message "*residue xxx, atom yyyy not found in conversion table*" indicates that no match was found for the given residue and atom labels in the conversion table on unit 21 used to assign atom types and force field parameters. (Default assignment is *bg_exe:dreidii300.cnv*. See Appendix I.2 for details.) If no match is found, an atom type is assigned based on the atom label. Since this may not be correct, the atom types should be displayed using ATOM TYPE under the EDIT ID function on the Builder menus (Section 3.1.13) before any energy calculations are performed.

Note: If you are planning to run energy calculations on RNA, DNA or peptide structures using the Molecular Simulations implementation of the AMBER force field, then additional steps must be taken after reading in the Brookhaven file (see Appendix O). These procedures produce structures with atom types and charges that allow usable energy expressions to be set up.

2.1.3 CAMBRIDGE

To read a Cambridge crystallographic data base file stored on disk, follow the procedure described in Section 2.1 above, picking CAMBRIDGE for the file type. The standard three-letter extension used for Cambridge file names is *cmb*. The Cambridge file type readable by the program is the "FDAT" file (formatted data file).

2.1.4 CHEMLAB CHM

To read a CHEMLAB CHM file stored on disk, follow the procedure described in Section 2.1 above, picking CHEMLAB CHM for the file type. The standard three-letter extension used for CHEMLAB CHM file names is *chm*.

2.1.5 MACROMODEL

To read a Macromodel file stored on disk, follow the procedure described in Section 2.1 above, picking MACROMODEL for the file type. The standard three-letter extension used for Macromodel file names is *mcm*.

2.1.6 MOLEDIT

To read a Moledit file stored on disk, follow the procedure described in Section 2.1 above, picking MOLEDIT for the file type. The standard three-letter extension used for Moledit file names is *mle*.

2.1.7 MOLFILE

To read a CHEMLAB MOL file stored on disk, follow the procedure described in Section 2.1 above, picking MOLFILE for the file type. The standard three-letter extension used for CHEMLAB MOL file names is *mol*.

2.1.8 MOPAC

The coordinates and atomic charges from a MOPAC file can be used to update an existing file. To read a MOPAC file stored on disk, pick READ, then pick MOPAC for the file type from the adjacent menu. A list of file names is displayed in the Files menu and the prompt *Pick the file to be updated* appears in the status box. After a file has been picked, the user is asked whether the coordinate and charge arrays should be updated, before being asked to enter the MOPAC data file name. The standard three-letter extension used for MOPAC file names is *mpc*.

User Input:

1. Set the READ defaults to the desired settings (see AUTO GROUP, HALF BOND, and CENTER NEXT in Sections 2.6 - 2.8 below).
2. Pick READ from the IN-OUT menu.
3. Pick MOPAC for the file type from the File types menu.
4. Pick the file to be updated from the Files menu.

5. Pick YES or NO to update coordinate arrays.
6. Pick YES or NO to update charge arrays.
7. Enter the MOPAC data file name in the dialog box. Verify the selection by pressing <Enter> or by clicking the OK button. Cancel aborts the operation. (See Section 1.6 for details on entering file names.)

2.1.9 USER

The USER option executes user-written routines. This allows the user to read in coordinate files stored in special formats. Contact Molecular Simulations Customer Support for more information.

2.1.10 FF PARAMETERS

To read a force field parameter file stored on disk, follow the procedure described in Section 2.1 above, picking FF PARAMETERS for the file type. The standard three-letter extension used for force field parameter file names is *par*.

2.1.11 GROUP

To read a Molecular Simulations group file stored on disk, follow the procedure described in Section 2.1 above, picking GROUP for the file type. The standard three-letter extension used for group file names is *grp*. For more information on group files see Section 4.1.

Note: Group files may be read in either binary or ASCII format.

Warning: The appropriate coordinate files (*e.g.*, *bgr* or *bkv* files) must be read before the group file. If multiple coordinate files were used, they must be read in the same order as when the group file was created. Reading the group file resets the Groups menu; all the old groups are replaced by the groups that are being read in.

2.1.12 ROTATE/TRANSLATE

To read a Molecular Simulations rotation/translation transformation file stored on disk, pick READ, then pick ROTATE/TRANSLATE for the file type from the adjacent menu. The name of the rotation/translation file must then be entered in the dialog box. The standard three-letter extension used for rotation/translation transformation file names is *rot*. The display will be updated to reflect the rotations and translations read from the *rot* file. For more information on transformation files see Section 3.10.

Warning: The appropriate coordinate files (*e.g.*, *bgr* or *bkv* files) and group (*grp*) file must be read before the rotation/translation transformation file.

User Input:

1. Pick READ.
2. Pick ROTATE/TRANSLATE for the file type from the File types menu.

3. Enter the name of the rotation/translation file in the file name entry dialog box. Verify the entry by pressing <Enter> or clicking the OK button. (See Section 1.6 for details on entering file names.)

2.1.13 DOT

To read a Molecular Simulations dotted-surface file stored on disk, follow the procedure described below. The standard three-letter extension used for dotted-surface file names is *dot*. For more information on the procedure for reading *dot* files see SURFACE in Section 4.3.

User Input:

1. Pick READ from the IN-OUT menu.
2. Pick DOT for the file type from the File types menu.
3. If the dot surface is to be displayed, pick YES; otherwise, pick NO.
4. Pick the group to be created from the Groups menu.
5. If you are adding a group, enter the name of the new group in the dialog box. Verify the entry by clicking the OK button (click Cancel to abort the operation).
6. If you want to display the calculated property, select YES; otherwise, pick NO.
7. If you want to group the surface, select YES, then pick the groups to be considered from the Groups menu; otherwise, select NO.
8. Enter the dot file name in the file name entry dialog box. Verify the selection by pressing <Enter> or by clicking the OK button. (See Section 1.6 for details on entering file names.)

2.1.14 VECTOR

To read a vector file stored on disk, pick READ, pick VECTOR from the adjacent menu for the file type, then pick the group to be created from the Groups menu. The program then asks for the name of the vector data file to be read in and the name of the group to be created. The standard three-letter extension used for vector file names is *vec*.

User Input:

1. Pick READ.
2. Pick VECTOR for the file type from the adjacent menu.
3. Pick the group to be created from the Groups menu.
4. Enter the vector file name in the file name entry dialog box. Verify the selection by clicking the OK button. (See Section 1.6 for details on entering file names.)
5. Enter the name of the group to be created in the dialog box. Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.2 WRITE

The WRITE operation allows the user to save files. This is generally done using the BY FILE option listed under the write defaults on the IN-OUT menu. However, for BioDesign General Format files, a BY GROUP option is also provided which allows the user to select specific groups to be saved from the Groups menu. This is especially useful for extracting pieces from a very large file. The BY FILE or BY GROUP option must be toggled on before picking WRITE.

When WRITE is picked, the menu shown in Figure 2.3 appears, displaying the types of files that may be output. The coordinate file types are listed first, followed by three other file types. The procedure for saving each of these file types is described in Sections 2.2.1 - 2.2.13 below. All require the user to enter a name for the output file, including the directory information if needed. The program automatically adds on the appropriate three- or four-letter extension at the end of the file name if the user does not include it. For example, if the user entered *data/test* as the file name for a BioDesign General Format file, the program would add on *.bgf* and the filename would become */data/test.bgf*. This automatic file name extension feature can be overridden by turning off the L_AUTOEXT switch (see the Defaults file description in Appendix D.9). The standard file name extensions are given in the descriptions for each of the file types below and are listed in Appendix D.4.

2.2.1 BIODESIGN

BioDesign General Format files contain atomic coordinate and connectivity information as well as force-field information. The user may either write out existing files using the BY FILE default, or create a new file based on groups using BY GROUP. After choosing the appropriate default, the user picks WRITE, then selects the BioDesign General Format file type (BIODESIGN) from the adjacent menu. If BY FILE is being used, the files to be included are also selected from the Files menu (multiple files may be selected). All those files toggled will be combined in the output. This is useful for combining a number of small files into a single file.

If BY GROUP is being used, the groups to be included are selected from the Groups menu. The union of all the atoms in all the groups selected will be combined in the output. This option is especially useful for extracting pieces from a very large file.

The program then asks for the name of the output file and for remarks. The remarks are placed at the start of the output file. The standard three-letter extension for BioDesign General Format file names is *bgf*.

File types:
BioDesign
Brookhaven
CSSR
Chemlab Chm
Macromodel
Moledit
Molfile
Mopac
Big Strain 2
user
FF parameters
group
rotate/translate
return

Figure 2.3:
File types menu
(WRITE)

User Input:

1. Toggle on BY FILE or BY GROUP under the write defaults on the IN-OUT menu.
2. Pick WRITE.
3. Pick BIODESIGN for the file type from the adjacent menu.
4. BY FILE: Pick the file(s) to be output from the Files menu. Terminate by picking RETURN. (This step is skipped if there is only one file in the system).
BY GROUP: Pick the groups to be output from the Groups menu. Terminate by picking RETURN.
5. Enter a name for the output file in the dialog box. Include the directory information if needed (*bgf* is used for the extension). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).
6. BY GROUP: Enter a descriptor for the file (up to 14 characters). Terminate by pressing <Enter> or clicking the OK button.
7. Enter remarks. Terminate by pressing <Enter> or clicking the OK button.

2.2.2 BROOKHAVEN

To write a Brookhaven coordinate file, the basic procedure outlined in Section 2.2.1 is followed except that the BY FILE default must be used, BROOKHAVEN is chosen for the file type, and only one file can be written. See Appendix D for more information on the format of Brookhaven files written by the program. The standard three-letter extension used for Brookhaven file names is *bkv*.

User Input:

1. Toggle on BY FILE under the write defaults on the IN-OUT menu.
2. Pick WRITE.
3. Pick BROOKHAVEN for the file type from the adjacent menu.
4. Pick the file to be output from the Files menu. Terminate by picking RETURN. (This step is skipped if there is only one file in the system).
5. Enter a name for the output file in the dialog box. Include the directory information if needed (*bkv* is used for the extension). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).
6. Enter a header for the output file. Terminate by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).
7. Enter remarks. Terminate by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.2.3 CSSR

To write a CSSR (Cambridge Structure Search and Retrieval) coordinate file, follow the procedure given in Section 2.2.2, picking CSSR for the file type. No header or remarks are required. CSSR originated as the file format used by the S. E. R. C. Daresbury Labs; this format is used to

transfer data into the CERIUS or Chem-X¹ programs. (CERIUS can also read BioDesign General Format files.) The standard four-letter extension used for CSSR files is *cssr*.

2.2.4 CHEMLAB CHM

To write a CHEMLAB CHM file, follow the procedure given in Section 2.2.2, picking CHEMLAB CHM for the file type. No header or remarks are requested. The standard three-letter extension used for CHEMLAB CHM file names is *chm*.

2.2.5 MACROMODEL

To write a Macromodel file, follow the procedure given in Section 2.2.2, picking MACROMODEL for the file type. No header is required, and instead of entering remarks, the user is asked to enter a title for the output file. The standard three-letter extension used for Macromodel file names is *mcm*.

2.2.6 MOLEDIT

To write a MOLEDIT file, follow the procedure given in Section 2.2.2, picking MOLEDIT for the file type. No header is required, and instead of entering remarks, the user is asked to enter a title for the output file. The title is used for the group name when the file is read back in. The standard three-letter extension used for MOLEDIT file names is *mle*.

2.2.7 MOLFILE

To write a CHEMLAB MOL file, follow the procedure given in Section 2.2.2, picking MOLFILE for the file type. No header or remarks are required. The standard three-letter extension used for CHEMLAB MOL file names is *mol*.

2.2.8 MOPAC

To write a MOPAC coordinate file, the procedure given in Section 2.2.2 is followed except that MOPAC is chosen for the file type and the program asks whether the user wants to minimize the geometry with MOPAC. This creates a command file that will run MOPAC. The standard three-letter extension used for MOPAC file names is *mpc*.

User Input:

1. Toggle on BY FILE under the write defaults on the IN-OUT menu.
2. Pick WRITE.
3. Pick MOPAC for the file type from the adjacent menu.
4. Pick the file to be output from the Files menu. (This step is skipped if there is only one file in the system).
5. Pick YES or NO to perform geometry optimization or not.

¹Chem-X is a product of Chemical Design Ltd.

6. Enter a name for the output file in the dialog box. Include the directory information if needed (*mpc* is used for the extension). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.2.9 BIG STRAIN 2

To write a file for the BIG STRAIN program, follow the procedure described below, choosing BIG STRAIN 2 for the file type. The program also asks the user to choose the minimization method, the type of force field and the time limit for the BIG STRAIN calculation. The standard extension for a BIG STRAIN 2 file is *bgs*.

User Input:

1. Toggle on BY FILE under the write defaults on the IN-OUT menu.
2. Pick WRITE.
3. Pick BIG STRAIN 2 from the File types menu.
4. Pick the file to be output from the Files menu. (This step is skipped if there is only one file in the system).
5. Pick the minimization method.
6. Pick the type of force field.
7. Enter the time limit in seconds for the calculation.
8. Enter a name for the output file in the dialog box. Include the directory information if needed (*bgs* is used for the extension). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.2.10 USER

The USER option executes user-written routines. This allows the user to save coordinate files in their own special formats. Contact Molecular Simulations Customer Support for more information.

2.2.11 FF PARAMETERS

To write a parameter file, follow the procedure given in Section 2.2.2, picking FF PARAMETERS for the file type. No header or remarks are required. The standard three-letter extension used for parameter file names is *par*.

2.2.12 GROUP

Group information may be saved in a group file. All the groups currently in the system are saved. To write a group file, pick WRITE, then select GROUP for the file type. After entering a name for the output file, the user is asked to specify whether a type 1 or type 2 group file is to be saved. The differences between type 1 and type 2 group information are described below:

Type 1 - Saves all the sequences of toggles that defined the various groups. When type 1 group files are read, the program regenerates the groups based on the saved sequences of toggles. This allows the same group file to work for different coordinate files.

Type 2 - Saves the list of atoms and connections for each group as well as the sequence of toggles used to define the various groups. Type 2 group files regenerate groups faster than type 1 because the groups are regenerated using the atom lists, not the toggles used to define the group. Type 2 group files, however, should only be used for coordinate files with the same set of atoms and connections. Groups created with distance criteria (see Section 4.1.6) may not refer to the current coordinate set, but will provide the same connections as in the original file. This feature may be desirable when performing energy calculations.

Group files are saved in ASCII format. The standard extension used for group file names is *grp*.

Warning: In order to read a group file, the appropriate coordinate file(s) must also be saved.

User Input:

1. Pick WRITE.
2. Pick GROUP for the file type from the adjacent menu.
3. Enter a name for the groupings file in the dialog box. Include the directory information if needed (*grp* is used for the extension). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).
4. Enter the output file type (1 or 2).
5. Enter remarks. Terminate by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.2.13 ROTATE/TRANSLATE

To write a Molecular Simulations rotation/translation transformation file, pick WRITE and select ROTATE/TRANSLATE for the file type. The name of the file to be saved is then entered in the File name entry dialog box. The standard three-letter extension used for rotation/translation file names is *rot*.

User Input:

1. Pick WRITE.
2. Pick ROTATE/TRANSLATE for the file type from the File types menu.
3. Enter the name of the rotation/translation file. Include the directory information if needed (*rot* is used for the extension). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.3 COPY FILE

The COPY FILE option is used to make an internal copy of a file. The AUTO GROUP and HALF BOND defaults apply to copied files, but the centering functions (CENTER NEXT *etc.*) do not (the current transformations are used). Before picking COPY FILE, the user should first set these defaults to the desired settings (see Sections 2.6 and 2.7 below). The user picks COPY FILE, selects the file to be copied from the Files menu, and enters a descriptor for the new file. The new file name will appear in the Files menu when it is displayed next, and if AUTO GROUP was on, the copied structure appears on the screen.

User Input:

1. Set the AUTO GROUP and HALF BOND defaults as desired (see Sections 2.6 and 2.7).
2. Pick COPY FILE.
3. Pick the file to be copied from the Files menu. Terminate by selecting RETURN.
4. Enter a descriptor for the new file (up to 14 characters). Verify the entry by pressing <Enter> or clicking the OK button (click Cancel to abort the operation).

2.4 CAPTION

The operations on the CAPTION menu shown in Figure 2.4 allow the user to prepare a displayed structure for hardcopy reproduction. Captions (*e.g.*, titles or labels), arrows, borders, and boxes can be added to the images on the molecular display. Their color, location, size, and font can also be specified. The type of object to be added or changed is chosen using SELECT OBJECTS. Four types are available: CAPTIONS, BORDERS, ARROWS, and BOXES. The currently active type is displayed beneath SELECT OBJECTS on the menu. An associated menu is also displayed listing the objects available. The default object type is CAPTIONS, so initially the Captions menu also appears listing 20 possible captions that may be defined. Five functions may be performed on the active object type. The object must first be defined (DEFINE) before any of the other functions may be used. The user may then turn the object on or off (DISPLAY), change its color (COLOR), move it to a different position or resize it (ORIENT). If it is a caption, the font may also be changed (CHANGE FONT). The default font and color applied to the objects being defined may be changed using the DEFAULT FONT and DEFAULT COLOR options. When leaving the CAPTION menu, the user must choose to keep the objects displayed or have them removed from view.

User Input:

1. If the defaults for font or color are not appropriate, change them using the DEFAULT FONT and DEFAULT COLOR options (see Sections 2.4.7 and 2.4.8).
2. If not working with the current object type, use SELECT OBJECTS to change the object type and bring up the appropriate menu (Captions, Borders, Arrows, or Boxes).
3. Perform the desired function(s) using the DEFINE, DISPLAY, COLOR, ORIENT, or CHANGE FONT options (see Sections 2.4.2 - 2.4.6 below). (An object must be defined before any of the other functions can be performed on it).
4. Repeat steps 1 -3 as needed.
5. Pick RETURN to exit the CAPTION menu.
6. Pick KEEP DISPLAY to keep the objects displayed or pick REMOVE DISPLAY to have them removed from view.

Caption:
select objects
captions
Function:
define
display
color
orient
change font
default font
Times-Roman
default color
return

Figure 2.4:
CAPTION menu

2.4.1 SELECT OBJECTS

The **SELECT OBJECTS** option allows the user to change the type of object that is to be added or modified. The currently active object type is displayed in red beneath **SELECT OBJECTS** on the menu. The user picks **SELECT OBJECTS** and then chooses the new object type from the Object types menu shown in Figure 2.5. Four types are available: **CAPTIONS**, **BORDERS**, **ARROWS**, and **BOXES**. The user completes the selection by picking **RETURN**. The appropriate menu (Captions, Borders, Arrows, or Boxes) will then appear listing the objects available (see Figure 2.6 - 2.9). The **DEFINE**, **DISPLAY**, **COLOR**, **ORIENT**, and **CHANGE FONT** functions may then be performed (see Sections 2.4.2 - 2.4.6 below). (An object must be defined before any of the other functions can be performed on it).

Selections:
captions
borders
arrows
boxes
return

Figure 2.5: Object types menu

User Input:

1. Pick **SELECT OBJECT**.
2. Pick the new object type from the Object types menu (**CAPTIONS**, **BORDERS**, **ARROWS**, or **BOXES**). Pick **RETURN** to complete the selection.

Captions:
caption 1
caption 2

⋮

caption 20
return

Figure 2.6:
Captions menu

Borders:
border 1
border 2

⋮

border 20
return

Figure 2.7:
Borders menu

Arrows:
arrow 1
arrow 2

⋮

arrow 20
return

Figure 2.8
Arrows menu

Boxes:
box 1
box 2

⋮

box 20
return

Figure 2.9:
Boxes menu

2.4.2 DEFINE

The **DEFINE** option allows the user to define objects for the activated object type (**CAPTIONS**, **BORDERS**, **ARROWS**, **BOXES**). If **CAPTIONS** is active, **DEFINE** requires that the desired object be picked from the Captions menu and that a string of text be entered for that object (up to 80 characters). If any other type is active, definition only involves picking the desired object from the menu. As each object is defined, it is displayed on the screen and appears highlighted in the menu. An object must be defined before any of the other functions (**DISPLAY**, **COLOR**, **ORIENT**, or **CHANGE FONT**) can be performed on it.

User Input:

1. Pick DEFINE.
2. Pick the object to be defined from the Captions, Arrows, Borders, or Boxes menu.
3. If a caption was selected, enter a character string (up to 80 characters).
4. Repeat steps 2 - 3 to define additional objects.
5. Terminate the function by picking RETURN.

2.4.3 DISPLAY

The DISPLAY option allows the user to turn objects on or off (have them displayed or not) by toggling them on the currently active menu (Captions, Borders, Arrows, or Boxes).

User Input:

1. Pick DISPLAY.
2. Toggle the objects on the menu on or off.
3. Terminate the function by picking RETURN.

2.4.4 COLOR

The COLOR option allows the user to change the color of an object. After picking COLOR, the user selects the object whose color is to be changed from the Captions, Borders, Arrows or Boxes menu. He then picks the new color from those listed in the Colors menu as shown in Figure 2.10. If none of these colors is appropriate, the user may specify the exact color by picking RGB VALUE and using the sliders to specify values for the red, green and blue components, as well as for hue, saturation, and value (see Section 4.5 for more information on how to use the RGB VALUE option).

User Input:

1. Pick COLOR.
2. Pick the object whose color is to be changed from the Captions, Borders, Arrows or Boxes menu.
3. Pick the new color from the Colors menu or pick RGB VALUE and use the sliders to specify the exact color (see Section 4.5 for more information on RGB VALUE).
4. Pick RETURN to complete the selection.
5. Repeat steps 2-4 for another object or pick RETURN on the Captions, Borders, Arrows or Boxes menu to terminate the function.

Colors:
purple
magenta
pink
red
orange
yellow
chartreuse
green
aqua
cyan
light blue
blue
white
RGB value
return

Figure 2.10
Colors menu

2.4.5 ORIENT

The ORIENT option allows the user to move and scale an object on the display. (Borders cannot be oriented.) The object can be scaled in x and y and translated in x and y. The scale factors for these functions can also be adjusted. After toggling the object to be oriented from the menu, the dials and mouse take on new orient functions as shown in Table 2.1.

User Input:

1. Pick ORIENT.
2. Pick the object to be oriented from the Captions, Arrows or Boxes menu.
3. Move and scale the object using the dials or mouse.
4. Repeat steps 2-3 for another object or pick RETURN from the Captions, Arrows or Boxes menu to terminate the function.

Orient Function	Dials	Mouse
x translate	dial 5	middle button (left/right mvt)
y translate	dial 6	middle button (up/down mvt)
x/y translate scale factor	dial 8	middle button + control key
x scale	dial 1	right button (left/right mvt)
y scale	dial 2	right button (up/down mvt)
x/y scale scale factor	dial 4	right button + control key

Table 2.1: Use of dials and mouse for ORIENT functions

2.4.6 CHANGE FONT

The CHANGE FONT option allows the user to change the font for a selected caption (CAPTIONS must be the active object type). The possible fonts are displayed in the Fonts menu shown in Figure 2.11. Each time a different font is picked, the caption is displayed in the new font. The default font will still apply to subsequently created captions however.

User Input:

1. Pick CHANGE FONT.
2. Pick the caption whose font is to be changed from the Captions menu.
3. Pick the new font from those displayed in the Fonts menu. Pick RETURN to complete the selection.
4. Repeat steps 2 - 3 for another caption or pick RETURN on the Captions menu to terminate the operation.

Selections:	
Boston	International
Cyrillic	Iris
Courier	Screen
Courier-Bold	Screen-Bold
Courier-Bd-Obl	Symbol
Courier-Oblq	Symbol-Medium
Helvetica	Times-Bold
Helvetica-Bold	Times-Bd-Itlic
Hvica-Bd-Oblq	Times-Italic
Hvica-Medium	Times-Roman
Hvica-Oblq	Type
Icon	
	return

2.4.7 DEFAULT FONT

The DEFAULT FONT option allows the user to change the default font for newly-defined captions. The possible fonts are displayed in the Fonts menu shown in Figure 2.11. The default font is displayed beneath DEFAULT FONT on the CAPTION menu. Initially the default font is Times-Roman.

Figure 2.11: Fonts menu

User Input:

1. Pick DEFAULT FONT.
2. Pick the new default font from those listed in the Fonts menu.
3. Pick RETURN to terminate the operation.

2.4.8 DEFAULT COLOR

The DEFAULT COLOR option allows the user to change the default color for newly-defined captions, arrows, borders, and boxes. The possible colors are displayed in the Colors menu (see Figure 2.10). If none of these colors is appropriate, the user may specify the exact color by picking RGB VALUE and using the sliders to specify values for the red, green and blue components, as well as for hue, saturation, and value. Initially the default color is blue.

User Input:

1. Pick DEFAULT COLOR.

2. Pick the new default color from the Colors menu or pick RGB VALUE and use the sliders to specify the exact color (see Section 4.5 for more information on RGB VALUE).
3. Pick RETURN to terminate the operation.

2.5 PRINT SCREEN

The PRINT SCREEN menu shown in Figure 2.12 allows the user to generate an encapsulated PostScript file of the image currently on the screen so that it can be output to a printer or imported to another application. The file created may include information for the whole display area or a selected portion. The user can control the color of the image generated by specifying GRAY, BLACK&WHITE, or COLOR. GRAY (the default) uses 8 bits/pixel, BLACK&WHITE uses 1 bit/pixel, and COLOR uses 24 bits/pixel. The higher values give more shades of gray or more colors, but also take longer to generate. The image printed may be POSITIVE (colors are as shown on the screen) or NEGATIVE (colors are inverted). The CREATE PS FILE option is used to generate the PostScript file.

Print screen:
gray
negative
create PS file
return

Figure 2.12:
PRINT SCREEN
menu

2.5.1 GRAY/BLACK&WHITE/COLOR

The user can control the color of the image generated by specifying GRAY, BLACK&WHITE, or COLOR. GRAY uses 8 bits/pixel, BLACK&WHITE uses 1 bit/pixel, and COLOR uses 24 bits/pixel. The higher values give more shades of gray or more colors, but also take longer to generate. GRAY (8bits/pixel) is the default.

User Input:

1. Toggle the option until the desired color type is displayed (GRAY, BLACK&WHITE, or COLOR for 8, 1, or 24 bits/pixel respectively).

2.5.2 NEGATIVE/POSITIVE

This option allows the user to select whether the image printed will be POSITIVE (colors are as shown on the screen) or NEGATIVE (colors are inverted). The default is NEGATIVE. Typically, one would use the NEGATIVE option when printing a molecular structure (the screen image of the molecule is generally a light color against a black background and one would probably want the printout to be dark against a white background).

User Input:

1. Toggle the menu item until the desired printing mode is displayed (POSITIVE or NEGATIVE).

2.5.3 CREATE PS FILE

CREATE PS FILE allows the user to generate an encapsulated PostScript file of the images currently on the screen so that it may be output to a printer or imported into another application. After picking CREATE PS, the user enters a name for the PostScript file (the *ps* extension is automatically added). The file created may include information for the whole screen (including menus) or a selected portion. The user specifies the size and location of the rectangle to be included by picking

its upper left and lower right corners. The PostScript file generated may be printed at any time using the appropriate system print commands. For example, on some UNIX systems the command `lpr -Plw filename` would print the PostScript file specified by *filename* on the LaserWriter printer (`-Pprinter` specifies a particular printer and is optional; if omitted, the default printer is used).

User Input:

1. Set the color (bits/pixel) and printing mode (negative or positive) to the desired values.
2. Pick CREATE PS FILE.
3. Enter a name for the PostScript file.
4. Pick the upper left corner of the area to be included.
5. Pick the lower right corner of the area to be included.

2.6 AUTO GROUP

The AUTO GROUP option is an off/on toggle which allows the user to automatically create a group for the file to be read in or copied. The structure is treated as a single group which includes every atom and bond in the file. If the HALF BOND toggle is on, the group is displayed using the half-bond colors; otherwise, a single color is used (see HALF BOND, Section 2.7 below). The group appears in the Groups menu and is given the same name as the file descriptor. The AUTO GROUP option applies to coordinate files only. The default is on.

User Input:

1. Toggle HALF BOND on if the group is to be displayed using the half-bond colors; toggle it off if a single color is to be used.
2. Toggle AUTO GROUP on to automatically create a group; toggle it off if this is not wanted (see Section 4.1 for information on how to create your own groups).

2.7 HALF BOND

Groups may be displayed with bonds in all one color or in a half-bond representation. If HALF BOND is on when a group is created, each atom will be assigned a CPK color (note that carbons are aqua not black), and the color of each heteronuclear bond will change color halfway. If HALF BOND is off, then the bonds in the group will all be one color. That color can be changed using the CHANGE COLORS operation described in Section 4.5. HALF BOND is only applicable if AUTO GROUP is on (see Section 2.6 above). The default is on.

User Input:

1. Toggle HALF BOND on if the group is to be displayed using the half-bond colors; toggle it off if a single color is to be used.

2.8 **CENTER NEXT/CENTER ALL/NO CENTER**

This option is a 3-way toggle which allows the user to have the graphics display:

- 1) centered on the file being read in (CENTER NEXT)
- 2) centered on all the files in the system including the one being read in (CENTER ALL)
- 3) left as is (NO CENTER); the current atomic coordinates to display transformations are retained.

CENTER NEXT and CENTER ALL also scale the structure so that it fills the display. The centering functions are applied after the READ operation is executed (see Section 2.1 above); they are not available for the COPY FILE operation. The default is CENTER NEXT.

User Input:

1. Toggle this option until the desired centering method is displayed.

2.9 **BY FILE, BY GROUP**

BY FILE and BY GROUP are mutually exclusive toggles which apply when using the WRITE operation to save coordinate files (see Section 2.2). The default is BY FILE and must be used when writing out all coordinate file types except BioDesign General Format. BioDesign General Format files may also be written out BY GROUP. This option allows the user to select specific groups to be saved from the Groups menu and is especially useful for extracting pieces from a very large file. The option desired must be toggled on prior to picking WRITE.

User Input:

1. Pick BY FILE or BY GROUP (*bfg* files only) prior to picking WRITE.

Chapter 3

Build Operations

The Builders:

The operations on the BUILD menu shown in Figure 3.1 allow the user to sketch new structures and build structures based on fragment libraries which consist of directories of BioDesign General Format (*bgf*) files. Currently, there are Builders for organic molecules, peptides, DNA and RNA, lipids, carbohydrates, crystals, and solvents. The Builders are accessed using the ORGANIC, PEPTIDE, RNA-DNA, DNA-OLD, LIPID, CARBOHYDRATE, CRYSTAL, and SOLVATE options. Each of the Builders automatically reads its own set of fragment files in a special directory. These fragment libraries allow the user to build up a large structure by combining peptide residues, fatty acids, small organic molecules, *etc.*

The fragment library for organic molecules contains 28 organic fragments. These are located under the *bg_org_frg* directory. There are an additional 17 functional groups available in the *bg_org_fun* directory. The fragment libraries for peptides are located under the *bg_pep_l* and *bg_pep_d* directories. They contain the standard 20 amino acids in d- and l-form. The RNA-DNA Builder uses DNA nucleotides to build single, double and triple stranded chains. These are found in three fragment libraries, *bg_dna_single*, *bg_dna_double*, and *bg_dna_triple*. The single and double libraries contain the four standard DNA nucleotides and nucleotide pairs in both the A and B form, and the triple library contains seven polynucleotide triplets. *Note:* The old DNA Builder still functions as before, building double-stranded DNA based on the nucleotide sequence specified by the user.

The fragment libraries for lipids contain 4 common lipid backbone structures under the *bg_lipid_bck* directory, 4 head groups under the *bg_lipid_hg* directory, and 12 fatty acids under the *bg_lipid_fa* directory. The fragment library for carbohydrates contains 52 monosaccharides including the α and β conformations of D-glucose. The fragments contained in each of these libraries are listed in

Build:
organic
peptide
RNA-DNA
DNA old
lipid
carbohydrate
crystal
solvate
rotate bond
dock
modify H
moment
convert
connect
return

Figure 3.1:
BUILD menu

Appendix B. Fragment libraries can be augmented by user-defined *bgf* files. Separate user-defined libraries (or any of the libraries supplied by the program) can be used in any of the Builders.

The Crystal Builder functions somewhat differently in that crystallographic structures are built based on data entered by the user. A unit cell is constructed based on the data specified. This includes the space group, unit cell dimensions, and coordinates of the unique atoms. There are 230 space groups to choose from (these are listed in Appendix M). The atoms can be represented with symbols of various types, sizes and color. The unit cell may be extended to form a crystal and bonds may be formed across the cell boundaries. The Crystal Builder is only available for users who do not also have POLYGRAF. (POLYGRAF has a Crystal Builder, which is identical to the one provided here except that it also does energy calculations on periodic systems).

A number of functions are available on all the Builder menus which are useful in modifying the molecule after it has been created. These include the EDIT, EDIT ID, MECHANICS, SIMULATE and ANALYZE options. The EDIT functions are used to add and delete atoms, to alter connections between atoms and to modify the geometry of the system. EDIT ID is used to modify the labels, force field types, and charges on specified atoms. These options are described under the Organic Builder in Sections 3.1.12 and 3.1.13. The SIMULATE and MECHANICS operations can be used to modify the conformation of the molecule (see Chapter 5) and ANALYZE may be used to access all the analytical functions (see Chapter 6).

There are certain functions which are specific to each of the Builders. The Lipid Builder allows the user to extend the structure by replicating the molecule in the x and z directions to form a lipid monolayer or a lipid bilayer. In the Peptide Builder, the dihedral ϕ , ψ , and ω angles for peptides may be changed on a residue by residue basis using the MODIFY option. Up to four dihedral angles may also be changed for each side-chain. The Peptide Builder also has INSERT and DELETE options which allow the user to insert or delete amino acid residues into the peptide. The Carbohydrate Builder allows the user to extend the structure by replicating the molecule in one, two, or three directions and to modify the dihedral angles of the glycosidic bonds of all or part of the polysaccharide. The Solvent Builder may be used to build a solvent shell around a structure.

Other Build Options:

Additional options are also provided on the BUILD menu which serve to complete the building process. Functions are available which allow the user to rotate bonds (ROTATE BOND) and to dock two structures and interactively monitor interatomic distances, force vectors, and energy (DOCK). MODIFY H can be used to add missing hydrogens or delete hydrogens; one can also change the mass of hydrogens to that of deuterium or tritium in order to improve the stability of dynamics simulations. The MOMENT option allows the user to calculate the center of mass and principle axes of a structure and to transform the coordinates if desired. A CONVERT function is provided for converting coordinate files that do not include force field data, and CONNECT generates connectivity data for input files with no or incorrect bonding information.

3.1 ORGANIC

The operations on the ORGANIC menu shown in Figure 3.2 allow the user to construct organic molecules based on fragment libraries. The default fragment and functional group libraries are contained in two directories which are automatically read in when the user first enters the Organic Builder (a list of the fragments and their associated directories is given in Appendix B). The user accesses these fragments by selecting the BASE FRAGMENT and FUNCTIONAL GRP options. He then picks one of the fragments from the associated library which appears in the adjacent menu.

The Organic Builder can also read user-defined fragment libraries of BioDesign General Format (*bgf*) files as well as the other fragment libraries used by the program (*bg_pep_l* and *bg_pep_d*, the amino acid residue libraries for the Peptide Builder; *bg_carbo*, the carbohydrate fragment library; or *bg_lipid_bck*, *bg_lipid_fa*, *bg_lipid_hg*, the libraries for the Lipid Builder). These are accessed with the USER BASE FRAG option. The USER FUNCT GRP option accesses a user-defined functional groups library; each of the fragments in this library must have a hydrogen with the label XX (formatted as in the *bg_org_fun* library). This hydrogen is deleted when the functional group is condensed onto the existing molecule. *Note:* The user-defined directories must be read in with the READ LIBRARY option before attempting to access the fragments in these libraries.

When using the BASE FRAGMENT and USER BASE FRAG options, the selected fragment is loaded into the file and displayed in a pseudo-random location and orientation on the screen. Connections between fragments can be made by picking an atom on each substructure; the Builder automatically connects the fragments with appropriate bond lengths and angles. The dihedral angle between the joined fragments can be readily adjusted. When using the FUNCTIONAL GRP and USER FUNCT GRP options, the user is prompted to pick the hydrogen atom on the screen to indicate where the functional group is to be attached. The hydrogen atom picked will be replaced by the new functional group.

There are four operations which join fragments: BOND (accessed under EDIT), CONNECT FRAG, REPLACE and CONDENSE. The simplest is BOND where the user selects two atoms and the program draws a bond between them. The next simplest is CONNECT FRAG which forms a bond between two atoms, one selected from each fragment. Sometimes one or two atoms and associated bonds need to be deleted in order to form a bond between fragments. The REPLACE option allows the user to remove one atom and bond, while the CONDENSE option allows the user to remove two atoms and bonds, one on each structure. Since the fragments in the *bg_org_frg* directory include all hydrogens, they are generally

Organic:
setup
new atom ids
read library
base fragment
functional grp
user base frag
user funct grp
read fragment
copy
condense
replace
connect frag
draw
default atom
C
edit
edit id
mechanics
simulate
analyze
return

Figure 3.2:
ORGANIC menu

joined using the CONDENSE option, in which the user selects the two hydrogen atoms to be replaced by the new bond connecting the fragments.

The DRAW option uses the tablet to draw structures on the screen. The EDIT option allows the user to add and delete atoms, to alter connections between atoms, and to modify the geometry of the system. EDIT ID is used to modify the labels, force field types and charges on specified atoms, and to create user-defined monomers for use in the Polymer Builder. MECHANICS and SIMULATE may be used to modify the conformation of the structure, and ANALYZE provides analytical operations.

Upon toggling ORGANIC, the SETUP operation is automatically invoked and the user is prompted to specify the file to be used.

3.1.1 SETUP

The SETUP operation is automatically invoked each time the Organic Builder is entered. If there are no files in the system, the user merely enters a descriptor for the new file which will be created. A group of the same name is also created. If any files have already been input or created, then the Files menu appears and the user selects the file to be used from those listed in the menu. He may either modify one of the existing files or create a new one. If a new file is being created, the user chooses NEW FILE and then enters a descriptor for it.

User Input:

1. Pick SETUP (needed only if already in the Organic Builder and setting up again).
2. Pick the file on which to operate from the Files menu. Pick NEW FILE if creating a new file. This step is skipped if there are no files in the system.
3. Enter a descriptor (up to 8 characters) for the file if the file is new.

3.1.2 NEW ATOM IDS

The NEW ATOM IDS option allows the user to change the residue name, residue number and chain designator that will be used for the new atoms. The default values are listed in the box beneath the appropriate item on the New atom ids menu. An example is shown in Figure 3.3. (Defaults for a *new* file are ORG for residue name, 1 for residue number, and A for chain designator.) If the defaults are not acceptable, the user picks the item to be changed (RESIDUE NAME, RESIDUE NUMBER, or CHAIN DESIGNATOR) and enters a new value in the dialog box. When satisfied with all values, the user picks RETURN. ABORT aborts the operation. This information is used in writing Brookhaven or BioDesign General Format (*bgf*) files (see Appendix D) and can be used later in creating groups.

User Input:

1. Pick NEW ATOM IDS.
2. Pick the menu item to be changed from the New atom ids menu (RESIDUE NAME, RESIDUE NUMBER, or CHAIN DESIGNATOR).
3. Enter the new value in the dialog box. The residue name must be less than 3 characters, the residue number must be 0-9999, and the chain designator must be a capital letter or a blank (blank is indicated with an underscore).
4. When the defaults are acceptable, pick RETURN. (ABORT aborts the operation).

New atom ids:
residue name
ORG
residue number
1
chain designator
A
abort
return

Figure 3.3: New atom ids menu

3.1.3 READ LIBRARY

The READ LIBRARY option allows the user to read the directory information for any of four libraries selected from the Libraries menu shown in Figure 3.4: the base fragment library (BASE FRAGMENT), the functional group library (FUNCTIONAL GRP), a user-defined base fragment library (USER BASE FRAG), or a user-defined functional group library (USER FUNCT GRP). The directory information must be read before the options which access these libraries can be used (see Section 3.1.4). The base fragment library is contained in the *bg_org_frg* directory and the functional group library is contained in the *bg_org_fun* directory. These two directories are automatically read in when the user first enters the Organic Builder. The only time the user would need to read one of these default directories again would be if the user had added *bgf* fragment files to it during a session. The user-defined directories *must* be read in, however, in order to allow access to the user-defined libraries.

User Input:

1. Pick READ LIBRARY.

Selections:
base fragment
functional grp
user base frag
user funct grp
return

Figure 3.4: Libraries menu

2. Pick the library to be read in (BASE FRAGMENT, FUNCTIONAL GRP, USER BASE FRAG, or USER FUNCT GRP) from the adjacent menu.
3. Enter the name of the directory (for the user-defined fragment libraries). See Section 1.6 for details on entering file names.

3.1.4 BASE FRAGMENT, FUNCTIONAL GRP, USER BASE FRAG, USER FUNCT GRP

When one of these four options is toggled on, a new menu is displayed listing the fragments in the associated library in alphabetical order. BASE FRAGMENT brings up the base fragment library and FUNCTIONAL GRP brings up the functional group library (see Figures 3.5 and 3.6). A more complete description of the fragments contained in these two libraries is given in Appendix B.3 and B.4. USER BASE FRAG and USER FUNCT GRP bring up the user-defined base fragment and functional group libraries respectively (provided they have been read in using the READ LIBRARY option described in Section 3.1.3 above).

The user selects a fragment from the menu. When using the BASE FRAGMENT and USER BASE FRAG options, the selected fragment is loaded into the file and displayed in a pseudo-random location and orientation on the screen. Connections between fragments can be made by picking an atom on each substructure; the Builder automatically connects the fragments with appropriate bond lengths and angles. When using the FUNCTIONAL GRP and USER FUNCT GRP options, the user is prompted to pick the hydrogen atom on the screen which indicates where the functional group is to be attached. The hydrogen atom picked will be replaced by the new functional group.

User Input:

1. Pick BASE FRAGMENT, FUNCTIONAL GRP, USER BASE FRAG, or USER FUNCT GRP.
2. Pick the fragment(s) or functional group(s) from the library menu.
3. Functional group libraries only: pick the hydrogen atom(s) to be replaced on the molecule display. Pick off the molecule to complete the replacement(s).
4. Repeat steps 2 and 3 to add additional fragments or functional groups.
5. Pick RETURN to complete the operation.

Selections:	
acetylene	furan
ammonia	h3po4
anthracene	hcl
benzene	hcn
cyclobutane	hono2
cyclohexane	indene
cyclopentane	methane
cyclopropane	naphthalene
decalin_cis	phenalene
decalin_trans	phenanthrene
ethane	steroid
ethylene	tetralin
fluorene	water
formaldehyde	
formic_acid	return

Figure 3.5: Base fragment library menu

Selections:
acetyl
amino
carboxy
chloro
cyano
ethyl
formyl
hydroxy
methyl
nitro
phenyl
phosphino
sulfo
t_butyl
thiol
tosyl
vinyl
return

Figure 3.6:
Functional group library menu

3.1.5 READ FRAGMENT

This operation allows the user to read in a fragment file and display its atoms on the screen. The fragment is displayed in a pseudo-random location and orientation. A fragment file can be any Brookhaven or BioDesign General Format file (see Appendix D) in which the atoms are numbered from 1 to the number of atoms in the file.

User Input:

1. Pick READ FRAGMENT.
2. Enter the name of the fragment file.

3.1.6 COPY

The COPY operation allows the user to copy an entire molecule or some fragment of a molecule. After toggling COPY, the user is prompted to define the fragment (or molecule) to be copied. See Section 1.5.1 on how to define a fragment (pick a range of atoms). The copied structure can then be moved using the MOVE option and attached to another structure using the BOND, CONNECT FRAG, REPLACE, or CONDENSE options. The MOVE and BOND options are accessed under EDIT.

User Input:

1. Pick COPY.
2. Define the fragment by picking starting and stopping atoms on the molecule display as described in Section 1.5.1.

3.1.7 CONDENSE

The CONDENSE option allows the user to join two fragments by replacing two atoms and associated bonds with a new bond connecting the two fragments. The program prompts the user to select the atom to be deleted on each structure. For example, biphenyl can be created by picking a hydrogen on each of two benzene fragments. The first fragment is moved to the proper position and connected to the second fragment picked. (*Warning:* The second fragment must contain at least four atoms. It is not possible to condense an alkane with water to form an alcohol, but it is possible to form the alcohol by condensing water onto an alkane.)

User Input:

1. Pick CONDENSE.
2. Pick the first atom to be deleted. (Pick off the molecule to abort.)
3. Pick the second atom to be deleted.
4. Repeat steps 2 and 3 to join additional fragments or pick off the molecule to complete the operation.

3.1.8 REPLACE

The REPLACE option allows the user to join two fragments with a bond that replaces one atom and associated bonds on one of the fragments. The program prompts the user to select the atom to be replaced and then the replacement atom on the other structure. For example the atoms of methylethyl ether can be assembled by picking a hydrogen on ethane and then the oxygen on formaldehyde.

User Input:

1. Pick REPLACE.
2. Pick the atom to be replaced. (Pick off the molecule to abort.)
3. Pick the replacement atom.
4. Repeat steps 2 and 3 to join additional fragments or pick off the molecule to complete the operation.

3.1.9 CONNECT FRAG

The CONNECT FRAG option allows the user to join two fragments with a bond between selected atoms. The first substructure is attached, with the appropriate geometry, to the second substructure.

User Input:

1. Pick CONNECT FRAG.
2. Pick the atom on the first substructure. (Pick off the molecule to abort.)
3. Pick the atom on the second substructure.
4. Repeat steps 2 and 3 to join additional fragments or pick off the molecule to complete the operation.

3.1.10 DRAW

The DRAW option enables the user to employ the tablet to draw a molecule or to modify an existing molecule. The atoms drawn on the display are assigned the atom type which is displayed beneath DEFAULT ATOM on the menu. The atom type will determine the parameters used in the force fields for energy and dynamics calculations.

Atom type designators consist of five characters (see Appendix I.3.1). The first two indicate the atomic symbol, the third indicates the hybridization (3 for sp^3 , 2 for sp^2 , 1 for sp^1 and R for resonant as in benzene), the fourth indicates the number of implicitly bonded hydrogens, and the fifth indicates differences between closely related atom types. For example, the default atom type C_32 is used for methylene carbons.

If the currently active atom type is not correct, it can be changed by toggling on DEFAULT ATOM and selecting the proper atom type from those listed in the menu (see Figure 3.7). New atom positions can then be added. Each new atom will be bonded to the previous atom picked and placed in a plane which contains the previous atom picked and is parallel to the plane of the screen. If the user picks an existing atom, DRAW creates a bond between that atom and the previous atom picked or created. Thus double and triple bonds can be entered by redrawing the same bond twice or thrice.

The active atom type can be changed at any time during the drawing process, but it may be more efficient to change the atom types after the entire structure is drawn by using the ATOM TYPE operation described in Section 3.1.13.

To exit from drawing, toggle the red display border.

User Input:

1. Use DEFAULT ATOM to change the atom type if the current default is not desired (see Section 3.1.12).
2. Pick DRAW.

3. Pick the first atom position. If the first position picked is an existing atom, no new atom will be created; however, connectivity will start from this position.
4. Pick the next position. The atom created (or picked) at this position will be bonded to the most recently picked atom.
5. Repeat the previous step until current bonding circuit is complete. Each new atom will be bonded to the previous atom.
6. To terminate the bonding circuit, pick an atom twice.
7. Repeat steps 3-6 as often as necessary to complete the structure(s) or until a new orientation of the molecule is required.
8. Pick the red border to stop drawing.

3.1.11 DEFAULT ATOM

The menu item beneath DEFAULT ATOM displays the default atom type which will be used with the next DRAW operation (see Section 3.1.10). The current default may be changed by toggling DEFAULT ATOM and selecting a new atom type from those listed in the menu (see Figure 3.7).

User Input:

1. Pick DEFAULT ATOM.
2. Select the atom type to be used with the next DRAW operation from the menu.

Selections:		
H_	N_3	Cl
H__A	N_22	Ga3
H__b	N_21	Ge3
B_3	N_2	As3
B_2	N_R2	Se3
C_34	N_R1	Br
C_33	N_R	In3
C_32	N_1	Sn3
C_31	O_32	Sb3
C_3	O_31	Te3
C_22	O_3	I_
C_21	O_2	Na
C_2	O_R1	Ca
C_R2	O_R	Ti
C_R1	O_1	Fe
C_R	F_	Zn
C_11	Al3	Tc
C_1	Si3	Ru
N_33	P_3	
N_32	S_31	
N_31	S_3	return

Figure 3.7:
Atom types menu (Dreiding-II force field)

3.1.12 EDIT

The operations on the EDIT menu shown in Figure 3.8 allow the user to edit the structure built using the Builders. Functions are available which allow the user to drag selected atoms to a new position (DRAG) or to drag and have the structure automatically move in response so that its energy is optimized (DRAG-CLEAN). If desired, the user may keep selected atoms fixed during dynamic dragging. This may be useful in forcing a structure over an energy barrier. Other options allow the user to move or rotate substructures, invert stereochemistry, add missing hydrogens, and delete atoms. A SELECT option is provided which allows the user to specify which atoms these functions will operate on. One may select a substructure or range of atoms, select atoms individually, or specify that only hydrogens, heteroatoms, or lone pairs be included. One may also draw in new structures, add or delete bonds, rotate portions of a structure about covalent bonds, and specify bond lengths and angles. Options are also provided to minimize the structure, center it on the screen, and define atom types based on connectivities.

Edit:
select
show
drag
drag-clean
move
invert
add H
delete
draw
bond
rotate bond
set bond
delete bond
clean
center
auto type
defaults
default atom
C_32
return

Figure 3.8:
EDIT menu

SELECT

The **SELECT** option is used to specify which atoms are to be included when using the **DRAG**, **DRAG-CLEAN**, **MOVE**, **INVERT**, **ADD H** and **DELETE** options. After picking **SELECT**, the Selection methods menu shown in Figure 3.9 appears. The user then chooses from the eight different selection methods listed. All atoms may be included (**ALL**) or only a selected substructure (**SUBSTRUCTURE**). A range of atoms may be specified (**RANGE**) or atoms may be selected individually (**ATOMS**). Special types of atoms may also be specified: **HETEROATOMS** (all atoms except carbons and hydrogens), **HYDROGENS**, and **LONE PAIRS**. A **CONTINUOUS** option is also provided which allows the user to continue to perform the **DRAG**, **DRAG-CLEAN**, **INVERT**, **ADD H**, and **DELETE** operations on individual atoms selected one at a time. The default selection method is **CONTINUOUS**.

The selected atoms may be seen by choosing **SHOW** from the **EDIT** menu (selected atoms are marked with red crosses). None will appear if **CONTINUOUS** was chosen.

User Input:

1. Pick **SELECT**.
2. Pick the selection method desired from the Selection methods menu (see Figure 3.9).
3. **SUBSTRUCTURE**: Pick an atom from the substructure to be included (all connected atoms will be included).

RANGE: Pick the base atom. Pick stopping atom(s) to delimit the range. Pick a non-atom and then pick **EXECUTE** to complete the operation or pick **CONTINUE** to continue (**RETURN** will abort the range selection). (See Section 1.5.1 on picking a range of atoms).

ATOMS: Pick the atom(s) to be included. Pick a non-atom and then pick **EXECUTE** to complete the operation or pick **CONTINUE** to continue (**RETURN** will abort the atom selection). (See Section 1.5.1 on picking a list of atoms).
4. If desired, pick **SHOW** on the **EDIT** menu to show which atoms have been selected.

SHOW

The **SHOW** option is used to show which atoms have been selected with the **SELECT** option. The selected atoms are marked with red crosses. Only these atoms will be included when using the **DRAG**, **DRAG-CLEAN**, **MOVE**, **INVERT**, **ADD H** and **DELETE** options. None will appear if **CONTINUOUS** was chosen.

Selections:
all
substructure
range
atoms
heteroatoms
hydrogens
lone pairs
continuous
return

Figure 3.9:
Selection methods
menu

DRAG

The DRAG option is used to change the position of the atoms selected under SELECT. The user picks DRAG to activate the DRAG function. He then uses the *middle* mouse button to drag the selected atoms to the desired location. The bonds to the atom(s) are also dragged along.

If the CONTINUOUS option under SELECT was chosen, then the same procedure is followed except that the user must first pick the atom to be dragged (using the *left* mouse button) before dragging it with the *middle* mouse button. In this way, the user can continue to pick and drag individual atoms one at a time.

User Input:

1. Pick DRAG.
2. CONTINUOUS method only: use the left mouse button to pick the atom to be moved and activate the DRAG function.
3. Use the middle mouse button to drag the selected atom(s) to the desired location.
4. CONTINUOUS method only: repeat steps 2 and 3 for additional atoms as desired.
5. Pick off the molecule to terminate the DRAG process.

DRAG-CLEAN

DRAG-CLEAN functions in the same way as DRAG above except that the entire structure is automatically minimized in response to the new position of the dragged atom(s). Minimization continues until convergence is reached or until the user stops the process by picking off the molecule. If FIX SELECTED is also on, then the selected atoms will remain fixed during the minimization while the rest of the structure moves. (See FIX SELECTED under DEFAULTS below).

MOVE

The MOVE option is used to change the position of the atoms selected under SELECT. After picking MOVE, the selected atoms may be translated or rotated using the middle or right mouse buttons respectively. This is often useful in moving one substructure with respect to another. The structure being moved is shown in white (its original position is still displayed). Another menu is also displayed providing the user with the option of having the coordinates of the moved atoms updated to the new position (UPDATE) or returned to their original position (ABORT) (see Figure 3.10). If UPDATE was chosen, the bonds to the moved atoms are also dragged along. Picking RETURN updates and returns you to the EDIT menu. A ROTATE ALL toggle is also provided which allows the user to readily switch between moving the entire display (ROTATE ALL on) or just the selected atoms (ROTATE ALL off). MOVE functions similarly to DRAG except that it also allows rotations.

Selections:
rotate all
update
abort
return

Figure 3.10:
MOVE update
menu

Note: MOVE requires a minimum of two atoms to be selected (they must also be connected if doing rotations). If the CONTINUOUS option under SELECT is active (the default) or only one

atom has been selected, the user is prompted to select more atoms; that is, he must choose a selection method other than CONTINUOUS and specify at least two atoms to be moved (see SELECT, above).

User Input:

1. Use the SELECT function to pick the atoms to be moved. At least two atoms must be picked (the CONTINUOUS method cannot be used).
2. Pick MOVE.
3. Toggle ROTATE ALL on to move all the atoms on the display, toggle it off to move only the selected atoms.
3. Use the middle or right mouse buttons to translate or rotate the selected atom(s) to the desired location.
4. Pick UPDATE to update the coordinates of the moved atoms, pick ABORT to return them to their previous position, or pick RETURN to update and quit.

INVERT

The INVERT operation allows the user to invert stereochemistry. The inversion takes place for the atoms specified under SELECT. If ALL was chosen, then the entire structure will be inverted. If the selection method is CONTINUOUS, then the user picks the chiral atoms to be inverted one at a time. Currently the program cannot handle certain cases involving chiral atoms which are part of fused rings, since bonds must be broken and reconnected to generate the inverted structure. If the program cannot invert the selected atom, it issues a warning message.

User Input:

1. Pick INVERT.
2. If CONTINUOUS was chosen under SELECT, then pick the chiral atom to be inverted. Repeat for another chiral atom or pick off the molecule to terminate the function.

ADD H

The ADD H operation adds missing hydrogens to those atoms selected under SELECT. If CONTINUOUS was chosen, the user can continue to add missing hydrogens to atoms by selecting them one at a time.

User Input:

1. Pick ADD H.
2. If CONTINUOUS was chosen under SELECT, then pick the atom where the missing hydrogens are to be added. Repeat for another atom or pick off the molecule to terminate the function.

DELETE

The DELETE option deletes those atoms selected under SELECT. If CONTINUOUS was chosen, the user picks the atoms to be deleted from the screen.

User Input:

1. Pick DELETE.
2. If CONTINUOUS was chosen under SELECT, then pick the atom to be deleted from the screen. Repeat for another atom or pick off the molecule to terminate the function.

DRAW

The DRAW operation is described in Section 3.1.10 above.

BOND

The BOND option allows the user to add bonds to a structure. The atoms are not repositioned, unlike the CONNECT FRAG operation which is intended for joining fragments. This option is typically used to add the additional bonds needed for multiply bonded systems.

User Input:

1. Pick BOND.
2. Pick the first and second atoms to be bonded.
3. Repeat the above two steps as required.
4. Pick off the molecule to terminate.

ROTATE BOND

The ROTATE BOND operation is described in Section 3.9.

SET BOND

The SET BOND option allows the user to set bond lengths, bond angles, or dihedral angles to a specified value using the menu shown in Figure 3.11. The user picks atoms from the display to specify the bond or angle to be changed; the selected atoms are marked with colored crosses.

A. DISTANCE

The DISTANCE option allows the user to specify the bond distance between two selected atoms. After picking the two atoms which define the bond from the screen, the current bond distance is displayed in the dialog box. The user enters the new value and the program rebuilds the structure with the new bond distance.

Set bond:
distance
angle
dihedral
return

Figure 3.11:
SET BOND menu

User Input:

1. Pick DISTANCE.
2. Pick two atoms from the screen to define the bond.

3. Enter the new bond length in Å.

B. ANGLE

The ANGLE option allows the user to specify the angle between two bonds. After picking the first, second (central), and third atoms from the screen, the current bond angle is displayed in the dialog box. The user enters the new value and the program rebuilds the structure with the new bond angle. The method of changing the angle is that the third atom and the structure attached to it are moved to the correct position.

Note: If all three atoms are in the same ring, the angle may not be changed.

User Input:

1. Pick ANGLE.
2. Pick three atoms from the screen to define the angle (in the proper order).
3. Enter the new bond angle in degrees.

C. DIHEDRAL

The DIHEDRAL option allows the user to specify dihedral angles by selecting four connected atoms on the screen. The atoms must be picked in the proper order (the second and third atoms selected represent the common bond atoms). The picked atoms are sequentially marked with red, yellow and blue crosses. After picking the fourth atom, a dialog box appears giving the current dihedral angle value. The user then enters the new value and the program rebuilds the structure using the dihedral angle specified. The method of changing the dihedral angle is that the structure attached to the third atom is rotated to the correct position.

Note: If the two central atoms are in the same ring, the dihedral angle may not be changed.

User Input:

1. Pick DIHEDRAL.
2. Pick four connected atoms making up the dihedral angle to be specified (in the proper order).
3. Enter the new dihedral angle in degrees.

DELETE BOND

The DELETE BOND operation allows the user to delete a bond by picking the atoms comprising the bond. *Note:* DELETE BOND must be used twice to delete a double bond.

User Input:

1. Pick DELETE BOND.
2. Pick the two atoms that make up the bond.
3. Repeat step 2 as often as required.
4. Pick off the molecule to terminate the function.

CLEAN

Toggling CLEAN performs an energy minimization on the structure. It is the same as using the MINIMIZE option on the MECHANICS menu (under SIMULATE) with all atoms movable (see Section 5.7). NUMBER OF STEPS of minimization are performed (the default is 50 steps). If the user wishes to keep certain atoms fixed, these may be specified using the FIX SELECTED option under DEFAULTS. This option is described below.

CENTER

Toggling CENTER centers the structure in the middle of the display by modifying its coordinates so that the average value of x, y, and z each equal 0.

AUTO TYPE

The AUTO TYPE operation allows the user to define atom types based on connectivities. After toggling AUTO TYPE the element symbols for each atom are displayed on the screen and the Atomic elements menu appears listing the possible elements as shown in Figure 3.12. If no element types need to be changed, the hybridizations may be corrected by picking RETURN. To change the type of element for an atom, the user toggles the desired element on the menu and then picks the corresponding atom(s) on the screen. This may be done for different elements. The user just picks anywhere off the molecule and then repeats the process for the new element type. Once the atoms have been set to the desired elemental types, the user toggles RETURN to execute atom typing based on connectivities. The new atom types are displayed on the screen. If the number of the covalent bonds is less than the maximum, then the program assumes that the remaining bonds are to implicit hydrogens. Any atom type generated with AUTO TYPE can be changed by using the ATOM TYPE option under EDIT ID (Section 3.1.13).

User Input:

1. Pick AUTO TYPE.
2. If changing element types, pick the element symbol from the menu then pick the corresponding atom(s) on the screen. Terminate by picking off the molecule. Repeat as needed.
3. Pick RETURN on the Atomic elements menu to generate the corrected atom types.
4. Click anywhere to remove the atom type display.

Selections:	
H	Zn
B	Ga
C	Ge
N	As
O	Se
F	Br
Na	Tc
Al	Ru
Si	In
P	Sn
S	Sb
Cl	Te
Ca	I
Ti	
Fe	return

Figure 3.12: Atomic elements menu (Dreiding-II force field)

DEFAULTS

Only one default is currently available (FIX SELECTED). This is described below.

A. FIX SELECTED

FIX SELECTED is an on/off toggle which applies to the DRAG-CLEAN and CLEAN options. When on, the atoms specified under SELECT remain fixed while the rest are allowed to move as the structure is minimized during the DRAG-CLEAN operation. The selected atoms will also be treated as fixed if the user performs an energy minimization by toggling CLEAN. When FIX SELECTED is off, all atoms are allowed to move.

Defaults:
fix selected
return

Figure 3.13:
DEFAULTS menu

User Input:

1. Use the SELECT option to select the atoms which will be fixed.
2. Turn FIX SELECTED on to keep the selected atoms fixed. Turn FIX SELECTED off if all atoms are to be movable.

DEFAULT ATOM

The menu item beneath DEFAULT ATOM displays the default atom type which will be used with the next DRAW operation (see Section 3.1.10). The current default may be changed by toggling DEFAULT ATOM and selecting a new atom type from those listed in the Atom types menu (see Figure 3.8).

User Input:

1. Pick DEFAULT ATOM.
2. Select the atom type to be used with the next DRAW operation from the menu.

3.1.13 EDIT ID

The EDIT ID menu shown in Figure 3.14 is used to modify the labels and force field types on specified atoms, to reassign residue names and numbers and chain identifiers, and to reassign charges. The options are described below.

Edit id:
atom type
auto type
atom labels
auto labels
residue
chain
charges
auto charges
return

Figure 3.14:
EDIT ID menu

ATOM TYPE

The ATOM TYPE option allows the user to reassign atom types. When ATOM TYPE is toggled, a new menu appears listing the possible atom types. The Atom types menu for the Dreiding-II force field is shown in Figure 3.15. The current atom types are displayed next to each atom on the screen.

User Input:

1. Pick ATOM TYPE.
2. Pick the desired atom type from the menu.
3. Pick the atoms to receive this atom type. Pick off the molecule when done.
4. Repeat steps 2 and 3 until all atoms have been assigned the proper types.
5. Terminate by picking RETURN.
6. Click anywhere to remove the atom type display.

AUTO TYPE

The AUTO TYPE operation is described under EDIT in Section 3.1.12 above.

Selections:		
H_	N_3	Cl
H_A	N_22	Ga3
H_b	N_21	Ge3
B_3	N_2	As3
B_2	N_R2	Se3
C_34	N_R1	Br
C_33	N_R	In3
C_32	N_1	Sn3
C_31	O_32	Sb3
C_3	O_31	Te3
C_22	O_3	I_
C_21	O_2	Na
C_2	O_R1	Ca
C_R2	O_R	Ti
C_R1	O_1	Fe
C_R	F_	Zn
C_11	Al3	Tc
C_1	Si3	Ru
N_33	P_3	
N_32	S_31	
N_31	S_3	return

Figure 3.15: Atom types menu (Dreiding-II force field)

ATOM LABELS

Toggling ATOM LABELS shows the current atom labels on the display. A new label can be given to any atom by picking it on the display and then entering up to four characters in the dialog box.

User Input:

1. Pick ATOM LABELS.
2. Pick the atom whose label you wish to modify from the display.
3. Enter the new atom label (up to four characters).
4. Repeat steps 2 and 3 as required.

5. Terminate by picking off the molecule.
6. Click anywhere to remove the atom label display.

AUTO LABELS

Toggling AUTO LABELS generates a set of atom labels based on the current atom types. The user may click anywhere to remove the atom label display.

RESIDUE

The RESIDUE option is used to reassign residue names and numbers. The user first defines the molecular segment that will receive the new residue name and number by picking a range of atoms from the display (Section 1.5.1 describes how to pick a range of atoms). Once the substructure is defined, the user picks EXECUTE and then enters the new residue name and number. This process is repeated until all the residue information has been specified. The user then picks off the molecule to complete the function. This causes the program to sort the atoms based on the newly entered data so that they are in a new sequence. This is useful in creating groups where residue names and numbers are used to specify a segment of a molecule.

User Input:

1. Pick RESIDUE.
2. Specify the substructure by picking the starting and stopping atoms from the display. Pick off the molecule to indicate selection is complete.
3. Pick EXECUTE to complete the substructure specification (pick RETURN to abort the operation).
4. Enter the residue name (up to three-letters) followed by a space and the residue number (0-999) (*e.g.*, TYR 10).
5. Repeat steps 2 - 4 until all the new residue information has been specified.
6. Pick off the molecule to terminate the function.
7. Click anywhere to remove the residue display.

CHAIN

The CHAIN option allows the user to reassign chain designators. The user first defines the molecular segment that will receive the new designator by picking a range of atoms from the display (Section 1.5.1 describes how to pick a range of atoms). Once the substructure is defined, the user picks EXECUTE and then enters the new chain designator. Chain designators are denoted by a one-letter (upper case) or blank designator (use an underscore for a blank chain designator). This process is repeated until all the new chain designators have been specified. The user then picks off the molecule to complete the function. (Picking ABORT at any time will terminate the function). This causes the program to sort the atoms based on the newly entered data so that they are in a new sequence.

User Input:

1. Pick CHAIN.

2. Specify the substructure by picking the starting and stopping atoms from the display. Pick off the molecule to complete the substructure specification.
3. Pick EXECUTE to execute the operation or RETURN to abort the operation.
4. Enter the new chain designator: one-letter (upper case) or an underscore for a blank chain designation.
5. Repeat steps 2 - 4 until all the new chain designators have been specified.
6. Pick off the molecule to terminate.
7. Click anywhere to remove the chain display.

CHARGES

The CHARGES option allows the user to reassign atomic charges. The CHARGES menu shown in Figure 3.16 provides three methods for calculating atomic charges on selected groups of atoms: a charge equilibration method (Q EQUILIBRATE), a method described by Gasteiger (GASTEIGER) and a method described by Del Re (DEL RE). The user may also specify charges for individual atoms using the EDIT CHARGES option.

Charges:
Q equilibrate
Gasteiger
Del Re
edit charges
return

Figure 3.16:
CHARGES menu

A. Q EQUILIBRATE

This option uses a charge equilibration method¹ for calculating charges. After picking Q EQUILIBRATE, the menu shown in Figure 3.17 is displayed.

SETUP The SETUP option is used to select the groups whose charges are to be modified and to setup a group for displaying the charges on the screen. SETUP *must* be done if the CHARGES functions were accessed from the SIMULATE menu. However, if accessing CHARGES via EDIT ID from one of the Builders, then SETUP may be skipped. In this case, the group being built is automatically selected as the group to be modified.

TOTAL CHARGE The total charge of the group being modified is displayed under TOTAL CHARGE on the menu. The user may specify another value for TOTAL CHARGE, however, which will be applied when the charge distribution is calculated (Q EQUILIBRATE method only). The value is given in units of electrons, where -1 indicates that the system has a net negative charge (one extra electron).

Q equilibrate:
setup
total charge
0.0000
charge converge
0.0005
execute
return

Figure 3.17:
Q EQUILIBRATE
menu

¹ A. K. Rappé and W. A. Goddard III (submitted).

CHARGE CONVERGE If there are hydrogen atoms in the structure, the computation of charges is done iteratively. Iteration continues until the largest change in the charge on a hydrogen is less than the convergence criteria (CHARGE CONVERGE) or until 30 iterations have been done. The default value is 0.0005.

EXECUTE EXECUTE calculates the charges and shows the new values for each of the atoms on the display. The calculation can be interrupted with the abort key (Q EQUILIBRATE method only); charges from the last complete iteration are returned.

User Input:

1. Pick Q EQUILIBRATE.
2. Pick SETUP (if in a Builder and using the default groups, skip SETUP and go to step 5).
3. Pick the groups whose charges are to be modified from the Groups menu; terminate by picking RETURN (this step is skipped if there is only one group in the system).
4. Pick the group to be used for displaying the charges from the Groups menu; terminate by picking RETURN.
5. To change the total charge, pick TOTAL CHARGE and enter the new value (in electrons).
6. To change the convergence criteria, pick CHARGE CONVERGE and enter the new value (electrons).
7. Pick EXECUTE to calculate the charges and have them displayed on the screen.

B. GASTEIGER

This option uses the Gasteiger method² to calculate charges. When GASTEIGER is picked, a menu similar to that shown in Figure 3.17 is displayed (the menu is the same except that there is no CHARGE CONVERGE option). The SETUP and EXECUTE options function as described above for Q EQUILIBRATE. TOTAL CHARGE displays the total charge for the group being modified, but this value may not be changed by the user.

User Input:

1. Pick GASTEIGER.
2. Pick SETUP (if in a Builder and using the default groups, skip SETUP and go to step 5).
3. Pick the groups whose charges are to be modified from the Groups menu; terminate by picking RETURN (this step is skipped if there is only one group in the system).
4. Pick the group to be used for displaying the charges from the Groups menu; terminate by picking RETURN.
5. Pick EXECUTE to calculate the charges and have them displayed on the screen.

C. DEL RE

This option is the same as that described for GASTEIGER above, except that the Del Re method³⁻⁴ is used to calculate the charges.

² J. Gasteiger, *Tetrahedron*, **36** (1980) 3219-3228.

³ G. Del Re, *J. Chem Soc.*, (1958) 4031-4040.

User Input:

1. Pick DEL RE.
2. Pick SETUP (if in a Builder and using the default groups, skip SETUP and go to step 5).
3. Pick the groups whose charges are to be modified from the Groups menu; terminate by picking RETURN (this step is skipped if there is only one group in the system).
4. Pick the group to be used for displaying the charges from the Groups menu; terminate by picking RETURN.
5. Pick EXECUTE to calculate the charges and have them displayed on the screen.

D. EDIT CHARGES

The EDIT CHARGES option is used to assign charges to selected atoms (see Figure 3.18).

The SETUP option functions as described for Q EQUILIBRATE above. TOTAL CHARGE displays the total charge, but may not be changed by the user. After the groups have been SETUP, the ASSIGN CHARGES and ZERO CHARGES options become operational. These are described below.

User Input:

1. Pick EDIT CHARGES.
2. Pick SETUP (if in a Builder and using the default groups, skip SETUP and go to step 5).
3. Pick the groups whose charges are to be modified from the Groups menu; terminate by picking RETURN (this step is skipped if there is only one group in the system).
4. Pick the group to be used for displaying the charges from the Groups menu; terminate by picking RETURN.
5. Perform the ASSIGN CHARGES or ZERO CHARGES operations (see below).

Edit charges:
setup
total charge
0.000
assign charges
zero charges
return

Figure 3.18:
EDIT CHARGES
menu

ASSIGN CHARGES The ASSIGN CHARGES option allows the user to reassign atomic charges to selected atoms. The current charges are displayed for each atom on the screen. Atom charges are expressed in electrons. A Charges menu appears listing charge values ranging from -1.00 to +3.00 (see Figure 3.19). The user picks the new charge value from the menu, then selects the corresponding atom on the screen. A red cross is placed on the picked atom and the new charge value is displayed. If the charge value desired does not appear on the menu, it can be entered directly. There are four special items on the Charges menu for this purpose: USER1, USER2, USER3 and KEY. If KEY is toggled, the user is prompted to pick one of the USER entries and then to enter a value for the charge. This value is then displayed on the menu in place of the selected USER entry; it can be used to set charges.

⁴ G. Del Re, B. Pullman and T. Yonezawa, *Biochim. Biophys. Acta*, **75** (1963) 153-182.

User Input:

1. Pick ASSIGN CHARGES.
2. Pick the desired charge from the Charges menu. (To set a charge value that is not on the menu, pick KEY, pick one of the USER items, then enter the desired charge value).
3. Pick the atoms to receive this charge. Terminate by picking off the molecule.
4. Repeat steps 2 and 3 until all atoms have been assigned proper charges.
5. Terminate by picking RETURN on the Charges menu.

Charges:		
-1.00	-2.00	+0.60
-0.95	-0.15	+0.65
-0.90	-0.10	+0.70
-0.85	-0.05	+0.75
-0.80	+0.00	+0.80
-0.75	+0.05	+0.85
-0.70	+0.10	+0.90
-0.65	+0.15	+0.95
-0.60	+0.20	+1.00
-0.55	+0.25	+2.00
-0.50	+0.30	+3.00
-0.45	+0.35	user1
-0.40	+0.40	user2
-0.35	+0.45	user3
-0.30	+0.50	key
-0.25	+0.55	return

Figure 3.19: Charges menu

ZERO CHARGES The ZERO CHARGES option makes all the charges zero.

User Input:

1. Pick ZERO CHARGES to make all the charges zero.

3.1.14 MECHANICS

The MECHANICS option is designed to perform energy minimizations and/or molecular dynamics calculations on structures built with the Builders (under BUILD). Toggling MECHANICS brings up the MECHANICS menu and sets up the energy expression so that all the atoms in the group being built are included and are movable. The MECHANICS operations are described in Section 5.7. See Section 5.2 for more information on setting up energy expressions.

User Input:

1. Pick MECHANICS.
2. Perform the desired operations on the MECHANICS menu (see Section 5.7).

3.1.15 SIMULATE

The SIMULATE option, which allows the user to perform minimizations, dynamics, and conformational searches with general energy expressions, is described in Chapter 5.

3.1.16 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.2 PEPTIDE

The operations on the PEPTIDE menu (see Figure 3.20) allow the user to build peptides from primary sequences based on the fragment library contained in the *bg_pep_l* and *bg_pep_d* directories. After toggling PEPTIDE, the fragments for the D- and L-amino acids (20 each) are loaded into the program. The SETUP operation is automatically invoked and the user is prompted to specify the file to be used (a new file may be created or the peptide may be added to an existing file). If the file is new, the user enters a name for it. A group of the same name is also created.

The residues that define the primary sequence of the peptide are then selected using the L_LIBRARY option. The D_LIBRARY option is used to add D-amino acid residues, which are rare in natural peptides. The default is to build the peptide in an extended configuration ($\phi = 180^\circ$, $\psi = 180^\circ$). The default ϕ and ψ angles can be changed to values appropriate for α -helices, β -sheets, β -turns or to arbitrary values using the DEFINE ANGLE option.

Note: No special manipulations are required to build peptides when using the Molecular Simulations implementation of the AMBER force field. The atom types produced allow energy calculations to be run.

Peptide:
setup
l_library
d_library
user library
read library
define angle
modify
replace
insert
delete
edit
edit id
mechanics
simulate
analyze
return

Figure 3.20:
PEPTIDE menu

3.2.1 SETUP

The SETUP operation of the Peptide Builder is similar to the Organic Builder's SETUP function, which is described in Section 3.1.1. It allows the user to pick a new file for the peptide or add the peptide to an existing file. SETUP is automatically invoked each time the Peptide Builder is entered.

User Input:

1. Pick SETUP (needed only if already in the Peptide Builder and setting up again).
2. Pick the file on which to operate from the Files menu. Pick NEW FILE if creating a new file. This step is skipped if there are no files in the system.
3. Enter a descriptor (up to 8 characters) for the file if the file is new.

3.2.2 L_LIBRARY, D_LIBRARY

The L_LIBRARY option allows the user to add selected L-amino acid residues to the carboxyl (C) terminus of the peptide (see Figure 3.21). The residues are selected from the L_LIBRARY menu and are joined to the C-terminus according to the default values for the ϕ , ψ , and ω angles. After exiting the option, any necessary hydrogens or lone pairs are automatically added and the charges on the amino and carboxyl termini are corrected. The D_LIBRARY option does the same as L_LIBRARY but for D-amino acid residues.

User Input:

1. Pick L_LIBRARY or D_LIBRARY.
2. Pick the desired residues from the L_LIBRARY menu.
3. Pick RETURN to terminate the addition of residues.

3.2.3 USER LIBRARY

The USER LIBRARY option allows the user to read fragments from a user-defined directory of fragments (*bgi* files). The fragments contained in the directory are displayed in alphabetical order on the USER LIBRARY menu. Selecting a fragment from the menu loads it into a file and displays it on the screen. The fragments are displayed at pseudo-random locations and orientations.

Note: The directory information for the user-defined library must be read using READ LIBRARY before USER LIBRARY may be used.

User Input:

1. Pick USER LIBRARY.
2. Pick the fragment(s) or functional group(s) from the USER LIBRARY menu.
3. Pick RETURN to terminate.

Selections:
ala_l
arg_l
asn_l
asp_l
cys_l
gln_l
glu_l
gly_l
his_l
ile_l
leu_l
lys_l
met_l
phe_l
pro_l
ser_l
thr_l
trp_l
tyr_l
val_l
return

Figure 3.21:
L_amino acid
residues menu

3.2.4 READ LIBRARY

The READ LIBRARY option allows the user to read the directory information for one of three libraries selected from the READ LIBRARY menu: the L-amino acid library (L_LIBRARY), the D-amino acid library (D_LIBRARY), or a user-defined library (USER LIBRARY) (see Figure 3.22). The directory information must be read before the options which access these libraries can be used (see Sections 3.2.2 and 3.2.3). The directories containing the D- and L- libraries are automatically read in when the user first enters the Peptide Builder. The only time the user would need to read one of these default directories again would be if the user had added *bfg* fragment files to it during a session. The user-defined directory *must* be read in, however, in order to allow access to the user-defined libraries.

Selections:
l_library
d_library
user library
return

Figure 3.22:
Peptide libraries
menu

User Input:

1. Pick READ LIBRARY.
2. Pick the library to be read in (L_LIBRARY, D_LIBRARY or USER LIBRARY).
3. Enter the name of the directory (for the user-defined library). See Section 1.6 for details on entering file names.

3.2.5 DEFINE ANGLE

The operations on the DEFINE ANGLE menu shown in Figure 3.23 allow the user to change the default ϕ , ψ , and ω angles. The default values are 180° for ϕ and ψ and 0° for ω , which yields extended configurations. The angles can be changed to arbitrary values using the PHI, PSI, and OMEGA options. The angles can also be changed to values appropriate for a number of standard structures by toggling the appropriate operation listed in Table 3.1. To build a β -turn, two residues must be built in succession with the appropriate dihedral angles (1A,1B or 2A,2B).

The user may also change the default χ angles using the CHI1, CHI2, CHI3 and CHI4 options (up to four dihedral angles may be changed for each side-chain). The default χ values are all 60° .

Changes made using DEFINE ANGLE are not implemented until the MODIFY function is executed (see Section 3.2.6).

OMEGA, PHI, PSI, CHI1, CHI2, CHI3, CHI4

Toggling OMEGA, PHI, PSI, CHI1, CHI2, CHI3 or CHI4 allows the user to enter arbitrary values for the dihedral ϕ , ψ , or ω angles or for the side-chain dihedral χ angles.

User Input:

1. Pick OMEGA, PHI, PSI, CHI1, CHI2, CHI3 or CHI4.
2. Enter the new value for the ϕ , ψ , or ω angle or for the χ_1 , χ_2 , χ_3 , or χ_4 angle.
3. Use MODIFY to have the structure changed to reflect the new values (see Section 3.2.6 below).

Define angle:
omega=180.000
phi=180.000
psi=180.000
chi1=60.000
chi2=60.000
chi3=60.000
chi4=60.000
alpha helix
par beta
anti beta
extended
-50, -40
-75, -20
50, 105
100, 0
return

Figure 3.23:
DEFINE ANGLE
menu

ALPHA HELIX, PAR BETA, EXTENDED, -50 -40, -75 -20, 50 105, 100 0

Toggling one of these options changes the ϕ , ψ , and ω angles appropriately (as shown in Table 3.1). The structure will be altered to reflect the new values when the MODIFY function is executed (see Section 3.2.6 below).

User Input:

1. Pick ALPHA HELIX, PAR BETA, EXTENDED, -50 -40, -75 -20, 50 105, or 100 0.
2. Use MODIFY to have the structure changed to reflect the new values (see Section 3.2.6 below).

Operation	Structure	ϕ	ψ
ALPHA HELIX	α -Helix	-57	-47
PAR BETA	Parallel β -sheet	-119	113
ANTI BETA	Anti-parallel β -sheet	-139	135
EXTENDED	Extended	180	180
-50,-40	β -turn (1A)	-50	-40
-75,-20	β -turn (1B)	-75	-20
50,105	β -turn (2A)	50	105
100,0	β -turn (2B)	100	0

Table 3.1: DEFINE ANGLE operations for standard structures

3.2.6 MODIFY

The operations on the MODIFY menu shown in Figure 3.24 allow the user to modify the dihedral ϕ , ψ , or ω angles in the polypeptide being constructed or to change up to four side-chain dihedral χ angles. The default angles are changed by setting new default angles (DEFINE ANGLE), picking a range of residues (PICK START RES, PICK STOP RES), and then executing the change (EXECUTE).

DEFINE ANGLE

The DEFINE ANGLE option is described in Section 3.2.5.

PICK START RES, PICK STOP RES

The PICK START RES and PICK STOP RES options allow the user to pick the starting and stopping residues which define the range over which the ϕ , ψ , ω or χ angles will be changed.

User Input:

1. Pick PICK START RES or PICK STOP RES.
2. Pick the starting or stopping residue from the display. The residue number will appear below PICK START RES or PICK STOP RES on the menu.

EXECUTE

Toggling EXECUTE changes the ϕ , ψ , ω or χ angles to the current values set under DEFINE ANGLE for the residues between and including PICK START RES and PICK STOP RES.

Modify:
define angle
pick start res
- - -
pick stop res
- - -
execute
return

Figure 3.24:
MODIFY menu

3.2.7 REPLACE

The REPLACE operations shown in Figure 3.25 are designed to simulate site-directed mutagenesis and allow the user to replace single amino acid side-chains. Picking NEW RESIDUE displays a menu listing the available amino acid residues. After selecting the new side-chain from this list, the residue to be replaced is picked using the OLD RESIDUE option.

User Input:

1. Pick REPLACE.
2. If needed, use READ LIBRARY to read in the directory information for the library to be used.
3. Make sure the appropriate library is toggled on (LIBRARY for the program's side chain fragment library or USER LIBRARY for the user-defined side chain library).
4. Use NEW RESIDUE to select the new side chain.
5. Use OLD RESIDUE to select the residue to be replaced.
6. Pick EXECUTE to replace the old residue with the new one.

Replace:
read library
library
user library
new residue
- - -
old residue
- - -
execute
return

READ LIBRARY

The READ LIBRARY option allows the user to read the directory information from a user-defined side-chain library or from the program's side-chain fragment library *bg_side*. The directory information must be read before the USER LIBRARY option can be used. When the user picks REPLACE, the *bg_side* directory is read automatically. The only time the user would need to read the *bg_side* directory again would be when the user has added side-chain fragment files to *bg_side* during a session.

Figure 3.25:
REPLACE menu

User Input:

1. Pick READ LIBRARY.
2. Pick BIOGRAF to read the *bg_side* directory or USER to read a user-defined directory.
3. If USER was picked, enter the name of the directory for the user-defined fragment library.
See Section 1.6 for details on entering file names.

LIBRARY

If LIBRARY is toggled on, the NEW RESIDUE operation uses the side-chains in directory *bg_side*. LIBRARY is on by default.

USER LIBRARY

Picking USER LIBRARY makes the NEW RESIDUE operation use the side-chains in the user-defined directory.

NEW RESIDUE

Picking NEW RESIDUE displays a new menu listing the amino acid residues available. If LIBRARY was picked, then this menu will contain the side-chains in the *bg_side* directory as shown in Figure 3.26; if USER LIBRARY was picked, then the side-chains in the user-defined fragment library will be displayed. The program prompts the user to pick a new side-chain. The side-chain picked will then be indicated in the box below NEW RESIDUE on the REPLACE menu. The new side-chain will replace the side-chain selected with the OLD RESIDUE operation.

User Input:

1. Pick NEW RESIDUE.
2. Pick a new side-chain from the adjacent menu.
3. Use OLD RESIDUE to select the side chain to be replaced.
4. Pick EXECUTE to replace the old residue with the new one.

OLD RESIDUE

When OLD RESIDUE is picked, the program prompts the user to pick a residue to be replaced. The user picks an atom from the residue on the display. The atom picked is marked with a red cross and the residue name and number are indicated below OLD RESIDUE on the REPLACE menu. The EXECUTE operation will replace this residue with the side-chain selected during the NEW RESIDUE operation. Residues can be picked until the desired one is obtained.

User Input:

1. Pick OLD RESIDUE.
2. Pick an atom of the side-chain to be replaced.

EXECUTE

When EXECUTE is picked, the program replaces the old residue with the new one. The orientation is arbitrary, so the user will generally need to use the ROTATE BOND operation (see Section 3.9) to rotate the side-chain into a reasonable starting geometry and then minimize the energy to determine an optimal structure (see Section 5.7). If there are a large number of atoms in the program, the user can minimize the energy more quickly by defining two groups: (1) the side-chain to be replaced and (2) all residues within 5 to 10 Å of the side-chain to be replaced (see KEY ONE and DISTANCE under MAKE GROUPS/LIMITS/SELECTORS in Section 4.1.6). The geometry of the new side-chain can then be rapidly optimized by allowing the atoms of the first group to be movable (except the α -carbon which should be fixed using MODIFY EEX in Section 5.3) and the atoms of the second group to be fixed (see Section 5.2).

Selections:
ala
arg
asn
asp
cys
gln
glu
gly
his
ile
leu
lys
met
phe
pro
ser
thr
trp
tyr
val
return

Figure 3.26:
Amino acid
residues menu

User Input:

1. Pick EXECUTE to replace the old residue with the new one.

3.2.8 INSERT

The options on the INSERT menu shown in Figure 3.27 allow the user to insert amino acid residues into the peptide. The user specifies the point of insertion by choosing the INSERT AFTER option and selecting a residue on the display. The insertion will occur after the selected residue. The BUILD option is used to create the piece to be inserted. It provides a separate BUILD menu with options which function as in the Peptide Builder. Upon returning from BUILD, the new piece is automatically inserted. If the RENUMBER option had been selected (toggled to YES), then the residues in the modified structure will be renumbered to account for the insertion. If desired, the user may then change the backbone dihedral angles of the newly inserted piece (DEFINE ANGLE) and rebuild the structure using these new angle values (REBUILD).

The conformation of the new structure can be optimized using the SIMULATE option. The remaining options (N OFFSET, C OFFSET, SEPARATION, DISTANCE CUTOFF and CALC LEVEL) affect the way the energy optimization is done (either by specifying the movable and fixed atoms or by setting distance or dihedral angle constraints on the molecule). These options can therefore be used to control the shape the new structure will take.

DEFINE ANGLE

The DEFINE ANGLE option allows the user to set the default backbone dihedral angles for the newly inserted structure. This is described in Section 3.2.5.

INSERT AFTER RES

The INSERT AFTER RES option allows the user to specify the residue after which the insertion is to occur. This is done by selecting an atom on the screen; the name and number of the residue chosen is then displayed beneath INSERT AFTER RES on the INSERT menu.

User Input:

1. Pick INSERT AFTER RES.
2. Pick the residue to insert after from the display.

Insert:
define angle
insert after res
- - -
renumber
yes
build
rebuild
n offset=2
c offset=2
separation
0.000
distance cutoff
0.000
calc level
0
simulate
return

Figure 3.27:
INSERT menu

RENUMBER

The RENUMBER option allows the user to renumber the residues in the peptide to account for the insertion. It is a YES/NO toggle, with the default being YES.

User Input:

1. Toggle RENUMBER until the YES/NO box beneath it is set as desired.

BUILD

The BUILD option is used to create the piece to be inserted. After picking BUILD, the menu shown in Figure 3.28 comes up. The options on this menu function in the same way as those on the PEPTIDE menu described in Section 3.2. They are listed below. The RETURN option functions differently, however, in that it not only causes you to return to the previous level, but it also automatically inserts the newly built structure into the peptide. If you are not satisfied with the piece you inserted, you can always delete it using the DELETE option (see Section 3.2.9).

L LIBRARY, D LIBRARY, USER LIBRARY, DEFINE ANGLE, MODIFY: These options are used to construct the peptide piece to be inserted. They are described in Sections 3.2.2 - 3.2.6.

EDIT, EDIT ID: These two options are used to edit the peptide piece to be inserted. They are described in Sections 3.1.12 and 3.1.13.

MECHANICS, SIMULATE: MECHANICS performs energy minimizations, dynamics, and conformational searches on the structures built under BUILD and contained in the PEPTIDE group. The energy expression is automatically set up. MECHANICS is described in Section 5.7. The SIMULATE option performs the same operations using general energy expressions and is described in Chapter 5.

ANALYZE: The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

RETURN: The RETURN option is somewhat unique on this menu in that it not only causes you to return to the previous level (the INSERT menu), but it also inserts the piece just built into the peptide. You may wish to reposition the altered structure.

Build:
l_library
d_library
user library
define angle
modify
edit
edit id
mechanics
simulate
analyze
return

Figure 3.28:
BUILD menu

User Input:

1. Pick BUILD.
2. Use the options on the menu to create the piece to be inserted.
3. Pick RETURN to insert the newly built piece into the peptide and return to the INSERT menu.

REBUILD

The REBUILD option allows the user to implement changes in the backbone dihedral angles of the newly inserted residues after having inserted them into the peptide. The current ϕ , ψ , and ω angles are changed by setting new default angles with the DEFINE ANGLE option. REBUILD then rebuilds the structure using the new angle values.

User Input:

1. Pick REBUILD.

N OFFSET, C OFFSET

The N OFFSET option allows the user to specify the number of residues on the amino side of the the insertion point that are allowed to move during the energy calculations. Similarly C OFFSET specifies this number on the carboxy side of the insertion point. The default value for both options is 2.

User Input:

1. Pick N OFFSET or C OFFSET.
2. Enter the number of residues from the insertion point that are to be included in Level 2 energy calculations.

SEPARATION

The SEPARATION option sets the endpoint separation distance of the inserted peptide. This is the distance between the two ends of the inserted piece and is a distance constraint which is applied during the energy calculations. Upon returning from BUILD, the current value appears beneath SEPARATION on the menu. The default separation distance is 6.0 Å.

User Input:

1. Pick SEPARATION.
2. Enter the endpoint separation distance in Å.

DISTANCE CUTOFF

The DISTANCE CUTOFF option allows the user to specify an area of fixed atoms around the movable atoms when doing the energy calculations. All atoms that are within the distance cutoff value of any of the movable atoms will be included in the energy calculations as fixed atoms.

User Input:

1. Pick DISTANCE CUTOFF.
2. Enter the cutoff distance in Å.

CALC LEVEL

The CALC LEVEL option provides the user with three levels for doing the energy calculations (0, 1, and 2). Each level sets different distance or dihedral angle constraints and allows certain parts of the peptide structure to relax or remain fixed. It is intended that the user cycle through each calculation level in sequence starting with 0, doing energy calculations each time. The default level is 0.

Level 0: This level uses the separation distance constraint (SEPARATION) in an attempt to move the endpoints of the newly inserted peptide to the indicated value. This allows the protein to readjust to the new piece.

Level 1: Level 1 only permits the inserted piece to move; all other atoms remain fixed. A dihedral angle constraint ($\omega = 180^\circ$) is set in order to keep both of the terminal peptide bonds in the *trans* configuration.

Level 2: Level 2 not only allows the inserted piece to move, but also allows the residues specified by N OFFSET and C OFFSET to move. DISTANCE CUTOFF is used to select the fixed atoms. A dihedral angle constraint ($\omega = 180^\circ$) is also set as in Level 1.

The user could achieve the same results by setting up the proper groups and imposing the desired constraints each time. The calculation levels are provided here to simplify this process.

User Input:

1. Toggle CALC LEVEL until the calculation level desired is shown (0, 1, or 2).

SIMULATE

The SIMULATE option allows the user to perform minimizations, dynamics, and conformational searches to modify the conformation of the structure created. This is described in Chapter 5. However, the movable and fixed groups are already defined (with the N OFFSET, C OFFSET, and DISTANCE CUTOFF options), so it is not necessary to set up the energy expression.

3.2.9 DELETE

The options on the DELETE menu shown in Figure 3.29 allow the user to delete amino acid residues from the peptide. The user specifies the residues to be deleted by picking a starting and a stopping residue on the display (using the PICK START RES and PICK STOP RES options). The number of the chosen residue appears beneath these items on the menu. The starting and stopping residues and all those in-between will be deleted when EXECUTE is picked. A long bond is drawn between the carbonyl carbon of the N-terminal residue and the nitrogen of the C-terminal residue to connect the remaining fragments. If the RENUMBER option was selected (toggled to YES), then the residues in the new structure will be renumbered to account for the deletion.

The SIMULATE option allows the user to perform minimizations, dynamics, and conformational searches to modify the conformation of the structure created. The remaining options (N OFFSET, C OFFSET and DISTANCE CUTOFF) control which atoms are allowed to move and which must remain fixed when doing these energy operations. These options can therefore be used to control the shape the new structure will take. N OFFSET and C OFFSET allow the user to specify the number of residues that are allowed to move from each end of the deletion (N OFFSET sets the value for the N-terminal, C OFFSET sets it for the C-terminal). DISTANCE CUTOFF sets all the atoms within the specified cutoff distance from any of the movable atoms to be fixed.

Delete:
pick start res
- - -
pick stop res
- - -
renumber
yes
execute
n offset=2
c offset=2
distance cutoff
0.000
simulate
return

PICK START RES, PICK STOP RES

The PICK START RES and PICK STOP RES options allow the user to pick the starting and stopping residues which define the range of amino acid residues to be deleted. The number of the residue chosen will appear below PICK START RES or PICK STOP RES on the menu. The starting and stopping residues and all those in between will be deleted when EXECUTE is picked.

User Input:

1. Pick PICK START RES or PICK STOP RES.
2. Pick the starting or stopping residue from the display.

RENUMBER

The RENUMBER option allows the user to renumber the residues in the peptide to account for the deletion. It is a YES/NO toggle, with the default being YES.

User Input:

1. Toggle RENUMBER until the YES/NO box beneath it is set as desired.

Figure 3.29:
DELETE menu

EXECUTE

The EXECUTE option causes the deletion of the residues specified under PICK START RES and PICK STOP RES to occur. A long bond is drawn between the carbonyl carbon of the N-terminal residue and the nitrogen of the C-terminal residue to connect the severed ends. A dihedral angle constraint is automatically placed so that the peptide bond remains *trans* ($\omega = 180^\circ$).

User Input:

1. Pick EXECUTE.

N OFFSET, C OFFSET

These options are described under INSERT (Section 3.2.8) above.

DISTANCE CUTOFF

The DISTANCE CUTOFF option is described under INSERT (Section 3.2.8) above.

SIMULATE

The SIMULATE option is described in Chapter 5.

3.2.10 EDIT, EDIT ID

The EDIT and EDIT ID options are used to edit the structure being built. They are described in Sections 3.1.12 and 3.1.13.

3.2.11 MECHANICS, SIMULATE

MECHANICS performs energy minimizations, dynamics, and conformational searches on the structures built under BUILD. The energy expression is automatically set up with all atoms movable. MECHANICS is described in Section 5.7. The SIMULATE option performs the same operations using general energy expressions and is described in Chapter 5.

3.2.12 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.3 RNA-DNA

The options on the RNA-DNA menu allow the user to build single, double, or triple-stranded DNA or RNA from fragments contained in fragment libraries. The RNA-DNA menu is shown in Figure 3.30. After setting up a file to hold the new structure, the user specifies the type of chain to be built by selecting the SINGLE, DOUBLE, or TRIPLE option. He then specifies the nucleotide sequence by picking from a menu listing the nucleotide fragments.

When using the SINGLE option, the menu contains the four normal DNA nucleotides for single-stranded chains: adenylic, guanylic, cytidylic, and thymidylic acid. The DOUBLE menu contains the DNA nucleotide *pairs* for double stranded chains. Each new fragment selected is added on to the growing chain at the 5' end. The helices constructed may be in the A or B form. The program does not allow mixing of A and B form nucleotides within the same chain. Initially a DNA chain is built. However, this can be changed to RNA (and vice versa) using the MODIFY SUGAR option. The TRIPLE option allows the user to build triple-stranded homopolymers. Seven polynucleotide triplets are supported.

Because the chains are built from fragments, a 5' phosphate is partially present at the end of each strand when the Builder terminates. A BAP (β alkaline phosphatase) option is therefore provided to clean up the end groups. BAP "cleaves" the free (partial) phosphate groups from the 5' ends. In addition, BAP generates heteroatom hydrogens (*i.e.*, polar hydrogens), which are needed for hydrogen bonding. After BAPing the structure, -OH groups will be present at both the 5' and 3' ends of the DNA and RNA strands, as expected.

Note: The BAP option *must* be done if planning to use the Molecular Simulations implementation of the AMBER force field. BAP converts the H3 type hydrogens on adenine, guanine, and cytosine amino nitrogens to H2, the atom type required by the program to do energy calculations. For more information, see Appendix O.

A USER option is also provided which allows the user to build DNA or RNA chains from fragments supplied in a user-defined library. This user library must be read in with the READ LIBRARY option in order to make these fragments available.

Note: The old DNA Builder may also be used to build double-stranded DNA (see DNA-OLD, Section 3.4). The structures created will be the same as those constructed using the RNA-DNA Builder. Since the sequence is entered from the keyboard, the old DNA Builder may actually build these structures faster. However, it cannot be used to build RNA, or any single or triple-stranded chains, and it cannot incorporate any user-defined fragments.

RNA-DNA:
setup
read library
single
double
triple
user
BAP
modify sugar
edit
edit id
mechanics
simulate
analyze
return

Figure 3.30:
RNA-DNA menu

3.3.1 SETUP

The SETUP option is automatically invoked each time the RNA-DNA Builder is entered. It allows the user to pick a new file for the structure or add to an existing file. If the file is new, the user enters a name for it. A group of the same name is also created.

User Input:

1. Pick SETUP (needed only if already in the RNA-DNA Builder and setting up again).
2. Pick the file to be operated on or pick NEW FILE if creating a new file. (This step is skipped if there are no files in the system.)
3. Enter a descriptor (up to 8 characters) for the file if the file is new.

3.3.2 READ LIBRARY

The READ LIBRARY option allows the user to read the directory information for one of two libraries selected from the READ LIBRARY menu: the BioDesign library of nucleotide fragments for building single, double and triple strands (BIODESIGN), or a user-defined library (USER) (see Figure 3.31). The directories containing the BioDesign-supplied fragments are automatically read in when the user first enters the RNA-DNA Builder. The only time the user would need to read these directories again would be if he had added *bgf* fragment files to them during a session. The user-defined directory *must* be read in, however, in order to allow access to the user-defined library.

Selections:
BioDesign
user
return

Figure 3.31:
Nucleotide
libraries menu

User Input:

1. Pick READ LIBRARY.
2. Pick the library to be read in (BIODESIGN or USER).
3. Enter the name of the directory (for the user-defined library). See Section 1.6 for details on entering file names.

3.3.3 SINGLE, DOUBLE, TRIPLE

These options allow the user to build single, double, or triple-stranded DNA or RNA from fragments contained in fragment libraries⁵. The helices constructed may be in the A or B form. The user specifies the type of chain to be built by selecting the SINGLE, DOUBLE, or TRIPLE option. He then specifies the nucleotide sequence by picking from a menu listing the nucleotide fragments available (see Figure 3.32). When using the SINGLE option, the menu contains the four normal DNA nucleotides: adenylic, guanylic, cytidylic, and thymidylic acid (dA, dG, dC, and dT) in both the A and B form. The DOUBLE menu contains the DNA nucleotide *pairs* for double stranded

⁵ Coordinates for the single and double stranded fragments are the same as those used for the old DNA builder and were taken from Arnott and Hukins "Optimized Parameters for A-DNA and B-DNA", *Biochem. and Biophys. Research Comm.*, 47 (1972) 1504-1510. Coordinates for the triple stranded fragments were taken from Arnott *et al*, "Models of triple stranded polynucleotides with optimized stereochemistry", *Nucleic Acids Research*, 3 (1976) 2459-2470 with the following caveats (see footnotes 2 and 3).

chains (dA-dT, dC-dG, dG-dC, and dT-dA) also in both the A and B form. Each new fragment selected is added on to the growing chain at the 5' end. The program does not allow mixing of A and B form nucleotides within the same chain. Initially a DNA chain is built. However, this can be changed to RNA (and vice versa) using the MODIFY SUGAR option.

The TRIPLE option allows the user to build triple-stranded homopolymers. Seven polynucleotide triplets are supported, three with ribose as the sugar (I-A-I, U-A-U-11, U-A-U-12), three with deoxyribose (dC-dI-dC-1, dC-dI-dC-2, dT-dA-dT), and one with both (U-dA-U). The TRIPLE menu is also shown in Figure 3.32. A description of the fragments listed is given in Appendix B.7.

<table><tr><td>Selections:</td></tr><tr><td>A_dA</td></tr><tr><td>A_dC</td></tr><tr><td>A_dG</td></tr><tr><td>A_dT</td></tr><tr><td></td></tr><tr><td>B_dA</td></tr><tr><td>B_dC</td></tr><tr><td>B_dG</td></tr><tr><td>B_dT</td></tr><tr><td>return</td></tr></table>	Selections:	A_dA	A_dC	A_dG	A_dT		B_dA	B_dC	B_dG	B_dT	return	<table><tr><td>Selections:</td></tr><tr><td>A_dA-dT</td></tr><tr><td>A_dC-dG</td></tr><tr><td>A_dG-dC</td></tr><tr><td>A_dT-dA</td></tr><tr><td></td></tr><tr><td>B_dA-dT</td></tr><tr><td>B_dC-dG</td></tr><tr><td>B_dG-dC</td></tr><tr><td>B_dT-dA</td></tr><tr><td>return</td></tr></table>	Selections:	A_dA-dT	A_dC-dG	A_dG-dC	A_dT-dA		B_dA-dT	B_dC-dG	B_dG-dC	B_dT-dA	return	<table><tr><td>Selections:</td></tr><tr><td>I-A-I</td></tr><tr><td></td></tr><tr><td>U-A-U-11</td></tr><tr><td></td></tr><tr><td>U-A-U-12</td></tr><tr><td></td></tr><tr><td>U-dA-U</td></tr><tr><td></td></tr><tr><td>dC-dI-dC-1</td></tr><tr><td></td></tr><tr><td>dC-dI-dC-2</td></tr><tr><td></td></tr><tr><td>dT-dA-dT</td></tr><tr><td>return</td></tr></table>	Selections:	I-A-I		U-A-U-11		U-A-U-12		U-dA-U		dC-dI-dC-1		dC-dI-dC-2		dT-dA-dT	return
Selections:																																							
A_dA																																							
A_dC																																							
A_dG																																							
A_dT																																							
B_dA																																							
B_dC																																							
B_dG																																							
B_dT																																							
return																																							
Selections:																																							
A_dA-dT																																							
A_dC-dG																																							
A_dG-dC																																							
A_dT-dA																																							
B_dA-dT																																							
B_dC-dG																																							
B_dG-dC																																							
B_dT-dA																																							
return																																							
Selections:																																							
I-A-I																																							
U-A-U-11																																							
U-A-U-12																																							
U-dA-U																																							
dC-dI-dC-1																																							
dC-dI-dC-2																																							
dT-dA-dT																																							
return																																							
DNA nucleotides menu	DNA nucleotide pairs menu	Polynucleotide triplets menu																																					

Figure 3.32: Nucleotide fragment menus for single, double, and triple-stranded chains

The helix symmetry used for I-A-I is 12 with a turn angle of 30.0°⁶. Two types of triplets for poly(U)·poly(A)·poly(U) are provided, one with a helix symmetry of 11 (U-A-U-11), and one

⁶ The helix symmetry in Arnott *et al*¹ for poly(I)·poly(A)·poly(I) is described as 12-fold, however, the "turn angle" given in the corresponding table is 32.7° which indicates 11-fold symmetry (i.e., 360/11 = 32.7). Since the pitch and residue height for this structure are given as 39.5 Å and 32.9 Å, respectively (which implies 12 units/turn - 39.5/3.29 = 12), the helix symmetry was taken as 12 and a turn angle of 30.0° was used.

with 12 fold symmetry (U-A-U-12). Two types of triplets are also provided for poly(dC)-poly(dI)-poly(dC), one with oxygens on the 2' position of the sugars (dC-dI-dC-1) and one without 2' oxygens (dC-dI-dC-2)⁷. The user selects the fragment the appropriate number of times to create the structure of the desired size. For example, to build a pentamer of poly(U)-poly(dA)-poly(U), select U-dA-U five times.

User Input:

1. Pick SINGLE, DOUBLE, or TRIPLE.
2. SINGLE or DOUBLE option: Select the desired nucleotide or nucleotide pair. Continue by picking the next fragment in the sequence until the chain is completed. Mixing of A and B form nucleotides within the same chain is not allowed.
TRIPLE option: Select the desired nucleotide triplet the appropriate number of times to create the structure of the desired size.
3. Pick RETURN to terminate.

3.3.4 USER

The USER option is similar to the SINGLE, DOUBLE, and TRIPLE options described above except that the DNA or RNA chains are built from nucleotide fragments contained in a user-defined fragment library. The fragments contained in the library are displayed in alphabetical order. The directory information for this library must be read in using READ LIBRARY before the fragments in this library are accessible (see Section 3.3.2).

User Input:

1. Pick USER.
2. Pick the fragments in the desired sequence.
3. Pick RETURN to terminate.

3.3.5 BAP

Because the RNA-DNA Builder constructs chains from fragments, a 5' phosphate is partially present at the end of each strand when the Builder terminates. The BAP (β alkaline phosphatase) option is therefore provided to clean up the end groups. BAP "cleaves" the free (partial) phosphate groups from the 5' ends. In addition, BAP generates heteroatom hydrogens (*i.e.*, polar hydrogens), which are needed for hydrogen bonding. After BAPing the structure, -OH groups will be present at both the 5' and 3' ends of the DNA and RNA strands, as expected.

Note: The BAP option *must* be done if you are planning to use the Molecular Simulations implementation of the AMBER force field. BAP converts the H3 type hydrogens on adenine, guanine, and cytosine amino nitrogens to H2, the atom type required by the program to do energy calculations. For more information, see Appendix O.

⁷ In Arnott *et al*¹, the structure poly(dC)-poly(dI)-poly(dC) has a name that implies 2' deoxy sugars. However, the corresponding table reports coordinates for the 2' oxygens. Hence, BIOGRAF provides both types of triplets, one with and one without the 2' oxygens.

User Input:

1. Pick BAP.

3.3.6 MODIFY SUGAR

The MODIFY SUGAR option allows the user to modify the sugars in a selected chain from deoxyribose to ribose or vice versa. After picking MODIFY SUGAR, the user selects the chain to be modified by picking one of the atoms along the backbone of the chain. He then selects the type of sugar desired (RIBOSE or DEOXYRIBOSE) from the menu. All the sugars in the selected chain are modified accordingly.

Selections:
ribose
deoxyribose
return

User Input:

1. Pick MODIFY SUGAR.
2. Select the chain to be modified by picking one of the atoms along the backbone of the chain.
3. Select the type of sugar desired (RIBOSE or DEOXYRIBOSE) from the menu.

Figure 3.33:
Sugar types menu

3.3.7 EDIT, EDIT ID

The EDIT and EDIT ID options are used to edit the structure being built. They are described in Sections 3.1.12 and 3.1.13.

3.3.8 MECHANICS, SIMULATE

MECHANICS is used to perform energy minimizations, dynamics, and conformational searches on the structures built under BUILD. The energy expression is automatically set up with all atoms movable. MECHANICS is described in Section 5.7. The SIMULATE option performs the same operations using general energy expressions and is described in Chapter 5.

3.3.9 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.4 DNA-OLD

The DNA-OLD option is the same as the DNA option which was available in previous releases of the program (version 2.0 and earlier). The name has been changed since a new function is now available (RNA-DNA) which not only builds double stranded DNA, but also builds RNA and can construct single, double, and triple stranded chains (see Section 3.3). The DNA-OLD option is retained since it will often construct double-stranded DNA faster than the new option, and thus is still useful.

The options on the DNA-OLD menu shown in Figure 3.34 allow the user to build a double-stranded DNA helix from the four normal DNA nucleotides: adenylic, guanylic, cytidylic, and thymidylic acid (A, G, C, and T). The helix may be in the A or B form. The user specifies the form of the helix and then enters the nucleotide sequence for each strand. The sequence for the first strand is entered in 5'-3' order, and the sequence for the opposite strand in 3'-5' order. This means that the first nucleotide entered for the first strand is complementary to the first nucleotide entered for the second strand, the second nucleotide on the first strand is complementary to the second nucleotide on the second strand, and so forth.

The user can optimize the conformation of the structure by using MECHANICS (see Section 5.7) or SIMULATE (see Chapter 5) or edit the structure as described in Sections 3.1.12 and 3.1.13. The ANALYZE functions are also available from this menu (see Chapter 6).

DNA old:
setup
edit
edit id
mechanics
simulate
analyze
return

Figure 3.34:
DNA-OLD menu

3.4.1 SETUP

The SETUP option is automatically invoked when the DNA Builder is entered. It allows the user to pick a new file for the structure or add to an existing file. If the file is new, the user enters a name for it. A group of the same name is also created.

User Input:

1. Pick SETUP (needed only if already in the DNA Builder and setting up again).
2. Pick the file to be operated on or pick NEW FILE if creating a new file. (This step is skipped if there are no files in the system.)
3. Enter a descriptor (up to 8 characters) for the file if the file is new.
4. Enter the type of DNA helix to be built (A or B).
5. Enter the upper strand sequence in 5'-3' order. Use only the letters A, G, C, T (upper or lower case).
6. Enter the lower strand sequence in 3'-5' order. Use only the letters A, G, C, T (upper or lower case).

3.4.2 EDIT, EDIT ID

The EDIT and EDIT ID options are used to edit the structure being built. They are described in Sections 3.1.12 and 3.1.13.

3.4.3 MECHANICS, SIMULATE

MECHANICS is used to perform energy minimizations, dynamics, and conformational searches on the structures built under BUILD. The energy expression is automatically set up with all atoms movable. MECHANICS is described in Section 5.7. The SIMULATE option performs the same operations using general energy expressions and is described in Chapter 5.

3.4.4 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.5 LIPID

The operations on the LIPID menu shown in Figure 3.35 allow the user to build lipids from combinations of lipid backbone molecules from the *bg_lipid_bck* directory, fatty acids from the *bg_lipid_fa* directory, and head groups from the *bg_lipid_hg* directory. After toggling LIPID the program reads the lipid fragments from the three directories. SETUP is automatically invoked and the user is prompted for the file to contain the lipid that will be built. A new file may be created or the lipid may be added to an existing file. If the file is new, the user enters a name for it. A group of the same name is also created.

The user then selects a backbone structure for the lipid by toggling the BUILD option and selecting one of the backbone molecules from the choices on the menu. Next the menu shows a selection of fatty acids. The user selects one of the fatty acids to extend the lipid structure. The next selection can be either a second fatty acid or a head group, depending on the backbone structure that was used to start the lipid. The user continues by adding the remaining fatty acids and a head group if appropriate.

The lipid may be modified using the EDIT and EDIT ID functions and its conformation may be optimized using the MECHANICS or SIMULATE operations. The user can also create a lipid monolayer or bilayer by using the EXTEND option. This option makes copies of the lipid structure that form a lattice in the xz plane. The lattice is either one layer thick along the y axis, forming a lipid monolayer, or two layers thick, forming a lipid bilayer.

Lipid:
setup
read library
build
extend
edit
edit id
mechanics
simulate
analyze
return

Figure 3.35:
LIPID menu

3.5.1 SETUP

The SETUP option is automatically invoked when the Lipid Builder is entered. It allows the user to pick a new file for the structure or add to an existing file. If the file is new, the user enters a name for it. A group of the same name is also created.

User Input:

1. Pick SETUP (needed only if already in the Lipid Builder and setting up again).
2. Pick the file to be operated on or pick NEW FILE if creating a new file. (This step is skipped if there are no files in the system.)
3. Enter a descriptor (up to 8 characters) for the file if the file is new.

3.5.2 READ LIBRARY

Toggling READ LIBRARY reads the directory information for the three lipid libraries. The lipid backbone molecules are read in from the *bg_lipid_bck* directory, fatty acids from the *bg_lipid_fa* directory, and head groups from the *bg_lipid_hg* directory. These three directories are automatically read in when the user first enters the Lipid Builder. The only time the user would need to

read one of these default directories again would be if the user had added *bgr* fragment files to it during a session.

User Input:

1. Pick READ LIBRARY to read the three lipid directories.

3.5.3 BUILD

The BUILD option is used to assemble a complex lipid molecule from the simpler fragments read from the *bgr* lipid libraries. The fragments are selected from the Lipid BUILD menus (see Figure 3.36). A list of the fragments is given in Appendix B. Fatty acids are joined to hydroxyl groups or amino groups on the backbone molecules to form esters or amides. A head group may be joined to the backbone group by a phosphoester, acetal or ether linkage.

Selections:
glycerol
p_glycerol
p_sphingosine
sphingosine
return

Lipid backbone
menu

Selections:
arachidate
arachidonate
behenate
laurate
lignocerate
linoleate
linolenate
myristate
oleate
palmitate
palmitoleate
stearate
return

Lipid fatty acids
menu

Selections:
choline
ethanolamine
glycerol
inositol
return

Lipid head groups
menu

Figure 3.36: Lipid BUILD menus

User Input:

1. Pick BUILD.
2. Pick the backbone group.

3. Pick a fatty acid group or head group.
4. Repeat the previous instruction until the fatty acid or head group menu is no longer displayed or pick RETURN if you wish to stop.

3.5.4 EXTEND

The EXTEND option allows the user to build a monolayer or bilayer composed of a number of copies of the lipid. The layer can be made as large as desired in the x direction (across the display) and the z direction (along the viewer's line of sight out from the display). It is one layer thick in the y direction (up and down) if the user chooses to create a monolayer and two layers thick if the user chooses to create a bilayer. The starting lipid molecule is rotated 180° about the z axis to produce the lipid molecules in the bottom layer of a bilayer. Thus the molecules in the bottom layer face the opposite way from the molecules in the top layer.

After picking EXTEND, the user selects the type of lipid structure he wishes to create (MONOLAYER or BILAYER) from the menu shown in Figure 3.37. The program automatically aligns the molecule so it has x, y, and z as its principal axes. The axes are also displayed. The EXTEND menu appears and the user is prompted to change the limits of the extension or the separation distance of the layers by toggling the items on this menu (see Figure 3.38).

The extension values displayed refer to the number of rows to be added in each direction (relative to the current structure):

MIN X	-	to the left	MAX X	-	to the right
MIN Z	-	in front of	MAX Z	-	behind

The current structure is considered to be at position 0 in each direction. Therefore, the values to the left and in front of the structure (MIN X and MIN Z) are always negative, while the values to the right and behind the structure (MAX X and MAX Z) are always positive. The default values specify that one row of the molecule be added in the x and z directions. To change these limits, the user toggles the MIN or MAX item(s) on, and then hits the + (plus) or - (minus) box above or below the items to increase or decrease the displayed value.

The spacing between the layers may be changed by specifying a new separation distance (the default value is 2.0 Å). This value determines the distance between the closest points of adjacent layers. To change the separation distance, the user toggles SEPARATION on, and then enters the

Selections:
monolayer
bilayer
return

Figure 3.37: Lipid structure-type menu

Extend:
+
min x = -1
min y = 0
min z = -1
-
+
max x = 1
max y = 0
max z = 1
-
separation
2.000
abort
execute

Figure 3.38: EXTEND menu

new value. The extended structure consists of individual molecules that are not covalently bonded. The CONNECT option on the main BUILD menu may be used to cross-link the units.

User Input:

1. Pick EXTEND.
2. Pick MONOLAYER or BILAYER.
3. To change the extension limits:
 - Toggle the MIN or MAX item(s) to be changed from the EXTEND menu.
 - Toggle the + (plus) or - (minus) box to increase or decrease the displayed value.
4. To change the separation distance between adjacent layers:
 - Toggle SEPARATION on.
 - Enter the desired separation distance in Å .
5. Adjust the orientation of the molecule until it has the correct orientation for extension. Typically this will be when one of the fatty acids is oriented along the y axis.
6. Pick EXECUTE to initiate the extension or pick ABORT to abort the procedure and return to the LIPID menu.

3.5.5 EDIT, EDIT ID

The EDIT and EDIT ID options are used to edit the structure being built. They are described in Sections 3.1.12 and 3.1.13.

3.5.6 MECHANICS, SIMULATE

MECHANICS is used to perform energy minimizations, dynamics, and conformational searches on the structures built under BUILD. The energy expression is automatically set up with all atoms movable. MECHANICS is described in Section 5.7. The SIMULATE option performs the same operations using general energy expressions and is described in Chapter 5.

3.5.7 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.6 CARBOHYDRATE

The options on the CARBOHYDRATE menu shown in Figure 3.39 allow the user to build complex polysaccharides from the monosaccharides in the *bg_carbo* directory. After toggling CARBOHYDRATE the program reads the monosaccharides from the *bg_carbo* directory. SETUP is automatically invoked and the user is prompted for the file to contain the structure that will be built.

The user can use the DEFINE and REPEAT functions to create a complex polysaccharide with a repeating structure. The DEFINE option allows the user to define the sequence of monosaccharides that is repeated, and the REPEAT option is used to specify how many copies of this sequence occur in the polysaccharide. The user can also use the LIBRARY function to add monosaccharides one at a time to the new molecule.

The LINKAGE option is used to specify the linkage atoms of successive monosaccharides and to get branch points in the polysaccharide. The DEFINE ANGLE and MODIFY CARBO functions can be used to create branch points in the polysaccharide and to modify the dihedral angles about the bonds that link successive monosaccharides. Modifying the dihedral angles in the glycosidic linkages with the MODIFY CARBO function is the standard method for changing the conformation of the polysaccharide.

3.6.1 SETUP

The SETUP option is automatically invoked when the Carbohydrate Builder is entered. It allows the user to pick a new file for the structure or add to an existing file. If the file is new, the user enters a name for it. A group of the same name is also created.

User Input:

1. Pick SETUP (needed only if already in the Carbohydrate Builder and setting up again).
2. Pick the file to be operated on or pick NEW FILE if creating a new file. (This step is skipped if there are no files in the system.)
3. Enter a descriptor (up to 8 characters) for the file if the file is new.

Carbohydrate:
setup
read library
linkage
1 ⇌ 4
define angle
modify carbo
carbo library
user library
define
repeat
extend
edit
edit id
mechanics
simulate
analyze
return

Figure 3.39:
CARBOHYDRATE
menu

3.6.2 READ LIBRARY

The READ LIBRARY option allows the user to read the directory information for one of two libraries selected from the Carbohydrate libraries menu: the carbohydrate library (CARBO LIBRARY) contained in the *bg_carbo* directory, or a user-defined library (USER LIBRARY) (see Figure 3.40). The *bg_carbo* directory is automatically read in when the user first enters the Carbohydrate Builder. The only time the user would need to read this directory again would be if the user had added *bgf* fragment files to it during a session. The user-defined directory *must* be read in, however, in order to allow access to the user-defined library.

Selections:
carbo library
user library
return

Figure 3.40:
Carbohydrate
libraries menu

User Input:

1. Pick READ LIBRARY.
2. Pick the library to be read in (CARBO LIBRARY or USER LIBRARY).
3. Enter the name of the directory (for the user-defined library). See Section 1.6 for details on entering file names.

3.6.3 LINKAGE

The LINKAGE option is used to pick branch points on the growing polysaccharide and to define or change the linkage carbons on successive monosaccharides. The selected linkage is displayed beneath LINKAGE on the CARBOHYDRATE menu.

User Input:

1. Pick LINKAGE.
 2. Pick YES on the adjacent menu to show that you want to pick a new branch point or pick NO to show that you want to define a new linkage.
 3. If YES (picking a new branch point):
 - a) Pick the atom to which the next monosaccharide will be added. You must pick an oxygen atom that can be part of normal glycosidic linkages.
 - b) Pick the number of the second atom of the new linkage from the menu. This number is the atom number of the linkage carbon atom on the next monosaccharide to be added to the molecule.
- If NO (defining a new linkage):
- a) Pick the number of the first atom of the new linkage. This number is the number of the linkage carbon atom on the most recently added monosaccharide unit of the existing polysaccharide.
 - b) Pick the number of the second carbon atom of the new linkage from the menu. This is the number of the linkage carbon atom on the next monosaccharide to be added.

3.6.4 DEFINE ANGLE

The DEFINE ANGLE function shown in Figure 3.41 allows the user to determine the dihedral angles about the bonds connecting successive monosaccharide units. There are two dihedral angles that can be set: D1, the dihedral angle about the bond between the first monosaccharide and the oxygen of the glycosidic linkage, and D2, the dihedral angle about the bond between the oxygen of the glycosidic linkage and the next monosaccharide. The MODIFY CARBO option is used to implement the changes specified under DEFINE ANGLE.

User Input:

1. Pick DEFINE ANGLE from the CARBOHYDRATE menu.
2. Pick D1 to set the first dihedral angle of the glycosidic linkage or D2 to set the second dihedral angle of the glycosidic linkage, or toggle RESET to set the values for both angles back to 180°.
3. Enter the new value for the dihedral angle.

Define angle:
reset
D1=180.00
D2=180.00
return

Figure 3.41:
DEFINE ANGLE
menu

3.6.5 MODIFY CARBO

The MODIFY CARBO menu shown in Figure 3.42 allows the user to modify the dihedral angles to the settings he has chosen with the DEFINE ANGLE function. The entire polysaccharide can be changed to use the new dihedral angles, or the user can select a part of the molecule to use the new angles. If the user wants to modify only part of the molecule, he must select a starting and ending monosaccharide unit to be modified. The dihedral angles from the starting monosaccharide to the ending monosaccharide unit are modified. Thus the conformation of the linkage between the ending unit and the next unit after it is not changed.

The MODIFY CARBO function displays a menu with the options DEFINE ANGLE and MODIFY DIHEDRAL. The DEFINE ANGLE option is same as the DEFINE ANGLE function of the main CARBOHYDRATE menu and is described in the previous section. The MODIFY DIHEDRAL option is used to make the actual changes to the molecule.

User Input:

1. Pick MODIFY CARBO from the CARBOHYDRATE menu.
2. Pick DEFINE ANGLE to specify the dihedral angles or pick MODIFY DIHEDRAL to change the conformation of the polysaccharide.
3. If you have picked MODIFY DIHEDRAL, pick YES to modify the glycosidic bonds of the entire polysaccharide, or pick NO to modify only part of the polysaccharide.
4. If modifying only part of the polysaccharide, pick the starting monosaccharide unit to be modified, then pick the last monosaccharide unit to be modified.

Modify carbo:
define angle
modify dihedral
return

Figure 3.42:
MODIFY CARBO
menu

3.6.6 CARBO LIBRARY

The CARBO LIBRARY option allows the user to select monosaccharides contained in the *bg_carbo* directory. The monosaccharides are displayed in alphabetical order as shown in Figure 3.43 (see Appendix B.5 for a complete listing). Selecting a fragment from the menu loads it into a file and displays it on the screen. The fragments are displayed at pseudo-random locations and orientations.

User Input:

1. Pick CARBO LIBRARY.
2. Pick the desired monosaccharides from the Monosaccharides menu.
3. Pick RETURN to complete the selection.

Selections:		
ad_2dribose	ad_ribose	bd_glucose
ad_allose	ad_sorbose	bd_gulose
ad_altrose	ad_tagatose	bd_idose
ad_arabinose	ad_talose	bd_lyxose
ad_fructose	ad_xylose	bd_mannosamn
ad_galactose	al_fucose	bd_mannose
ad_galactsamn	al_rhamnose	bd_mansa_nac
ad_galnac	al_sorbose	bd_psicose
ad_glcamn_nac	bd_2dribose	bd_ribose
ad_glucosamn	bd_allose	bd_sorbose
ad_glucose	bd_altrose	bd_tagatose
ad_gulose	bd_arabinose	bd_talose
ad_idose	bd_fructose	bd_xylose
ad_lyxose	bd_galactose	bl_fucose
ad_mannosamn	bd_galactsamn	bl_rhamnose
ad_mannose	bd_galnac	bl_sorbose
ad_mansa_nac	bd_glcamn_nac	
ad_psicose	bd_glucosamn	return

Figure 3.43: Monosaccharides menu

3.6.7 USER LIBRARY

See Section 3.2.3 for a description of the USER LIBRARY option.

3.6.8 DEFINE

The DEFINE function allows the user to define a monomeric unit composed of one or more monosaccharides from the *bg_carbo* library. The user can select and modify the monomeric unit as long as the DEFINE toggle is on. After toggling DEFINE off, the user can copy the monomeric unit an arbitrary number of times by using the REPEAT function. Note that the monomeric unit used by the REPEAT function can be an oligosaccharide.

User Input:

1. Toggle DEFINE on.
2. Use the CARBO LIBRARY and USER LIBRARY functions to select fragments to form the monomer, or other functions to modify the monomer.
3. When the monomeric unit is complete, toggle DEFINE off then use the REPEAT option to build the polysaccharide (see REPEAT, Section 3.6.9 below).

3.6.9 REPEAT

The REPEAT option allows the user to create a polysaccharide composed of an arbitrary number of copies of the monomeric unit defined with the DEFINE option described above.

User Input:

1. Pick REPEAT on the CARBOHYDRATE menu.
2. Enter the number of copies of the monomer to include in the polysaccharide.

3.6.10 EXTEND

The EXTEND option allows the user to build a lattice composed of a number of copies of the polysaccharide that has been constructed with the fragments in the user-defined or carbohydrate libraries. The lattice can extend in one, two, or three directions and can be made as large as desired. When the user picks EXTEND, the program automatically aligns the molecule so it has x, y, and z as its principal axes. The axes are displayed on the screen. The EXTEND menu appears and the user is prompted to change the limits of the extension or the separation distance of the units by toggling the items on this menu (see Figure 3.44).

The extension values displayed refer to the number of rows to be added in each direction (relative to the current structure):

MIN X	- to the left	MAX X	- to the right
MIN Y	- below	MAX Y	- above
MIN Z	- in front of	MAX Z	- behind

The current structure is considered to be at position 0 in each direction. Therefore, the values to the left, below, and in front of the structure (MIN X, Y, and Z) are always negative, while the values to the right, above, and behind the structure (MAX X, Y, and Z) are always positive. The default values specify that one row of the molecule be added in each direction (a 3 x 3 x 3 lattice will be constructed). To change these limits, the user toggles the MIN or MAX item(s) on, and then hits the + (plus) or - (minus) box above or below the items to increase or decrease the displayed value.

The spacing between the units may be changed by specifying a new separation distance (the default value is 2.0 Å). This value determines the distance between the closest points of adjacent molecules. To change the separation distance, the user toggles SEPARATION on, and then enters the new value.

The extended lattice consists of individual polysaccharide units that are not covalently bonded. The CONNECT option on the main BUILD menu may be used to cross-link the units.

User Input:

1. Pick EXTEND.
2. To change the extension limits:
 - Toggle the MIN or MAX item(s) to be changed.
 - Toggle the + (plus) or - (minus) box to increase or decrease the displayed value.
3. To change the separation distance between adjacent chains:
 - Toggle SEPARATION on.
 - Enter the desired separation distance in Å on the host terminal.
4. Adjust the orientation of the molecule until it has the correct orientation for extension.

Extend:
+
min x = -1
min y = -1
min z = -1
-
+
max x = 1
max y = 1
max z = 1
-
separation
2.000
abort
execute

Figure 3.44:
EXTEND menu

5. Pick EXECUTE to initiate the extension or pick ABORT to abort the procedure and return to the CARBOHYDRATE menu.
6. Pick anywhere to display the extended structure.

3.6.11 EDIT, EDIT ID

The EDIT and EDIT ID options are used to edit the structure being built. They are described in Sections 3.1.12 and 3.1.13.

3.6.12 MECHANICS, SIMULATE

MECHANICS performs energy minimizations, dynamics, and conformational searches on the structures built under BUILD. The energy expression is automatically set up with all atoms movable. MECHANICS is described in Section 5.7. The SIMULATE option performs the same operations using general energy expressions and is described in Chapter 5.

3.6.13 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.7 CRYSTAL

The CRYSTAL option allows the user to create crystallographic structures. This option is only available for users who do not also have POLYGRAF. (POLYGRAF uses a Crystal Builder which is identical to this one except that it also does energy calculations on periodic systems).

When the user first enters the Crystal Builder, the SETUP option is automatically invoked. SETUP sets up a file to hold the new data. Currently, the Crystal Builder requires that the atomic coordinates be placed in the first file. If there are no files in the system, a new one will be created. However, if a file is already present, the user may choose to work with the current file (pick CURRENT when doing SETUP) or delete the current file and start a new one (pick RESET). If starting a new file, the user enter a name for it.

The CRYSTAL menu then appears as shown in Figure 3.45. The operations on this menu allow the user to enter and display crystallographic data based on the space group, the dimensions of the unit cell, and the coordinates of the unique atoms. The atoms can be represented with symbols of various types, sizes and color.

The user can read in the crystallographic data using the INPUT FILE option or he may choose INPUT COORD and enter the data directly. If entering the data directly he may first use DEFAULTS to set the type, size, and color, of the symbols that will be used to represent the atoms on the screen. INPUT COORD is then used to specify the unit cell parameters and the coordinates of the unique atoms in the unit cell. After the space group name and dimensions of the unit cell have been entered, the program generates and displays the unit cell on the screen. The coordinate information for the unique atoms is then entered along with some optional parameters (force field type or element name, charge, color, and symbol type). The coordinates of the symmetry-equivalent atoms are generated and displayed automatically by the program.

Once the unit cell and its atoms have been constructed, it can be extended arbitrarily in three dimensions using EXTEND. One may bond atoms across the unit cell boundaries with CONNECT. EXTRACT MOLEC may be used to extract separate molecules from the crystal and create a regular non-periodic file. The CHARGES option may be used to specify charges, and EDIT and EDIT ID may be used at any time to modify the structure. The ANALYZE functions are also available to identify atoms, calculate distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure volumes, check proteins for irregular structures, match structures by least squares fits, plot trajectory file data, and change and merge plot files.

Crystal:
setup
input file
input coord
extend cell
unextend cell
connect
extract molec
defaults
charges
space group
edit
edit id
analyze
return

Figure 3.45:
CRYSTAL menu

Note: Currently, the Crystal Builder requires that the atomic coordinates be placed in the first file. If there are no files in the system, a new one will be created. However, if a file is already present, the user may choose to work with the current file (pick CURRENT when doing SETUP) or delete the current file and start a new one (pick RESET).

3.7.1 SETUP

The SETUP option allows the user to set up a file to hold the crystallographic data. Currently, the Crystal Builder requires that the atomic coordinates be placed in the first file. If there are no files in the system, a new one will be created. However, if a file is already present, the user may choose to work with the current file (pick CURRENT from the adjacent menu) or delete the current file and start a new one (pick RESET). If a new file is being created, the user is prompted for a descriptor for it.

Selections:
reset
current
return

Figure 3.46:
Reset/current
menu

Note: If you don't want to work with the current file, but wish to save it, pick CURRENT, pick RETURN to get out of the Crystal Builder, then write out the file using the IN-OUT functions from the main menu.

User Input:

1. If any files are currently in the system, pick RESET from the adjacent menu to delete the current file(s) and start a new one or pick CURRENT to work with the current file.
2. If creating a new file, enter a descriptor for it (up to 14 characters).

3.7.2 INPUT FILE

The INPUT FILE option is used to read in crystallographic data from a cell input file. The user enters the name of the input file in the dialog box. He may either select the desired file from the listings in the browser box or he may type in the file name directly via the New Dir and New File buttons (see Section 1.6 for details on entering file names).

The input file must be in the proper (.cel) format. The INPUT COORD option creates a *save.cel* file in the proper format which can be input as is or edited if desired (see Section 3.7.3 below).

The required format is as follows. The first line consists of the word 'CRYSTAL' (must be uppercase). This is followed by an arbitrary number of comment lines, each beginning with the comment indicator '!'. The remaining lines contain the unit cell parameters and coordinate information. The same sequence of parameters is used as when inputting the data with the keyboard. The space group name is entered on the next line (the internal number must also be included for space groups with more than one case, see Appendix M). This is followed by the unit cell dimensions (lengths - A, B, C in Å, and angles - α , β , γ in degrees) on the following line. One or more spaces must be used to separate each of the items entered. The next line contains all the information for the first unique atom to be added. The atom label is first, followed by the x, y, and z fractional coordinates. The remaining items on this line are optional and may be entered in any order; these are the force field type or element name, charge, color, and symbol type. All the parameters

for each unique atom must be on the same line and again, one or more spaces must be used to separate each of the items entered. The information for each unique atom added is entered on a new line. Blank lines may be placed between the descriptions for different unique atoms. When saving the file, use *cel* for the extension.

The following example is an input file for a NaCl crystal:

```
CRYSTAL
! Input file for NaCl crystal.
Fm3m      267
5.0      5.0      5.0      90.0      90.0      90.0
Na      0.0      0.0      0.0      Na      red      6-point      1.0

Cl      0.5      0.0      0.0      Cl      -1.0
```

Note: The program can also read data in the old Unit Cell file format.

3.7.3 INPUT COORD

The INPUT COORD option allows the user to specify the unit cell parameters and the coordinates of the unique atoms in the unit cell (the coordinates of the symmetry-equivalent atoms are generated and displayed automatically by the program). The user first enters the name of the space group and the dimensions (lengths - A, B, C and angles - α , β , γ) of the unit cell. (See the list of 230 possible space groups in Appendix M) For some space groups the user also must identify the unique axis of the cell, or select the origin of the unit cell or the geometry of the axes.

After the user has entered the cell parameters, the program generates and displays the unit cell. To help identify the a, b, and c axes, four corners of the unit cell are labeled with their a, b, and c values (in that order). Fractional coordinates are used, so the values are zero or one. The origin (zero point for all three directions) is indicated as "000", the corner labeled "100" indicates the positive direction for the a axis, while the "010" and "001" corners indicate the positive direction along the b and c axes respectively. Two groups are also created: the first one displays the unit cell and is labeled UNIT CELL; the second one will contain the unit cell atoms and is labeled UC ATOMS.

The coordinate information of the unique atoms may be entered directly from the keyboard, or it may be read in from an input file. If using the keyboard, the user first enters an atom label. The atomic coordinates are then entered in terms of x, y, and z fractional coordinates, *i.e.*, fractions of the unit cell distances (A, B, and C). There are four additional items which are optional and may be entered in any order. These are the force field type or element name, charge, color, and symbol type. The atom types for the Dreiding-II force field are listed in Figure 3.47. Colors are specified by name (see Figure 3.52 for a list of possible colors). The symbol types used to represent the atoms are also specified by name. One or more spaces must be used to separate each of the items entered. If nothing is entered for the optional parameters, the current values are automatically applied. All data entered from the keyboard is saved in a file (*save.cel*) which may be edited later and read in again using the INPUT FILE option.

User Input:

1. Pick INPUT COORD.
2. Enter the space group name (see Appendix M). If the program prompts for a case, enter the number 1 or 2 to determine the origin, unique axis, or choice of axes for the unit cell.
3. Enter the unit cell dimensions (lengths - A, B, C in Å, and angles - α , β , γ in degrees).
4. Enter the atom label (1 to 5 characters), x, y, and z fractional coordinates, and any other parameters (optional): force field type (1 to 5 characters) or element name (1 or 2 characters), charge, color, and symbol type. A list of possible force field types for the Dreiding-II force field is shown in Figure 3.47. Use one or more spaces to separate the items. Pick OK to complete the entry. Cancel aborts the entry.
5. Repeat step 4 for each unique atom added.

Note: If INPUT COORD is toggled again to enter additional unique atom data, the program uses the current space group information and skips prompts 2 and 3. If a different space group is desired, it is necessary to create a new file (use SETUP with the RESET option).

Selections:		
H_	N_3	Cl
H_A	N_22	Ga3
H_b	N_21	Ge3
B_3	N_2	As3
B_2	N_R2	Se3
C_34	N_R1	Br
C_33	N_R	In3
C_32	N_1	Sn3
C_31	O_32	Sb3
C_3	O_31	Te3
C_22	O_3	I_
C_21	O_2	Na
C_2	O_R1	Ca
C_R2	O_R	Ti
C_R1	O_1	Fe
C_R	F_	Zn
C_11	Al3	Tc
C_1	Si3	Ru
N_33	P_3	
N_32	S_31	
N_31	S_3	return

Figure 3.47:
Atom types menu (Dreiding-II force field)

3.7.4 EXTEND CELL

The **EXTEND CELL** option allows the user to build a lattice composed of a number of copies of the unit cell. The lattice can extend in the a, b, and c directions and can be made as large as desired. When the user picks **EXTEND CELL**, the menu shown in Figure 3.48 appears and the user is prompted to change the limits of the extension by toggling the items on this menu.

The extension values displayed refer to the number of rows to be added in each direction (relative to the current structure). The current structure is considered to be at position 0 in each direction. **MIN A**, **MIN B**, and **MIN C** are the number of rows to be added in the negative a, b, and c directions, respectively; these values are always negative. **MAX A**, **MAX B**, and **MAX C** are the number of rows to be added in the positive a, b, and c directions respectively; these values are always positive. The default values specify that one row of unit cells be added in each direction (a 3 x 3 x 3 lattice will be constructed). To change these limits, the user toggles the **MIN** or **MAX** item(s) on, and then hits the + (plus) or - (minus) box above or below the items to increase or decrease the displayed value.

When satisfied with the new values, the user picks **EXECUTE**. When the extension is complete, two new groups are created: **EXTENDED CELLS** which shows all the extended cells, and **EXTENDED ATOMS** which shows the atoms within all the extended cells. These groups are shown in Figure 3.49.

The extended crystal consists of individual unit cells that are not covalently bonded. The **CONNECT** option may be used to cross-link the units.

User Input:

1. Pick **EXTEND CELL**.
2. To change the cell limits:
 - a) Toggle the **MIN** item(s) or **MAX** item(s) to be changed from the **EXTEND CELL** menu.
 - b) Toggle the + (plus) or - (minus) box to increase or decrease the displayed value.
3. Pick **EXECUTE** to initiate the extension or pick **ABORT** to abort the procedure and return to the **CRYSTAL** menu.

3.7.5 UNEXTEND CELL

The **UNEXTEND CELL** option is used to undo the **EXTEND CELL** operation, that is, to remove the extended cells and extended atoms.

Extend cell:
+
min a = -1
min b = -1
min c = -1
-
+
max a = 1
max b = 1
max c = 1
-
abort
execute

Figure 3.48:
EXTEND CELL
menu

Display:
unit cell
uc atoms
extended cells
extended atoms
return

Figure 3.49:
Display groups
menu

User Input:

1. Pick UNEXTEND CELL to undo the EXTEND CELL operation.

3.7.6 CONNECT

The CONNECT option is used to make bonds between atoms. CONNECT can be used for extended or non-extended systems and can make bonds between atoms in adjacent unit cells. The menu is shown in Figure 3.50. It functions in the same way as the CONNECT option on the main BUILD menu (see Section 3.14).

3.7.7 EXTRACT MOLEC

The EXTRACT MOLEC option is used to extract separate molecules from a crystal and create a regular non-periodic file. This is useful in getting a single structure from x-ray crystal data. The user picks an atom from the molecule to be saved. All the atoms connected to it will also be saved. As the molecule is picked, it is deleted from the screen. The user may continue to pick additional molecules in the same manner. The selection process is terminated by picking any point on the screen other than an atom. The user enters a name for the file which will hold the extracted data. This new file will appear in the Files menu. A group of the same name is also created.

User Input:

1. Pick EXTRACT MOLEC.
2. Pick an atom on the screen from the molecule to be saved.
3. To save another molecule, repeat step 2.
4. Terminate the selection process by picking any point other than an atom.
5. Enter a descriptor for the new file. Click OK to complete the entry or Cancel to abort.

Connect:
setup
connect all
all cells
min distance
0.800
max distance
2.000
criteria
bond radii
scale
1.200
execute
delete bonds
return

Figure 3.50:
CONNECT menu

3.7.8 DEFAULTS

The DEFAULTS option allows the user to specify the type, size, and color of the symbols which will be used to represent the next atoms to be added to the unit cell. He may also specify whether the atoms will be connected by bonds and if so, whether the half-bond colors will be used. The charge assigned to the next atoms can also be specified and outlying atoms may be repositioned within the boundaries of the unit cell. The options used to implement these functions are shown in the DEFAULTS menu (see Figure 3.51).

HALF BOND

The HALF BOND option allows the user to determine whether the bonds connecting the atoms will be displayed in all one color or in a half-bond representation. HALF BOND only applies if the CONNECT option described below is toggled on. If HALF BOND is on, then each atom will be assigned a CPK color and the color of each heteronuclear bond will change color halfway (note that carbons are aqua not black). If HALF BOND is off, then the bonds will all be one color. The default is on.

User Input:

1. Toggle HALF BOND on or off.

CONNECT

The CONNECT option is an on/off toggle which allows the user to connect the atoms being added to any other atoms (including each other) as long as they meet the criteria specified (see the CONNECT option described in Section 3.14). If the HALF BOND option is on, the bonds connecting the atoms will be displayed with the half-bond colors; otherwise, one color will be used.

User Input:

1. Toggle CONNECT on to have the atoms connected or toggle it off to leave them unconnected. The default is off.

SYMBOL

Toggling SYMBOL cycles through the symbol types which may be used to represent the next atoms added to the unit cell (these are displayed in the box below SYMBOL on the menu). The allowed symbols are DOTS, tetrahedron (TETRAHEDRON), octahedron (OCTAHEDRON), 6-point star (6-POINT), 12-point star (12-POINT) and 18-point star (18-POINT). The default is 6-POINT.

Defaults:
half bond
connect
symbol
6-point
scale by
0.100
color by
atom
charge
0.000
color
into cell
return

Figure 3.51:
DEFAULTS menu

User Input:

1. Toggle SYMBOL until the desired symbol type is displayed in the menu box.

SCALE

The SCALE option allows the user to specify the size of the symbols in Å.

User Input:

1. Pick SCALE.
2. Enter the scale factor in Å (default is 0.10 Å).

COLOR BY ATOM/ELEMENT/FORCE FIELD

The COLOR BY ATOM/ELEMENT/FORCE FIELD option is a three-item toggle which allows the user to specify that each unique atom added will have a different color (ATOM), or that the next symbols will be colored according to element type (ELEMENT), or force field type (FORCE FIELD). If COLOR BY ATOM is on, each unique atom added will be assigned a different color (the next color). If COLOR BY ELEMENT is on and the element type of the next atoms is different, then the next color will be assigned to them. Similarly, if COLOR BY FORCE FIELD is on and the force field type of the next atoms is different, then the next color will be assigned to them. The next color is always the second color down the color list starting from the current color (see Figure 3.52). The user may change the next color using the COLOR option described below. The default is to COLOR BY ATOM.

User Input:

1. Toggle COLOR BY to change the coloring method.

CHARGE

The CHARGE option is used to change the charge value which will be used for the next unique atom added. The current value is displayed beneath CHARGE on the DEFAULTS menu. The default value is 0.000.

User Input:

1. Pick CHARGE.
2. Enter the new charge value.

COLOR

The COLOR option allows the user to specify the next color which will be used when another unique atom is added (if coloring by ATOM), or when the atoms being added are of a different

Colors:
purple
magenta
pink
red
orange
yellow
chartreuse
green
aqua
cyan
light blue
blue
white
RGB value
return

Figure 3.52:
Colors menu

element type (if coloring by ELEMENT) or force field type (if coloring by FORCE FIELD). A new menu is displayed which lists the possible colors (see Figure 3.52).

User Input:

1. Pick COLOR.
2. Pick a color from those listed on the menu or pick RGB VALUE and use the sliders to specify the exact color (see Section 4.5 for more information on using RGB VALUE).
3. Pick RETURN to verify the selection.

INTO CELL

INTO CELL is an on/off toggle. When on, any atoms that would have been positioned outside the boundaries of the unit cell will be placed inside. The position of each atom is determined from its periodic image. The default is on.

User Input:

1. Toggle INTO CELL on if you want atoms that would have been positioned outside the boundaries of the unit cell to be placed inside. Toggle INTO CELL off if this is not wanted.

3.7.9 CHARGES

The CHARGES option allows the user to reassign atomic charges. It is described under EDIT ID in Section 3.1.13.

3.7.10 SPACE GROUP

The SPACE GROUP menu item is merely a label. The current space group name is displayed in the menu box beneath it after it has been input by the user.

3.7.11 EDIT, EDIT ID

The EDIT option is used to add and delete atoms, to alter connections between atoms, and to modify the geometry of the system. This option is described in Section 3.1.12. EDIT ID is used to modify the labels, force field types, and atomic charges and is described in Section 3.1.13.

3.7.12 ANALYZE

The ANALYZE option allows the user to identify atoms and linkages, calculate bond distances, angles and planes, analyze energies and forces, and check for "bad" van der Waals contacts. One may also measure molecular volumes, check proteins for irregular structural features, and match structures by least squares fits. Functions are also provided to analyze and plot trajectory file data, and display, merge, and change the appearance of plot files. The analyze functions are described in Chapter 6.

3.8 SOLVATE

The SOLVATE operations allow the user to construct solvation shells of water or organic solvent around and within a molecule. The positions of the solvent molecules can be subsequently optimized, or the solvent molecules can be used in molecular dynamics calculations (see Chapter 5). The solvent molecules are placed on a grid built around the structure being solvated. The grid may be cubic, tetragonal, diamond shaped or hexagonal. The user can vary the spacing of the grid, how closely it approaches the molecule being solvated, and how far it extends from the molecule. These operations are implemented using the options on the SOLVATE menu shown in Figure 3.53.

The SETUP option is used to select which groups are to be solvated and which groups are to be used for contact exclusions. The grid of solvent molecules is generated around atoms in the groups specified to be solvated. In order to get rid of solvent molecules which may overlap with atoms in the solvated groups, the user specifies the groups to be used for contact exclusions. Solvent molecules that would be too close to atoms in these groups are deleted. One would generally want to exclude the groups which are being solvated; an exclude self (EXCLUDE SELF) toggle is therefore provided which, when on, automatically does this. If there is more than one group in the system, the user is prompted to select any additional groups to be excluded. The user also chooses a file that will hold the coordinates for the solvent molecules. The solvent information can be added to an existing file or to a new one.

After completing the SETUP, the user sets the solvation parameters. This is done with the next seven options on the menu. SOLVENT is used to choose the solvent type. The solvent molecules can be water, water with no hydrogens, or any of a number of organic solvents. The type to be used is chosen from a menu listing the available solvent types. GRID TYPE is used to select the type of grid that will be used (cubic, tetragonal, diamond or hexagonal). GRID SPACE sets the spacing of the grid, INNER CUTOFF determines how close the grid comes to the molecule being solvated, and OUTER CUTOFF determines how far out it extends. If the AUTO GROUP toggle is on, then a group will be created to hold the solvent and if HALF BOND is on, then the solvent molecules will be displayed with the half-bond colors. The default settings use water as the solvent with a diamond shaped grid spaced 2.8 Å apart (twice the van der Waals radius of crystalline water). This value is also used for the inner cutoff distance, while 5.6 Å is the outer cutoff default. The default settings automatically create a group and use the half-bond colors.

EXECUTE generates the solvent shell. If AUTO GROUP is on, a group is also created and the solvent molecules are displayed. The solvent group can be toggled on or off as desired using the global DISPLAY button.

Solvate:
setup
solvent
water
grid type
diamond
grid space
2.800
inner cutoff
2.800
outer cutoff
5.600
box solvent
auto group
half bond
exclude self
execute
return

Figure 3.53:
SOLVATE menu

Warning: AUTO GROUP must be on for the solvent molecules to be visible on the display. Otherwise, the MAKE GROUPS function can be used to create a group for them so that they may be visible (see Section 4.1).

User Input:

1. Pick SOLVATE.
2. Make sure EXCLUDE SELF is on unless you're not excluding the solvated groups (see EXCLUDE SELF, Section 3.8.10).
3. Use SETUP to select the groups to be solvated, the groups to be used for contact exclusions, and the file that will hold the solvent (see SETUP, Section 3.8.1 below).
4. Set the solvation parameters to the desired settings (see SOLVENT, GRID TYPE, GRID SPACE, INNER CUTOFF, OUTER CUTOFF, BOX SOLVENT, AUTO GROUP and HALF BOND in Sections 3.8.2 - 3.8.9 below).
5. Pick EXECUTE to generate the solvent molecules.
6. To create additional solvent shells with different parameter settings repeat steps 2 - 5.

3.8.1 SETUP

The SETUP option is used to select which groups are to be solvated and which groups are to be used for contact exclusions. The grid of solvent molecules is generated around atoms in the groups specified to be solvated. In order to get rid of solvent molecules which may overlap with atoms in the solvated groups, the user specifies the groups to be used for contact exclusions. Solvent molecules that would be too close to atoms in these groups are deleted. One would generally want to exclude the group which is being solvated; an exclude self (EXCLUDE SELF) toggle is therefore provided which, when on, automatically does this. If there is more than one group in the system, the user is prompted to select any additional groups to be excluded. The groups are selected from the Groups menu. The user then chooses the file that will hold the coordinates for the solvent molecules from the Files menu. He can add the solvent information to an existing file or save it in a new one. If a new one is picked, he enters a descriptor for it. This name will then appear in the Files menu.

Note: If the user changes the EXCLUDE SELF toggle, then SETUP must be done again before picking EXECUTE, since the positions of the atoms to be excluded are calculated when the groups are setup.

User Input:

1. Pick SETUP.
2. If there is only one group in the system, skip to step 3. (This group is automatically selected as the group to be solvated.)
If there is more than one group in the system:
 - a) Pick the groups to be solvated from the Groups menu. Terminate the selection by picking RETURN.
 - b) Pick the groups to be used for contact exclusions from the Groups menu. If EXCLUDE SELF is on, the groups specified in 2a are automatically selected; if any additional

groups are to be used for contact exclusions, they must be selected. Terminate the selection by picking RETURN.

3. Pick the file that will hold the solvent from the Files menu.
4. If a new file is picked, enter a name for it.

3.8.2 SOLVENT

The SOLVENT option is used to choose the solvent type. This is selected from a menu which appears listing the solvent types available (see Figure 3.54). The solvent molecules can be water, water with no hydrogens, or any of a number of organic solvents. Water (shown on the menu as WATER) has an oxygen atom containing implicit hydrogens (atom type O_32) and type H__A hydrogen atoms. Water with no hydrogens is shown on the menu as WATER_NO_H. It consists of the oxygen atom used for ordinary water without any hydrogen atoms. It cannot form hydrogen bonds. The default solvent type is water.

User Input:

1. Pick SOLVENT.
2. Select the solvent type from the menu.

Selections:
acetone
benzene
cyclohexane
cyclopentane
dimethylether
dmf
ethanol
ether
formamide
hexane
hexfisopropanol
methanol
nmp
pentane
phenol
toluene
water
water_no_H
return

Figure 3.54:
Solvents menu

3.8.3 GRID TYPE

The GRID TYPE option is used to select the type of grid which will be used to hold the solvent molecules: cubic, hexagonal, diamond shaped, or tetragonal. The user selects one of these grid types from those listed in the menu (see Figure 3.55). The default GRID TYPE is diamond.

User Input:

1. Pick GRID TYPE.
2. Select the grid type.

Selections:
cubic
hexagonal
diamond
tetragonal
return

Figure 3.55:
Grid types menu

3.8.4 GRID SPACE

The GRID SPACE option is used to set the grid spacing. The default is 2.8 Å, twice the van der Waals radius of water.

User Input:

1. Pick GRID SPACE.
2. Enter the desired grid spacing in Å.

3.8.5 INNER CUTOFF

The INNER CUTOFF option determines the inner cutoff distance, that is, how close the solvation shell comes to the molecule being solvated. The default is 2.8 Å, twice the van der Waals radius of water.

User Input:

1. Pick INNER CUTOFF.
2. Enter the inner cutoff distance in Å.

3.8.6 OUTER CUTOFF

The OUTER CUTOFF option determines the outer cutoff distance, that is, how far the solvation shell extends beyond the molecule being solvated. The program is limited to adding 20,000 solvent molecules and will not accept a cutoff distance that generates more than this value. The default is 5.6 Å, four times the van der Waals radius of water.

User Input:

1. Pick OUTER CUTOFF.
2. Enter the outer cutoff distance in Å.

3.8.7 BOX SOLVENT

The BOX SOLVENT option is an on/off toggle. When on, the solvent shell created around the molecule takes the shape of a box. This can be useful when working with periodic systems.

User Input:

1. Toggle BOX SOLVENT on to have the solvent shell box-shaped.

3.8.8 AUTO GROUP

The AUTO GROUP toggle determines whether a group will automatically be created to hold the solvent molecules. If AUTO GROUP is on, then when EXECUTE is picked, a solvent group is created and the solvent molecules are displayed. The solvent group is given the same name as the file that was selected to hold the coordinate information for the solvent (during SETUP). If AUTO GROUP is off, no group is created and the solvent molecules are not visible. In this case, the user must create a group using the MAKE GROUP functions in order to see the solvent molecules on the display (see Section 4.1). The default is on.

User Input:

1. Toggle AUTO GROUP on or off.

3.8.9 HALF BOND

The HALF BOND toggle determines how the solvent molecules will be displayed (this only applies if AUTO GROUP is on). If HALF BOND is on when the solvent group is created, then the half-bond colors are used: each atom is assigned a CPK color, and the color of each heteronuclear bond changes color halfway. If HALF BOND is off, then the bonds will all be one color. That color may be changed using the CHANGE COLORS operation discussed in Section 4.5. The default is on.

User Input:

1. Toggle HALF BOND on to use the half-bond colors, or off to use a single color.

3.8.10 EXCLUDE SELF

If the EXCLUDE SELF toggle is on, then the groups to be solvated are automatically selected as groups to be used for contact exclusions (see SETUP, Section 3.8.1). This will create a solvent shell around the solvated groups. If the EXCLUDE SELF toggle is off and no additional groups are specified to be excluded, then a complete solvent grid (without any holes in it) will be created. One would generally want to exclude the solvated groups; therefore the default for EXCLUDE SELF is on.

If the user changes the EXCLUDE SELF setting, SETUP must be done again before EXECUTE is picked, since the positions of the atoms to be excluded are calculated during SETUP.

User Input:

1. Toggle EXCLUDE SELF on or off.

3.8.11 EXECUTE

EXECUTE generates the solvent shell, and if AUTO GROUP is on, creates a group for it and draws it on the display. The solvent group can be toggled on or off as desired using the global DISPLAY button. It is given the same name as the file that was selected to hold the coordinate information for the solvent (during SETUP). A message indicating the number of solvent molecules generated appears in the text display.

User Input:

1. Pick EXECUTE to generate the solvent shell.

3.9 ROTATE BOND

The ROTATE BOND operations are used to rotate a portion of a molecule about a single bond. This is done using the options on the ROTATE BOND menu shown in Figure 3.56. After picking ROTATE BOND, this menu is displayed. The SETUP function is automatically invoked and the user is prompted to define the bond vector, that is, to select the two atoms of the bond that is to be rotated. The portion of the molecule to be rotated is located past the second atom picked. The user then selects the groups to be graphically updated from the Groups menu.

The rotations are done using the dials or mouse. Up to eight bonds may be rotated at the same time. The number of bonds to be simultaneously rotated is set with the NBONDS option. The additional rotatable bonds always appear downstream and adjacent to the defined bond. NBRANCHES specifies the number of branch points the program will remember (five is the maximum). This is used in conjunction with the FORWARD operation described below.

After the number of rotatable bonds and branch points are set, EXECUTE is picked. This generates a ghost image for the region being rotated. Labels appear on the structure to indicate which dial affects which bond (if your system does not have a dials unit, see Table 3.2 for a list of the mouse manipulations which correspond to each of the eight dials). The dihedral angle values are also displayed. Turning the dial or using the mouse appropriately causes the fragment (shown in red) to rotate while keeping the rest of the molecule fixed. The dihedral angle values are also updated. The original location of the fragment is still indicated. CLEAR may be used to return the fragment to this position. When satisfied with the new orientation, the user toggles UPDATE. The coordinates of the rotated atoms are updated and the old structure disappears.

The user may move the defined bond vector throughout the structure using the FORWARD and REVERSE toggles. Picking FORWARD advances the defined bond vector and all additional rotatable bonds by one bond. If a branch point is encountered, the program will record it and return to it when there are no more bonds in the forward direction. REVERSE reverses the bond vector (and hence the associated ghost image) to the opposite direction. The user may choose a new bond vector at any time by picking SETUP and selecting two new atoms on the display.

The FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, and ZERO ORIGIN options allow the user to set new global rotation points.

User Input:

1. Pick ROTATE BOND.
2. Select two atoms on the display to define the bond vector.

Rotate bond:
setup
nbonds=2
nbranches=1
execute
update
clear
forward
reverse
file origin
group origin
pick origin
zero origin
return

Figure 3.56:
ROTATE BOND
menu

3. Pick the groups to be updated from the Groups menu. Terminate the selection by picking RETURN. This step is skipped if ROTATE BOND was invoked from a Builder (the Builder group is used) or if there is only one group in the system.
4. Set NBONDS and NBRANCHES to the desired values.
5. Pick EXECUTE to generate the ghost image and enable the dials or mouse.
6. Use the appropriate device to rotate the ghost fragment(s) (see Table 3.2). If desired, use CLEAR to return the fragment(s) to the last updated position.
7. Pick UPDATE to save the current orientation.
8. Use the FORWARD and REVERSE options to move the bond vector throughout the structure, or pick SETUP and move the bond vector by choosing two atoms on the display.

Bond Rotated ¹	Dials	Mouse	Normal Function
defined bond	dial 8	middle button + control key	depth cueing
bond #2	dial 7	middle button + shift key	z translate
bond #3	dial 6	middle button (up/down mvt) ²	y translate
bond #4	dial 5	middle button (left/right mvt) ²	x translate
bond #5	dial 4	right button + control key	zoom
bond #6	dial 3	right button + shift key	z rotate
bond #7	dial 2	right button (left/right mvt) ²	y rotate
bond #8	dial 1	right button (up/down mvt) ²	x rotate

¹ Bonds 2-8 are sequentially the next bond down the structure, continuing in the direction defined by the bond vector.

² It is difficult to restrict mouse motion to *only* up/down or left/right (the adjacent bond may also be rotated). To obtain precise control when using the mouse, keep NBONDS at 2 and redefine the bond vector when necessary.

Table 3.2: Using the dials or mouse to rotate bonds

3.9.1 SETUP

The SETUP option is used to define the bond vector, that is, to specify the two atoms of the bond that is to be rotated. The user selects two atoms from the display. The portion of the molecule to be rotated is located past the second atom picked. This is also the direction the bond vector moves when the FORWARD option is used (see FORWARD in Section 3.9.7). The user is then prompted to select the groups to be updated. This tells the program which groups need to be graphically updated when EXECUTE is picked. This step is skipped, however, if ROTATE BOND was invoked from one of the Builders, since the Builder group is automatically chosen.

When the user first chooses ROTATE BOND, SETUP is automatically invoked and the user is prompted to define the bond vector. However, SETUP may be picked again later to move the bond vector to another location or to change the groups to be updated.

User Input:

1. Pick SETUP.

2. Select two atoms on the display to define the bond vector.
3. Pick the groups to be updated from the Groups menu. Terminate the selection by picking RETURN. This step is skipped if ROTATE BOND was invoked from a Builder (the Builder group is used).

3.9.2 NBONDS

The NBONDS option allows the user to set the number of bonds to be simultaneously rotated. The default is two (one for GKS systems), but up to eight bonds may be rotated at the same time (one for each dial). The additional rotatable bonds always appear downstream and adjacent to the defined bond (rings are bypassed). In assigning dials to bonds, the program selects dial eight for the first bond, dial seven for the second bond, and so on. This pattern preserves the most important dials (dials one through four corresponding to x, y, and z rotation and scaling). The corresponding mouse manipulations are listed in Table 3.2. NBONDS must be set prior to picking EXECUTE.

User Input:

1. Pick NBONDS.
2. Select the number of bonds to be simultaneously rotated (one to eight).

3.9.3 NBRANCHES

The NBRANCHES option allows the user to set the number of branch points that the program will remember. If a branch point is encountered when advancing the bond vector with the FORWARD command, the program will record it and return to it when there are no more bonds in the forward direction. Up to five branch points may be remembered. The default is one. If changing NBRANCHES, it must be set prior to picking EXECUTE.

User Input:

1. Pick NBRANCHES.
2. Select the number of branch points to be remembered (zero to five).

3.9.4 EXECUTE

Picking EXECUTE generates a ghost image for the region being rotated and enables the dials or mouse. Labels appear on the structure to indicate which dial affects which bond, and the dihedral angle values are shown. The current values for NBONDS and NBRANCHES are used.

User Input:

1. Pick EXECUTE.

3.9.5 UPDATE

Picking UPDATE saves the coordinates of the rotated atoms in the current orientation. The old structure disappears.

User Input:

1. Pick UPDATE to save the current orientation.

3.9.6 CLEAR

CLEAR is used to return the fragments being rotated to their last-updated position.

User Input:

1. Pick CLEAR to return the fragments being rotated to their last-updated position.

3.9.7 FORWARD, REVERSE

The FORWARD and REVERSE toggles are used to move the defined bond vector throughout the structure. Picking FORWARD advances the bond vector and all additional rotatable bonds by one bond in the direction specified by the bond vector. The program looks at the connectivity record of the second atom picked; the first atom found when going in the same direction becomes the second atom for the next bond. This process is repeated until there are no more atoms in the forward direction. If a branch point is encountered, the program will record it and return to it when there are no more bonds in the forward direction. The number of branch points saved is set by the user with the NBRANCHES option. Non-rotatable bonds (*e.g.*, rings) are skipped. Thus, the ordering of the atoms in the connectivity record determines the progression of the bonds. For peptides, the default ordering causes the progression to proceed down a side chain before proceeding down the main chain.

REVERSE reverses the bond vector (and hence the associated ghost image) to the opposite direction. The combination REVERSE, FORWARD, and REVERSE effectively moves the bond back one bond. Picking REVERSE also causes the program to reset its branch point stack.

User Input:

1. Pick FORWARD to advance the defined bond vector by one bond or pick REVERSE to reverse the bond vector (and the associated ghost image) to the opposite direction.

3.9.8 FILE ORIGIN

The FILE ORIGIN operation sets the center of rotation to the geometric center of a selected input data file. When FILE ORIGIN is picked the input data files appear in the Files menu. The user selects the file to be used for centering. If only one file has been read into the program, then it is selected automatically as the file origin.

User Input:

1. Pick FILE ORIGIN.
2. Pick the file from the Files menu.

3.9.9 GROUP ORIGIN

The GROUP ORIGIN operation sets the center of rotation to the geometric center of a set of groups selected from the Groups menu.

User Input:

1. Pick GROUP ORIGIN.
2. Pick the groups to be used for centering from the Groups menu.
3. Pick RETURN to terminate the list.

3.9.10 PICK ORIGIN

The PICK ORIGIN operation sets the center of rotation to an atom picked from the display.

User Input:

1. Pick PICK ORIGIN.
2. Pick the atom desired for the rotation center.

3.9.11 ZERO ORIGIN

Pick ZERO ORIGIN to set the center of rotation to $x,y,z = 0,0,0$.

3.10 DOCK

The operations on the DOCK menu shown in Figure 3.57 are designed to move one set of atoms with respect to another, as in docking an inhibitor with an enzyme or adsorbing a molecule on the surface of a crystal. Options are available which allow the user to control which groups are rotated and translated, to reposition the origin of rotation, to modify atomic coordinates of structures, and to reset the transformations. An interactive docking feature is available which allows the user to monitor interatomic distances, force vectors, and total energy as one of the structures is moved. The user may then perform energy minimizations or dynamics calculations on the system in the new docked position.

Dock:
execute
modify coords
rotate all
set origin
return

The user toggles EXECUTE, and is then asked to select the groups to be docked from the Groups menu. The user toggles the groups he wishes to move. After he has selected the movable groups, the dials (or mouse) will move only those groups. The user is then asked to select the groups to be used for analysis; if no groups are selected, then the user may continue with the docking operations. If any groups are selected, however, then the program goes into interactive docking mode and the ENERGY DOCK menu shown in Figure 3.59 is displayed. The interactive docking functions allow the user to monitor interatomic distances and force vectors as one of the structures is moved. The user may then perform energy minimizations or dynamics calculations on the system in the new docked position.

Figure 3.57:
DOCK menu

Toggling RETURN terminates the docking process. If interactive docking was used, the atomic coordinates are automatically updated to the docked configuration. Otherwise, the MODIFY COORDS option must be chosen. If the user does not elect to update the coordinates, the groups are returned to their original positions.

Note: To save the transformed atomic coordinates permanently, a new coordinate file must be written using IN-OUT as described in Chapter 2. This applies to both interactive and non-interactive docking modes.

Rotations and Translations:

Images on the graphics display are rotated, translated, zoomed in or out, and depth cued by using the dials or mouse. These functions are described for the different graphics systems in Chapter 1, Section 1.2. The axes are in terms of graphics display coordinates where

x is horizontal (positive to right)
y is vertical (positive up)
z is out of display (positive toward the user)

When a structure is read in, the (x, y, z) spatial coordinates of the atoms are stored on disk. The program converts these spatial coordinates into graphics display coordinates with a 4-by-4 transformation matrix denoted as RT (for rotation-translation). Rotating the dials or using the middle or

right buttons on the mouse changes the transformation matrix and thereby translates, rotates, or zooms the object on the display. Every time EXECUTE or ROTATE ALL is picked, the transformation generated by the dials or mouse is used to update the RT transformation, and the transformation associated with the dials or mouse is set to the unit matrix. The RT transformation is then used to update the atomic coordinates stored on disk. This is done automatically when using the interactive docking feature, otherwise the MODIFY COORDS option must be chosen. A detailed explanation of the procedures for constructing, modifying and writing RT transformations is presented in Appendix F.

User Input:

1. Pick DOCK.
2. Pick EXECUTE.
3. Pick the groups containing atoms to be docked from the Groups menu. Pick RETURN to terminate the list.
4. Pick the additional groups containing atoms to be used in analysis (if any groups are selected, the program goes into interactive docking mode).
5. Interactive docking:
 - a) Use the options on the ENERGY DOCK menu (see Section 3.10.5). Rotate and translate the group(s) to be docked. Use ROTATE ALL as needed to rotate and translate the entire structure.
 - b) Pick RETURN to terminate the interactive docking process.
- Non-interactive docking:
 - a) Rotate and translate the group(s) to be docked. Use ROTATE ALL as needed to rotate and translate the entire structure.
6. Pick RETURN to terminate the docking process.
7. Pick YES or NO from the adjacent menu to perform MODIFY COORDS or not (non-interactive docking only).
8. To save the transformed coordinates permanently, use WRITE under IN-OUT, Section 2.2.

3.10.1 EXECUTE

The EXECUTE option is used to initiate the docking operations.

3.10.2 MODIFY COORDS

Picking the MODIFY COORDS operation updates the atomic coordinates stored on disk by applying the RT transformation generated while docking. Thus, MODIFY COORDS allows the user to save an intermediate result under DOCK. If further manipulation of the structure yields a better arrangement, then the atomic coordinates can be updated again. MODIFY COORDS is not needed when doing interactive docking, as the atomic coordinates are automatically updated.

Note: The transformed coordinates are only saved permanently when they are written to disk using the IN-OUT operations described in Chapter 2.

3.10.3 ROTATE ALL

Toggling ROTATE ALL reconnects all groups to the dials (or mouse), so that the user can alternatively move selected groups (ROTATE ALL off) or all groups (ROTATE ALL on).

3.10.4 SET ORIGIN

The options on the SET ORIGIN menu are used to reset the center of rotation (FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, and ZERO ORIGIN), to center a selected file on the display (CENTER), to reset the dials or mouse to zero (RESET DIALS), and to reset the rotation/translation transformation matrices (RESET ALL).

FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN

These options are used to reset the center of rotation. FILE ORIGIN sets it to the geometric center of a selected input data file, GROUP ORIGIN sets it to the geometric center of a set of groups selected from the Groups menu, and PICK ORIGIN sets it to an atom picked from the display. ZERO ORIGIN sets the center of rotation to $x, y, z = 0,0,0$. These options are described in Sections 3.9.8 - 3.9.11.

CENTER

Toggling CENTER centers the selected file in the middle of the display by modifying its coordinates so that the average value of x, y , and z each equal 0.

User Input:

1. Pick CENTER.
2. Pick the file to be centered. This step is skipped if there is only one file in the system.

RESET DIALS

Picking RESET DIALS resets the dials or mouse to zero (zero rotation, zero translation, and zero zoom). All groups currently connected to the dials or mouse will move back to the positions they occupied before being last connected, *i.e.*, the last time ROTATE ALL or EXECUTE was toggled. (**Warning:** The molecule may disappear from the display. If so, zoom out until the molecule appears in view; use the x and y -translate functions to position the molecule in the center of the display, then zoom in to fill the display.)

RESET ALL

Picking RESET ALL resets the dials or mouse to zero (see above) and resets the RT transformation to the unit matrix. All groups currently connected to the dials or mouse will move back to the po-

Set origin:
file origin
group origin
pick origin
zero origin
center
reset dials
reset all
return

Figure 3.58:
SET ORIGIN
menu

sitions they occupied before the last time MODIFY COORDS was toggled, *i.e.*, to the current atomic coordinates stored on the host computer. (**Warning:** The molecule may disappear from the display. If so, zoom out until the molecule appears in view; use the x and y-translate functions to position the molecule in the center of the display, then zoom in to fill the display.

3.10.5 ENERGY DOCK (Interactive Docking)

When doing interactive docking (groups were selected for analysis) the ENERGY DOCK menu shown in Figure 3.59 is displayed. The EXECUTE option is automatically highlighted. The MODIFY COORDS, ROTATE ALL, FILE ORIGIN, GROUP ORIGIN, and PICK ORIGIN options all function as described for non-interactive docking (Sections 3.10.2 - 3.10.4 above). Several additional options are available on the ENERGY DOCK menu which provide interactive docking functionality. These are described below.

The AUTO DISTANCE option allows the user to monitor the distances between the atoms on the structure being moved and the other structures on the display. Lines are drawn connecting all the atoms within a specified cutoff distance. The minimum distance between the two structures is also displayed in the lower left corner. The AUTO ENERGY option interactively displays the force vectors for each atom on the structure being docked. The total energy of the system is also displayed in the lower left corner. The DOCK DEFAULTS option is used to change the docking defaults. The user may turn the display of the force vectors on or off, change the scale which determines the length of the force vectors, or reset the cutoff distance.

The MINIMIZE and SIMPLE DYNAMIC options allow the user to do energy minimizations and simple molecular dynamics calculations on the docked structures. These options function in the same way as those accessed under MECHANICS or DYNAMICS from the SIMULATE menu except that the energy quantities are not listed on the Simulation Results menu. However, the total energy of the system is displayed in the lower left corner for each step of the energy calculations. The ENERGY VAR option is also provided here so that the energy defaults may be readily changed if desired. The ROTATE BOND option allows the user to rotate a substructure about a specified bond and interactively monitor the effect of the rotation on the force vectors and interatomic distances. Options are also provided to internally save and restore the current atomic coordinates.

EXECUTE

The EXECUTE option is automatically highlighted indicating that interactive docking is operative.

Energy dock:
execute
modify coords
rotate all
auto distance
auto energy
dock defaults
energy var
minimize
simple dynamic
nstep=50
rotate bond
update list
xyz temp save
xyz restore
file origin
group origin
pick origin
return

Figure 3.59:
ENERGY DOCK
menu (interactive
docking only)

MODIFY COORDS

This option functions as described for non-interactive docking (see Section 3.10.2 above).

ROTATE ALL

This option functions as described for non-interactive docking (see Section 3.10.3). However, when doing interactive docking, ROTATE ALL must be on in order to use the MINIMIZE and SIMPLE DYNAMIC options (all the displayed groups must be able to move).

AUTO DISTANCE

The AUTO DISTANCE option is an on/off toggle which allows the user to monitor the distances between the atoms on the structure being moved and the other structures on the display. Lines are drawn connecting all the atoms within a specified cutoff distance. The cutoff distance is specified using the docking defaults option (see DOCK DEFAULTS, below). The minimum distance between the structures is also displayed in the lower left corner.

User Input:

1. Toggle AUTO DISTANCE on to monitor the distances as the structure is moved.

AUTO ENERGY

The AUTO ENERGY option is an on/off toggle which allows the user to interactively display the force vectors for each atom on the structure being docked. The total energy of the system is also displayed in the lower left corner. The display of the force vectors may be turned off if desired or the relative length of the arrows may be changed with the DISPLAY FORCE and FORCE SCALE options on the DOCK DEFAULTS menu described below.

User Input:

1. Toggle AUTO ENERGY on to monitor the force vectors and total energy as the structure is moved.

DOCK DEFAULTS

Toggling DOCK DEFAULTS brings up the menu shown in Figure 3.60. The DISPLAY FORCE and FORCE SCALE options apply to the AUTO ENERGY option described above, and allow the user to turn the display of the force vectors on or off or to change the relative length of the vectors. DISTANCE CUTOFF specifies the cutoff distance used to determine whether connecting lines are drawn (see AUTO DISTANCE above).

A. DISPLAY FORCE

If DISPLAY FORCE is toggled on, the force vectors for each of the atoms on the structure being moved will be displayed (AUTO ENERGY must also be on). If it is off, the force vectors will not be shown. The default is on.

User Input:

1. Toggle DISPLAY FORCE on to display the force vectors; toggle it off if this is not wanted.

B. FORCE SCALE

The FORCE SCALE option sets the relative length of the arrows used to display the force vectors. The default value is 0.05 Å.

User Input:

1. Pick FORCE SCALE.
2. Enter the new value in Å.

C. DISTANCE CUTOFF

The DISTANCE CUTOFF option sets the cutoff distance for drawing connecting lines between the atoms on the structure being moved and the other structures on the display. When an atom on the structure being docked comes within the cutoff distance of an atom on any of the other structures, a line will be drawn between the two atoms. The AUTO DISTANCE option must also be on. The default cutoff distance is 2.0 Å.

User Input:

1. Pick DISTANCE CUTOFF.
2. Enter the new cutoff distance in Å.

Dock defaults:
display force
force scale
0.050
distance cutoff
2.000
return

Figure 3.60:
DOCK
DEFAULTS menu

ENERGY VAR

The ENERGY VAR option allows the user to modify the simulation parameters used in energy calculations, minimization and dynamics runs, to change the parameters used in the computation of nonbonding and hydrogen bonding interactions, and to set the temperature defaults used in dynamics calculations. It functions in the same way as the DEFAULTS option on the SIMULATE menu (see Section 5.12).

MINIMIZE

The MINIMIZE option allows the user to minimize the energy of a structure. Toggling MINIMIZE initiates NSTEP steps of energy minimization. This option functions as described under MECHANICS from the SIMULATE menu (see Section 5.7) except that the energy quantities are not listed on the Simulation Results menu. However, the total energy of the system is displayed in the lower left corner for each step of the energy calculations.

Note: ROTATE ALL must be on in order to perform the minimizations (all the displayed groups must be able to move). Also, the nonbond interactions list should be updated every time the user moves the structure being docked significantly (about 1 Å). This is required in order to maintain proper energy and force values (see UPDATE LIST, below).

SIMPLE DYNAMIC

The SIMPLE DYNAMIC option allows the user to perform simple dynamics calculations on the docked structures. Toggling SIMPLE DYNAMIC initiates NSTEP steps of molecular dynamics. The function is called SIMPLE DYNAMIC because it does dynamics without quenching, annealing, or an impulse direction. This is the same as the ADIABATIC dynamics option under DYNAMICS from the SIMULATE menu (see Section 5.8) except that the energy quantities are not listed on the Simulation Results menu. However, the total energy of the system is displayed in the lower left corner and the force vectors are shown for each step of the energy calculations provided AUTO ENERGY is on.

Note: ROTATE ALL must be on in order to perform the dynamics calculations (all the displayed groups must be able to move). Also, the nonbond interactions list should be updated every time the user moves the structure being docked significantly (about 1 Å). This is required in order to maintain proper energy and force values (see UPDATE LIST below).

NSTEP

The NSTEP option allows the user to change the number of steps performed in an energy minimization or simple molecular dynamics calculation.

User Input:

1. Pick NSTEP.
2. Enter the new value for NSTEP (default is 50).

ROTATE BOND

The ROTATE BOND option is used to rotate a substructure about a specified bond. It functions as described in Section 3.9. An additional function is also provided, however, in that the user can interactively monitor the effect of the rotation on the force vectors and interatomic distances (AUTO ENERGY and/or AUTO DISTANCE must be on).

UPDATE LIST

Toggling the UPDATE LIST option updates the nonbond list to reflect the values which correspond to the current position of the structure being docked. The nonbond list includes all the atoms which have nonbond interactions; however, only pairs of atoms within a specified cutoff distance are included when the list is initially defined (see Appendix H). The nonbond list must therefore be updated every time the structure being docked is moved significantly (about 1 Å) in order to maintain proper energy and force values.

User Input:

1. Toggle UPDATE LIST to update the nonbond list.

XYZ TEMP SAVE, XYZ RESTORE

The XYZ TEMP SAVE and XYZ RESTORE options provide a safeguard allowing the user to internally save the current coordinates for all the atoms in all the files and then restore them later. These options are described in Section 5.7.10.

FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN

These options are used to reset the center of rotation. FILE ORIGIN sets it to the geometric center of a selected input data file, GROUP ORIGIN sets it to the geometric center of a set of groups selected from the Groups menu, and PICK ORIGIN sets it to an atom picked from the display. These options are described in Sections 3.9.8 - 3.9.11.

3.11 MODIFY H

The MODIFY H operations shown in Figure 3.61 allow the user to add missing hydrogens to a structure (ADD ALL H, ADD HET H), to delete all hydrogens (DELETE ALL H), and to change the mass of hydrogens to that of deuterium or tritium to improve the stability of molecular dynamics simulations (H MASS=D, H MASS=T).

Hydrogens are generally not included in crystallographic structures because they scatter X-rays too weakly. The hydrogen addition operations allow the user to add hydrogens to input files. Two choices are available: hydrogens can be added to all atoms (ADD ALL H), or hydrogens can be added only to heteroatoms such as nitrogen, oxygen, or sulfur which can be involved in hydrogen bonds (ADD HET H). If hydrogens are added to potential hydrogen bond donors having rotational degrees of freedom (*e.g.*, hydroxyls and waters), then the hydrogens are placed in positions that form the best hydrogen bonds. The DELETE ALL H operation deletes all hydrogens.

Modify H:
add all H
add het H
delete all H
H mass=H
H mass=D
H mass=T
return

Figure 3.61:
MODIFY H menu

3.11.1 ADD ALL H, ADD HET H

The hydrogen addition operations allow the user to add hydrogens to input files. Two choices are available: hydrogens can be added to all atoms (ADD ALL H option), or hydrogens can be added only to heteroatoms such as nitrogen, oxygen, or sulfur which can be involved in hydrogen bonds (ADD HET H option). If hydrogens are added to potential hydrogen bond donors having rotational degrees of freedom (*e.g.*, hydroxyls and waters), then the hydrogens are placed in positions that form the best hydrogen bonds.

The program uses two hydrogen atom types: H___A and H_ (see Appendix I). It adds H___A hydrogens to heteroatoms, and adds H_ atoms to more neutral atoms like carbon and silicon. Nonbonded (van der Waals or electrostatic) interactions are not allowed for H___A. Instead an explicit potential term is used to describe the hydrogen bond interactions (see Appendix G). Nonbonded interactions are allowed for H_, *i.e.*, charges and/or van der Waals potentials can be assigned. In the latter case, the user may want to turn off the explicit hydrogen bond potential using ENERGY VAR under DEFAULTS (see Section 5.12).

User Input:

1. Pick ADD ALL H to add missing hydrogens to all atoms or pick ADD HET H to add missing hydrogens just to heteroatoms.
2. Pick the file to be operated on from the Files menu.
3. Pick the file(s) which contain additional potential acceptors for determining optimal positions for hydrogen bonding. Pick RETURN to complete the selection.

3.11.2 DELETE ALL H

The DELETE ALL H operation deletes all hydrogen atoms from a selected file.

User Input:

1. Pick DELETE ALL H to delete all hydrogen atoms.
2. Pick the file to be operated on from the Files menu.

3.11.3 H MASS=H, H MASS=D, H MASS=T

These options allow the user to improve the stability of molecular dynamics simulations by altering the mass of the hydrogen atoms to that of deuterium (2.0140 amu, H MASS=D) or tritium (3.0161 amu, H MASS=T). An H MASS=H option is also provided so that the hydrogens may be returned to their normal mass value (1.0080 amu). After selecting one of these options, any energy expression that is subsequently set up will contain the appropriate masses for the movable hydrogen atoms. Be sure to set up the energy expression before running any molecular dynamics calculations (see SETUP EEX, Section 5.2). After setting up the energy expression, all subsequently run molecular dynamics calculations will use the selected hydrogen mass value (naturally this only applies to molecular dynamics calculations; energy minimizations are not affected).

Warning: After selecting the desired hydrogen mass option, be sure to set up the energy expression before running any molecular dynamics calculations. Although this technique is useful in speeding up dynamics calculations, please note that the atomic coordinates and energy levels obtained when using the heavier hydrogen mass values will be different than those obtained when using the normal mass value for hydrogen atoms.

User Input:

1. Pick H MASS=H, H MASS=D, or H MASS=T to change the hydrogen atom mass to 1.0080, 2.0140 or 3.0161 amu respectively.

3.12 MOMENT

The MOMENT option allows the user to calculate moments for a structure. The center of mass and principal axes (and moments) may be calculated for all movable atoms or the user may specify that the calculations be done for all atoms in a set of groups or a set of files. The user selects the appropriate option desired from those listed in the adjacent menu (MOVABLE, GROUPS, or FILES) (see Figure 3.62), then selects the groups to be updated after transformation. He may then choose to transform the coordinates to center of mass and principal axes or not.

User Input:

1. Pick MOMENT.
2. If calculating moments for all movable atoms, pick MOVABLE, then skip to step 4. If the calculations are to be done for a set of groups or files, pick GROUPS or FILES respectively.
3. FILES only: pick the files to be used in calculating moments from the Files menu. Complete the selection by picking RETURN. Pick any additional files whose atoms are to be transformed. Complete the selection by picking RETURN. This step is skipped if there is only one file in the system.
4. Pick the groups to be updated after transformation. Complete the selection by picking RETURN.
5. If the coordinates are to be transformed to center of mass and principal axes, pick TRANSFORM; otherwise, pick NOT TRANSFORM.

Selections:
movable
groups
files
return

Figure 3.62:
Set types for
calculation menu

Selections:
transform
not transform
return

Figure 3.63:
Transform
coordinates menu

3.13 CONVERT

The operations on the CONVERT menu shown in Figure 3.64 allow the user to read a new conversion file and to convert coordinate files. The program currently allows up to 99 different atom types in the force field. Each atom read from a coordinate file is defined with a five-character atom label. BioDesign General Format (*bfg*) coordinate files include the force field data or atom type for each atom. However, the coordinate data from Brookhaven or other sources do not have this force field information. Consequently, the program uses a conversion file in which the atom labels used on the input coordinate file are converted to the atom types and charges used for the force field. The default conversion file is *dreidiixxx.cnv* located in the *bg_exe* directory (where *xxx* = Molecular Simulation software version number $\times 100$; e.g., 300 = version 3.0). For further discussion of *dreidiixxx.cnv* and its format see Appendix I.

Convert:
read conv file
convert
edit residue
return

Figure 3.64:
CONVERT menu

3.13.1 READ CONV FILE

The READ CONV FILE option allows the user to read in a new conversion table, which converts the input atom labels to the atom types and charges used in energy calculations.

User Input:

1. Pick READ CONV FILE.
2. Enter the name of the file containing the conversion table.

3.13.2 CONVERT

The CONVERT option allows the user to convert selected coordinate files based on the current conversion table. Some input files have conversion information written in them and, hence, will not be converted when read into the program. This operation allows the user to force a conversion on such files in addition to reconverting files based on a new conversion table that has been read into the program.

User Input:

1. Pick CONVERT.
2. Pick the file to be converted from the Files menu.

3.13.3 EDIT RESIDUE

The EDIT RESIDUE option allows the user to change the name of a residue and to reconvert the atoms of that residue using the new name. For example there are three residue names for histidine in *dreidiixxx.cnv* corresponding to: 1) the standard histidine side chain with the heteroatom hydrogen on the δ nitrogen (HIS), 2) histidine with the heteroatom hydrogen on the ϵ nitrogen (HSD), and 3) protonated histidine with hydrogens on both the δ and ϵ nitrogens (HSP). The EDIT RESIDUE option can be used to change from one residue name to another.

User Input:

1. Toggle EDIT RESIDUE.
2. Pick the residue to be modified from the display.
3. Enter the new residue name.

3.14 CONNECT

The CONNECT option is used to make bonds between atoms. When periodic boundary conditions are being used (*e.g.*, in the Crystal Builder) bonds may also be made across adjacent unit cells. The options on the CONNECT menu shown in Figure 3.65 allow the user to specify how these connections are made. The MIN DISTANCE and MAX DISTANCE options are used to set the upper and lower limits on the acceptable bond distance between two atoms. A further condition is that the distance must be less than SCALE x CRITERIA. The user may set the SCALE and choose CRITERIA to be the equilibrium bond distance (BOND RADII) or the van der Waals radii (VDW RADII). The user may also specify that connections be allowed for all atoms that meet the required conditions (CONNECT ALL) or limit connections to atoms that are not already connected (NOT CONNECTED). A DELETE BONDS option is also available which deletes all bonds. When using periodic boundary conditions an additional two-way toggle appears on the menu which allows connections to be made across cell boundaries (ALL CELLS) or restricts connections to atoms within the unit cell (UNIT CELL ONLY).

CONNECT does not draw in multiple bonds on the display. If the atom types for these bonds are correct, then the energy calculations are not affected. The EDIT function can be used to draw in the multiple bonds so that they appear correctly on the display (see Section 3.1.12). CONNECT can be used for extended or non-extended systems.

3.14.1 SETUP

SETUP must be done before any connections can be made. The SETUP function allows the user to pick the files which are to be connected.

User Input:

1. Pick SETUP.
2. Pick the files to be connected from the Files menu. Pick RETURN to complete the file selection.

3.14.2 CONNECT ALL/NOT CONNECTED

This option is a two-way toggle. When CONNECT ALL is displayed, all atoms that meet the required conditions may be connected. When NOT CONNECTED is displayed, connections may only be made between atoms that are not already connected (for example, atoms that have been drawn in but not bonded yet). CONNECT ALL is the default.

User Input:

1. Toggle until the desired option is displayed (CONNECT ALL or NOT CONNECTED).

Connect:
setup
connect all
all cells
min distance
0.800
max distance
2.000
criteria
bond radii
scale
1.200
execute
delete bonds
return

Figure 3.65:
CONNECT menu

3.14.3 ALL CELLS/UNIT CELL ONLY

This option is a two-way toggle and is available only when periodic boundary conditions are being used (*e.g.*, in the Crystal Builder). When toggled to ALL CELLS, connections may be made across cell boundaries. When UNIT CELL ONLY is displayed, connections may only be made between atoms that are within the unit cell.

User Input:

1. Toggle until the desired option is displayed (ALL CELLS or UNIT CELL ONLY).

3.14.4 MIN DISTANCE, MAX DISTANCE

The MIN DISTANCE and MAX DISTANCE options are used to set the upper and lower limits on the acceptable bond distance between two atoms. In order to form a bond, the distance between the two atoms must be greater than MIN DISTANCE and less than MAX DISTANCE. The atoms must also meet the conditions specified by the CRITERIA option described below. The default settings are 0.8 Å for MIN DISTANCE and 2.0 Å for MAX DISTANCE.

User Input:

1. Pick MIN DISTANCE or MAX DISTANCE.
2. Enter the minimum or maximum separation distance in Å.

3.14.5 CRITERIA

The CRITERIA option sets the criterion; this is used in conjunction with the SCALE option described below to determine whether a bond is formed between two atoms. Two criteria are available: BOND RADII and VDW RADII. If BOND RADII is chosen, then the distance between the two atoms must be less than the sum of their equilibrium bond distance multiplied by the SCALE factor in order for a bond to be formed. If VDW RADII is chosen, then the distance must be less than the scaled sum of the van der Waals radii. In addition to the CRITERIA specified here, the two atoms must also satisfy the conditions specified by MIN DISTANCE and MAX DISTANCE described above. The default CRITERIA is BOND RADII.

User Input:

1. Toggle the option until the desired criterion is displayed (BOND RADII or VDW RADII).

3.14.6 SCALE

The SCALE option sets the scale factor; this is used in conjunction with the CRITERIA option described above to determine whether a bond is formed. The distance between two atoms must be less than the product of the SCALE factor and the distance obtained using the current criterion. The two atoms must also satisfy the conditions specified by MIN DISTANCE and MAX DIS-

TANCE. The default SCALE value is 1.200 when BOND RADII is the CRITERIA and 0.55 when VDW RADII is the CRITERIA.

User Input:

1. Pick SCALE.
2. Enter the scale factor.

3.14.7 EXECUTE

EXECUTE uses the criteria and limits specified by the other options on the CONNECT menu to determine bond formation and then connects all the appropriate atoms.

User Input:

1. Pick EXECUTE to connect the atoms.

3.14.8 DELETE BONDS

The DELETE BONDS option deletes all bonds.

User Input:

1. Pick DELETE BONDS to delete all bonds.

Chapter 4

Visualize Operations

The options on the VISUALIZE menu shown in Figure 4.1 allow the user to change the appearance of the images shown on the graphics display. The MAKE GROUPS operations are used to divide a structure into functional parts which may then be independently displayed and manipulated. One may also create special hydrogen bond groups or create ribbons to show the general folding pattern of proteins, peptides and polymers. RENDER GROUPS is used to change the way selected groups are rendered. The rendering mode may be changed from the default (vectors) to spheres, cylinders, or ball and stick, and the structure may be displayed all in one color or in a half-bond representation. Options are also available which specify the width of the vectors, the size of the cylinders and spheres, and the smoothness of the surfaces. SURFACE allows the user to construct dotted surfaces which show the structure's contact surface, electrical charge or electrostatic potential surface. SET ORIGIN is used to change the center of rotation of the display. CHANGE COLORS may be used to change the color of selected groups and STEREO renders the images on the display as a stereo pair for 3-dimensional viewing. ROCK and AUTO ROTATE cause objects on the display to automatically rock or rotate about the y axis.

Visualize:
make groups
render groups
surface
set origin
change colors
stereo
rock
auto rotate
return

Figure 4.1:
VISUALIZE menu

4.1 MAKE GROUPS

Any data to be shown on the graphics display must be defined in terms of a group. When MAKE GROUPS is picked, the MAKE GROUPS menu shown in Figure 4.2 is displayed. The operations on this menu allow the user to define groups in terms of collections of atoms and bonds, which can then be selectively shown on the graphics display. A wide range of other data can also be displayed in terms of groups including:

- atom labels (see Section 3.1.13),
- atom types (see Section 3.1.12),
- atomic charges (see Section 3.1.13),
- potential energy surfaces (see Section 5.10.10),
- general vector lists (see Section 2.1.14),
- dotted surfaces (see Section 4.3), and
- trajectories (see Section 5.14).

The first seven options on the menu are used to define the portion of a structure which the user wishes to display. They are listed in the order in which they should be used (starting with SETUP and moving down sequentially to CREATE; see Figure 4.2). SETUP is used to specify the group which is to be created or redefined. FILES is used to select the coordinate file(s) from which the group is to be built. FULLFILE is an on/off switch which indicates whether the group is to include all the atoms in the file (on) or only part of the file (off). The CHAINS, RESIDUES and LIMITS operations allow the user to delimit the group in terms of standard structural elements or primitives for biological macromolecules, such as chains and residues. The standard residues, structural elements and structural descriptors are listed in Appendix A. One may also limit the group by specifying a range of residues to be included or by using a distance constraint. If the whole file is to be used (FULLFILE is on), then the CHAINS, RESIDUES and LIMITS options are not applicable and are skipped. Once the group has been defined, the user then toggles CREATE to generate it. The sequence of steps used to construct a group is given below.

SETUP ⇌ FILES* ⇌ CHAINS* ⇌ RESIDUES* ⇌ LIMITS ⇌ CREATE (FULLFILE off)
 SETUP ⇌ FILES* ⇌ CREATE (FULLFILE on)

*If the default for FILES, CHAINS or RESIDUES is appropriate, the corresponding step(s) may be skipped.

Note: *The correct sequence is essential.* Changes may be made before the group is created by redoing the appropriate step and repeating the later steps in proper order. All of the steps need not be repeated since FILES, CHAINS, RESIDUES, PRIMITIVES and SELECTORS always corre-

Make groups:
setup
full file
files
chains
residues
limits
create
rename file
rename group
Special:
H bonds
ribbons
return

Figure 4.2:
MAKE GROUPS
menu

spond to their last setting. (PRIMITIVES and SELECTORS appear on the LIMITS menu shown in Figure 4.4) However, the range information specified under LIMITS is canceled every time SETUP is picked. Section 4.1.6 explains how to use the LIMITS option.

The remaining options on the menu include RENAME FILE which is used to change file descriptors, and RENAME GROUP which changes group names. A special group containing hydrogen bonds can be constructed using the H BONDS option. A RIBBONS option is also available which is useful in showing the general folding pattern for proteins, peptides, and polymers.

4.1.1 SETUP

To define a new group pick SETUP. The user is prompted to select the group to be defined from the Groups menu. One of the groups currently listed may be redefined or the user may create a new group by choosing ADD GROUP. A dialog box appears and the user is asked to enter a name for the group. He then picks OK to complete the entry (CANCEL aborts it).

User Input:

1. Pick SETUP.
2. Pick the group to be defined from the Groups menu. To create a new group choose ADD GROUP.
3. Enter a name for the group (up to 14 characters) in the dialog box.

4.1.2 FILES

The FILES operation allows the user to select which coordinate files will be considered in creating a group. Coordinate files may be read from disk or created in BUILD (Chapter 3). The program can handle up to 21 separate coordinate files. When FILES is picked, the Files menu appears listing the files in sequential order. Any combination of these files may be picked to construct a group.

User Input:

1. Pick FILES.
2. Toggle on the desired files from the Files menu.
3. Pick RETURN to complete the operation.

4.1.3 FULL FILE

To create a group containing all the atoms of any one coordinate file, toggle the FULL FILE switch on. If the FULL FILE switch is on, the group created will contain all the atoms in the file. The user must set up the group and select the file whose atoms will be used, but the CHAINS, RESIDUES, and LIMITS toggles are not active. In order to create a group containing only part of a file, the FULL FILE switch must be toggled off.

User Input:

1. Toggle FULL FILE on to create a group with all the atoms in a coordinate file, or toggle FULL FILE off and use the CHAINS, RESIDUES, and/or LIMITS options to create a group containing only part of a file.

4.1.4 CHAINS

Biological macromolecules often contain separate chains of peptides or nucleotides, which are denoted by a one-letter (upper case) or blank chain designator. When CHAINS is picked, the CHAINS menu appears listing the chain designators which are currently in the system. (Figure 4.3 shows the CHAINS menu for DNA). One can use all chains for each FILE by toggling on "?" or one can use a subset by toggling on the chain designators desired. (The underscore indicates a blank chain designation.)

User Input:

1. Pick CHAINS.
2. Pick the desired chains from the CHAINS menu (default is "?").
3. Pick RETURN to complete the selection.

Chains:
?
—
A
B
return

Figure 4.3 :
Chains menu
(example)

4.1.5 RESIDUES

Any three-letter name can be used as a residue name (a list of the standard residue names is given in Appendix A.1). When RESIDUES is picked, the RESIDUES menu appears listing the residues which are currently in the system. The group can be restricted to an arbitrary subset of these residues or ALL of them may be included.

User Input:

1. Pick RESIDUES.
2. Pick the desired residues from the RESIDUES menu (default is ALL).
3. Pick RETURN to complete the selection.

4.1.6 LIMITS

Groups may be restricted further by toggling the LIMITS operation which brings up the menu shown in Figure 4.4. Three criteria are used in LIMITS: PRIMITIVES, SELECTORS, and RANGE or DISTANCE. PRIMITIVES refers to the standard structural elements of a biological macromolecule, *e.g.*, main chain, side chain and carbonyl. (See Appendix A.2 for a complete listing of primitives.) SELECTORS refers to characteristics that the primitives possess, *e.g.*, a side chain can be hydrophobic or part of an α -helix. (See Appendix A.3 for a complete listing of selectors.) RANGE refers to a range of residue sequence numbers (*e.g.*, residue 25 through residue 134). For any particular set of selectors and primitives, one can restrict the range of residues or bases to be considered. The range can be specified by entering the residue numbers from the keyboard, by picking residues on the graphical display, or by using a distance constraint. Up to 20 combinations of ranges, primitives, and selectors can be used for each group. *Note:* The default is 0 to 999 (*i.e.*, all residues). The default is reset each time SETUP is toggled.

Grouping by residue sequence number is also very useful for nonbiological systems created under BUILD. Structures created with BUILD are assigned residue numbers entered under the SETUP operation (see Section 3.1.1).

Limits:
primitives
selectors
key res range
key one res
pick start res
0
pick stop res
0
pick one res
0
keep range
forget range
distance
return

Figure 4.4:
LIMITS menu

PRIMITIVES

Upon picking PRIMITIVES from the LIMITS menu, a list of the standard structural elements as shown in Figure 4.5 will appear. Any number of these may be toggled on. STD SET (standard set) includes all the elements listed except TRACE and LABELS CA and is used as the default. See Appendix A.2 for more information.

User Input:

1. Pick PRIMITIVES on the LIMITS menu.
2. Toggle on the primitive(s) desired (default is STD SET).
3. Pick CLEAR to clear selections and to reset the default.
4. Pick RETURN to complete the selection.

Primitives:
STD SET
TRACE
MAIN CHAIN
HMAIN
SIDECHAIN
HSIDE
CARBONYL
PHOSPHATE
HPHOSPHATE
SUGAR
HSUGAR
BASE
HBASE
METAL
WATER
HWATER
NONSTD
HNONSTD
LABELS CA
CLEAR
RETURN

Figure 4.5:
Primitives menu

SELECTORS

Upon picking SELECTORS from the LIMITS menu, a list of the standard descriptors will appear as shown in Figure 4.6. (See Appendix A.3 for more information.) Any number of these may be toggled on. *Note:* The structural characteristics such as HELIX, SHEET, TURN and ACTIVE SITE can be used to construct a group only if the information is contained in the Brookhaven file.

User Input:

1. Pick SELECTORS on the LIMITS menu.
2. Toggle on the desired selectors (default is ALL).
3. Pick CLEAR to clear selections and to reset the default.
4. Pick RETURN to complete the selection.

KEY RES RANGE

Specify range by entering the starting and stopping sequence numbers.

Warning: Nonstandard groups such as water and metal are given numbers in Brookhaven files so that the largest number may be greater than the number of residues or bases.

User Input:

1. Pick KEY RES RANGE on the LIMITS menu.
2. Enter the starting residue number in the dialog box (must be greater than -1).
3. Enter the stopping residue number in the dialog box (must be less than 1000).

KEY ONE RES

Specify a single residue by entering the sequence number.

User Input:

1. Pick KEY ONE RES on the LIMITS menu.
2. Enter the residue number in the dialog box (must be greater than -1 and less than 1000).

PICK START RES, PICK STOP RES

Specify a range by picking the starting and stopping residues off the display. Any atom of the residue (or base) may be picked. The identities of the picked residues are indicated in the menu boxes directly beneath PICK START RES and PICK STOP RES (see Figure 4.4). Both the residue name and number are shown (zero is used as the default).

Selectors:
ALL
HELIX
SHEET
TURN
ACTIVE SITE
POSITIVE
NEGATIVE
AROMATIC
FOBIC
FILIC
AMBIVALENT
CLEAR
RETURN

Figure 4.6:
Selectors menu

User Input:

1. Pick either PICK START RES or PICK STOP RES on the LIMITS menu.
2. Pick the residues from the display.
3. If an incorrect residue was picked, reactivate by toggling either PICK START RES or PICK STOP RES.
4. Pick KEEP RANGE after the correct residues are picked for both PICK START RES and PICK STOP RES.

PICK ONE RES

Specify a single residue by picking off the display. Any atom of the residue (or base) may be picked. The identity of the picked residue is indicated in the menu boxes directly beneath PICK START RES and PICK STOP RES (see Figure 4.4). Both the residue name and number are shown (zero is used as the default). If the wrong residue is picked, try again.

User Input:

1. Pick PICK ONE RES on the LIMITS menu.
2. Pick the residue from the display.
3. Reactivate by toggling PICK ONE RES if an incorrect residue was picked.
4. Pick KEEP RANGE if the correct residue was picked.

KEEP RANGE

Each time that KEY RES RANGE, KEY ONE RES, KEEP RANGE, or DISTANCE is toggled constitutes one range. A group can have up to 20 ranges for any particular choice of FILES, CHAINS, and RESIDUES. Each range can have different combinations of primitives and selectors. All of the ranges can be reset by toggling the FORGET RANGE option on the LIMITS menu.

FORGET RANGE

Toggling FORGET RANGE deletes all of the range information.

DISTANCE

A range of residues may be defined in terms of a distance constraint, *i.e.*, defined as all primitives within a specified distance of a selected group or set of groups. As an example, the user can define an active site as all residues within 10 Å of an inhibitor. This does require, however, that the base group or groups (*i.e.*, the group(s) containing the atoms from which the distances are measured) already exist. In the above example the base group is the group that contains the inhibitor molecule. A group built in this fashion will contain primitives that have at least one atom within the cutoff distance of at least one atom of the base group(s). (The primitives must also satisfy the criteria specified by FILES, CHAINS, and RESIDUES). Primitives comprising the base group(s)

will not be included in the newly created group. (In the above example, the inhibitor molecule will not be part of the new group.)

User Input:

1. Pick DISTANCE on the LIMITS menu.
2. Pick the group(s) containing base atoms from the Groups menu. Complete the selection by picking RETURN.
3. Enter the cutoff distance in Å in the dialog box.

4.1.7 CREATE

Toggling CREATE generates a group using the criteria defined in the SETUP, FILES, CHAINS, RESIDUES and LIMITS operations. While the group is being created, the message *Working...* appears in the Modules/Messages window. When the group has been created, the logo reappears and the new group is displayed in the graphics window. The group name also appears in the Groups menu. As groups are created, each is assigned a different color (unless HALF BOND is on). To change the color, use the CHANGE COLORS operation described in Section 4.5.

4.1.8 RENAME FILE

The RENAME FILE option is used to change the descriptor used to name files. After picking RENAME FILE, the user selects the file whose descriptor is to be changed from the Files menu and then enters the new name in the dialog box.

User Input:

1. Pick RENAME FILE.
2. Select the file whose descriptor is to be changed from the Files menu. Complete the selection by picking RETURN.
3. Enter the new file descriptor in the dialog box.

4.1.9 RENAME GROUP

The RENAME GROUP option is used to rename selected groups. After picking RENAME GROUP, the user selects the group whose name is to be changed from the Groups menu and then enters the new name in the dialog box.

User Input:

1. Pick RENAME GROUP.
2. Select the group whose name is to be changed from the Groups menu. Complete the selection by picking RETURN.
3. Enter the new group name in the dialog box.

4.1.10 H BONDS

A special group containing hydrogen bonds can be constructed using the H BONDS option. The group of hydrogen bonds is defined in terms of a primary and secondary set of groups. This special group contains all of the hydrogen bonds between atoms contained in a primary set of groups plus all of the hydrogen bonds connecting atoms in a primary set of groups with those in a secondary set. Before the special group can be created, all of the pertinent hydrogens must be explicitly included in the primary and secondary groups (see ADD ALL H, ADD HET H in Section 3.11.1).

User Input:

1. Pick HBONDS.
2. Pick the primary groups from the Groups menu. Complete the selection by picking RETURN.
3. Pick the secondary groups from the Groups menu. Complete the selection by picking RETURN.
4. Pick the group to be defined from the Groups menu. Complete the selection by picking RETURN.
5. Enter the name of the group to be created in the dialog box.
6. Enter the distance cutoff for hydrogen bonds (*recommended*: 3.5 Å to 4.0 Å).
7. Enter the angle cutoff for hydrogen bonds (*recommended*: 30° to 50°).

4.1.11 RIBBONS

The RIBBONS option allows the user to display the general folding pattern of proteins, peptides, and polymers. This option is only available on the IRIS, the Titan, and the RS6000. The algorithm used to create the ribbons is that described by M. Carson and C. E. Bugg.¹ After toggling RIBBONS, the user selects the file which is to be used from the Files menu. He then picks a group which will be used to display the ribbon from the Groups menu and enters a name for it. He specifies the total width for the ribbon (in Å) and the number of threads desired (up to 9). If a ribbon width of 0 is specified, the number of threads is automatically set to one. The ribbon then appears on the display. The color of the ribbon can be changed using the CHANGE COLORS operations (see Section 4.5), and the width of the lines used to draw each thread can be set using the VECTOR WIDTH option (see Section 4.2.3).

User Input:

1. Pick RIBBONS.
2. Pick the file to be used from the Files menu. This step is skipped if there is only one file in the system.
3. Pick the group to be used to display the ribbon from the Groups menu.
4. Enter a name for the group in the dialog box.
5. Enter the width for the ribbon (in Å). A ribbon width of 0 sets the number of threads to one.
6. Enter the number of threads for the ribbon (up to 9).

¹ M. Carson, and C. E. Bugg, *J. Mol. Graphics*, 4, 121 (1986).

7. If desired, use the **CHANGE COLORS** operations to change the ribbon color (see Section 4.5) or use **VECTOR WIDTH** to change the line width used to draw the threads (see Section 4.2.3).

4.2 RENDER GROUPS

The options on the RENDER GROUPS menu shown in Figure 4.7 control the way groups are rendered or presented on the display. The first four options determine which rendering mode is used: VECTORS (lines), CYLINDERS, SPHERES (CPK), or BALL & STICK. BALL & STICK is only available on the IRIS and the RS6000. HALF BOND determines whether the group is displayed with the half-bond colors or in a single color. For VECTORS rendering, the width of the vectors used to draw the structure may be changed using VECTOR WIDTH. Similarly, for the other modes, the size of the spheres and cylinders used to represent the atoms and bonds may be set using the BALL SCALE and CYLINDER SCALE options respectively. The DIVISIONS setting determines the number of triangles used to create the surfaces of the objects and thus controls their smoothness. Shading and specular reflection may also be controlled on the Titan, ESV, and DEC 5000 using the SHADING TYPE and SPECULAR INTEN options. These options are described in more detail below. The user sets the rendering mode and defaults to the desired values, then picks EXECUTE to implement the rendering.

The default rendering mode is VECTORS, using a vector width of 1 pixel (2 pixels on the IRIS).

User Input:

1. Pick RENDER GROUPS.
2. Pick the group to be rendered from the Groups menu. Pick RETURN to complete the selection.
3. Set the rendering mode and other defaults as desired.
4. Pick EXECUTE to implement the rendering.

4.2.1 VECTORS/ CYLINDERS/ SPHERES/ BALL & STICK

These four options determine which rendering mode will be used to render groups. VECTORS (the default) renders atoms and bonds with straight lines. The VECTOR WIDTH option may be used to change the width of the vectors (see Section 4.2.3). CYLINDERS mode renders atoms and bonds as hollow cylinders, SPHERES renders atoms and bonds as interconnected spheres (CPK modeling), and BALL & STICK renders atoms as spheres and bonds as cylinders. BALL & STICK is only available on the IRIS and the RS6000. The size of the spheres and cylinders used to represent the atoms and bonds may be set using the BALL SCALE and CYLINDER SCALE options respectively. Similarly, the type of shading used, the smoothness of the surfaces, and the degree of specularity may also be set using the SHADING TYPE, DIVISIONS and SPECULAR INTEN options. The SHADING TYPE and SPECULAR INTEN options are not available on the IRIS and the RS6000.

Render groups:
vectors
cylinders
spheres
ball & stick
half bond
vector width
2
cylinder scale
0.100
ball scale
0.600
shading type
vertex
divisions
0.300
specular inten
0.75
execute
return

Figure 4.7:
RENDER
GROUPS menu

User Input:

1. Pick the rendering mode to be used (VECTORS, CYLINDERS, SPHERES, or BALL & STICK). BALL & STICK is only available on the IRIS and RS6000.

4.2.2 HALF BOND

Groups may be displayed with bonds in all one color or in a half-bond representation. If HALF BOND is on when a group is created, each atom will be assigned a CPK color (note that carbons are aqua not black), and the color of each heteronuclear bond will change color halfway. If HALF BOND is off, then the bonds in the group will all be one color. That color can be changed using the CHANGE COLORS operation discussed in Section 4.5.

User Input:

1. Toggle on HALF BOND to achieve the desired representation.

4.2.3 VECTOR WIDTH

The VECTOR WIDTH option determines the width of the vectors used to draw the lines when rendering in VECTORS mode. The width is given in pixels. The default is 2 pixels for the IRIS and 1 pixel on all other machines.

User Input:

1. Pick VECTOR WIDTH.
2. Enter the vector width desired (in pixels).

4.2.4 CYLINDER SCALE, BALL SCALE

CYLINDER SCALE and BALL SCALE are scaling factors which determine the radius used when rendering the cylinders and spheres respectively. BALL SCALE is multiplied by the van der Waals radius of each of the atoms. CYLINDER SCALE, however, is independent of atom size. CYLINDER SCALE is applicable when rendering in CYLINDERS mode, and BALL SCALE is applicable when rendering in SPHERES mode. When rendering in BALL & STICK mode, both scaling factors apply. The default values (0.100 for CYLINDER SCALE and 0.600 for BALL SCALE) were chosen so that when using BALL & STICK mode, the bonds are always much smaller than the atoms.

User Input:

1. Pick CYLINDER SCALE or BALL SCALE.
2. Enter the desired value in the dialog box.

4.2.5 SHADING TYPE

The SHADING TYPE option is available on the Titan, ESV and DEC 5000 only. It is used to change the shading method used when rendering objects. On the ESV and DEC 5000, four different methods are available: NORMAL (the default), COLOUR, DOT PROD, and NONE (no shading method). On the Titan, three shading methods are provided: VERTEX (the default), SURFACE, and CONSTANT. The selected method is displayed beneath SHADING TYPE on the menu.

User Input:

1. Toggle SHADING TYPE until the desired shading method is displayed.

4.2.6 DIVISIONS

The DIVISIONS option determines the number of triangles used to create the surfaces of the rendered objects, and thus controls their apparent smoothness. The values range from 0 to 1.000 with higher numbers creating a smoother (but slower) result. The default value is 0.300.

User Input:

1. Pick DIVISIONS.
2. Enter the desired value (0 to 1.000) in the dialog box.

4.2.7 SPECULAR INTEN

The SPECULAR INTEN option is only available on the Titan, DEC 5000, and ESV. It controls the intensity of the specular reflection (the specular highlights or bright spots that are created from the reflection of a point source of light). The default value is 0.75.

User Input:

1. Pick SPECULAR INTEN.
2. Enter the desired value (0 to 1.0).

4.3 SURFACE

The operations on the SURFACE menu shown in Figure 4.8 allow the user to construct both van der Waals and solvent accessible molecular surfaces using the approach developed by Michael Connolly²⁻³ and based on the definitions for molecular surfaces proffered by Frederic Richards⁴. The generated surface can be color-coded by group or to show atom types, electrostatic potential, atom charges or the hydrogen bond donors and acceptors.

The SURFACE menu is divided into surface functions used for calculating surfaces and reading and writing surface files, and display functions which are used to define and display the surface. Surfaces can be calculated, and surface files written or read, without displaying the surface on the screen.

Note that a surface can be calculated and displayed without creating a permanent file on disk. A surface can be generated for the entire structure on display or just for selected groups in the structure. The user determines which groups in the structure are used to generate the surface when the surface is calculated. He can also decide which part of the surface to save, by groups, when a dot file is written by using the WRITE PART function.

The READ and READ PART functions read in existing surface files. The READ PART function makes it possible to select groups on the Groups menu and input the parts of an existing surface file that correspond to the atoms in these groups. The READ with DOT file option under IN-OUT (Section 2.1.13) can also be used to read dotted surface files. Note that the displayed property can be changed without recomputing the positions of the surface dots.

The format of the surface file generated by the program is given in Section 4.3.11. The three letter extension *dot* is recommended for dotted surface files.

Surface:
calculate
r ead
read part
Display:
setup
groups
prop defaults
property
execute
write
write part
return

Figure 4.8:
SURFACE menu

² M. L. Connolly, *Science*, **221** (1983) 709-713.

³ M. L. Connolly, *J. Appl. Crystal.*, **16** (1983) 548-558.

⁴ F. M. Richards, *Ann Rev. Biophys. Bioeng.*, (1977) 151-176.

4.3.1 CALCULATE

The CALCULATE operation allows the user to calculate the dot positions for dotted molecular surfaces. The surfaces are defined in terms of a probe sphere rolling along the van der Waals spheres of the outer (accessible) atoms of the molecule or crystal. If a probe radius of zero is used, then a van der Waals surface is generated. However, if the probe radius is set to 1.4 Å, the van der Waals radius of water, then a solvent-accessible surface is generated.

After toggling CALCULATE, the menu shown in Figure 4.9 appears and the user is prompted to select the groups to be surfaced. The surface will be calculated for all groups that are selected. The program also prompts the user to select groups to be excluded. This feature allows the elimination of unwanted surface dots. For example, calculating a surface for a portion of a structure can generate dots which are actually internal to the structure, since not all atoms were included in the surface calculation. If one excludes groups adjacent to the groups to be surfaced, then only the true external part of the surface will be computed. The dot coordinates for the last calculated surface are stored in memory. At this point, properties may be computed and results displayed for any individual group, or subset of the groups (see Sections 4.3.4 - 4.3.8).

Calculate:
density
8.000
probe radius
0.000
vdw scale
0.890
execute
return

Figure 4.9:
CALCULATE
menu

Note: Calculation of a dotted molecular surface for hundreds of atoms can require a substantial amount of computer time.

User Input:

1. Pick CALCULATE. Skip to step 4 if there is only one group in the system.
2. Pick the groups to be surfaced from the Groups menu. Pick RETURN to complete the selection.
3. Pick the groups to be excluded from the Groups menu (see above). Pick RETURN to complete the selection.
4. Define any desired changes to the default dot density, probe radius, and van der Waals scale factor (see DENSITY, PROBE RADIUS, and VDW SCALE below).
5. Pick EXECUTE.

DENSITY

The DENSITY option defines the surface density of dots, in dots per square Å.

User Input:

1. Pick DENSITY.
2. Enter the desired surface dot density, in dots per square Å.

PROBE RADIUS

This toggle allows the radius of the surface probe to be set. A value of zero, produces van der Waals surfaces. A value of 1.4 Å will produce a water solvent-accessible surface.

User Input:

1. Pick PROBE RADIUS.
2. Enter the probe radius value.

VDW SCALE

The van der Waals radii of atoms to be surfaced are scaled by this factor for the surfacing operation. A value of 0.89 yields the standard CPK radii (providing the probe radius is zero).

User Input:

1. Pick VDW SCALE.
2. Enter the van der Waals scale factor value.

EXECUTE

The EXECUTE toggle completes the surface calculation.

User Input:

1. Pick EXECUTE to compute the dot surface.
2. Use the DISPLAY functions to view the calculated surface.

4.3.2 READ

The READ option allows the user to read dotted surface or *dot* files from disk. After picking READ, the user is prompted for a file name. When the READ operation has completed, the surface data are in memory, including any property from the file data. The surface may be displayed with the options in the display section of the menu (see Sections 4.3.4 - 4.3.8).

User Input:

1. Pick READ.
2. Enter the name of the *dot* file to be read in (see Section 1.6 for details on entering file names).

4.3.3 READ PART

READ PART is like READ above, except that it allows the reduction of the surface data input to just those atoms contained within certain groups.

User Input:

1. Pick READ PART.
2. Enter the name of the *dot* file to be read in (see Section 1.6 for details on entering file names).
3. Pick the groups which include the atoms to be selected from the surface data file. Pick RETURN to complete the selection.

Surface Display Functions

There are five display options on the SURFACE menu which control the computation and display of surface properties: SETUP, GROUPS, PROP DEFAULTS, PROPERTY, and EXECUTE. The CALCULATE, READ or READ PART function must be executed before the display functions are operable. Typically the display options are picked sequentially starting with SETUP and working through to EXECUTE. If the whole calculated surface is to be displayed, then the GROUPS toggle may be omitted. Likewise, if the default property parameters are desired, the PROP DEFAULTS option may be skipped. Solid color dot surfaces can be displayed by picking SETUP and EXECUTE only. Note that the CHANGE COLORS operations can only be performed on solid color dot surfaces.

4.3.4. SETUP

SETUP allows one to choose the group to contain the surface display, and to name it. The new group is not displayable until the property computation has been executed.

User Input:

1. Pick SETUP.
2. Pick the group to be assigned to the surface data currently in memory from the Groups menu. Choose ADD GROUP if creating a new group. Pick RETURN to complete the selection.
3. Enter the name of the new surface group.

4.3.5 GROUPS

The GROUPS option selects a subset of the atoms in the last calculated surface for inclusion in the new surface group. For example, if the surface is to be broken into several groups, one first computes the surface over all the groups, and then uses the GROUPS function with the other display functions to select the individual parts for creation as separate display groups.

User Input:

1. Pick GROUPS.
2. Pick the group(s) containing the atoms to be included in the new surface display group from the Groups menu. Pick RETURN to complete the selection.

4.3.6 PROP DEFAULTS

Picking PROP DEFAULTS brings up the property defaults menu shown in Figure 4.10. The operations on this menu allow the user to change the parameters used in computing dot surface properties. These parameters apply only to the electrostatic potential and charge properties.

MAX ELEC VALUE

The maximum electrostatic potential parameter (MAX ELEC VALUE) defines the mapping of electrostatic potential into colors for the electrostatic potential property (see ELECTROSTATIC, Section 4.3.7). A calculated potential of MAX ELEC VALUE or greater is displayed as blue, and potentials less than or equal to $-(\text{MAX ELEC VALUE})$ are displayed as red. Potentials between these values are interpolated as hues between red (120) and blue (360) according to Table 4.1 (see Section 4.3.11).

User Input:

1. Pick MAX ELEC VALUE.
2. Enter the value for MAX ELEC VALUE in the dialog box.

MAX CHARGE

The maximum charge parameter (MAX CHARGE) defines the mapping of gross atomic charges into colors for the CHARGE property (see Section 4.3.7). Atomic charges between $-(\text{MAX CHARGE})$ and $+\text{MAX CHARGE}$ are interpolated between red and blue as in MAX ELEC VALUE above. Charges outside this range are displayed as red (negative) or blue (positive).

User Input:

1. Pick MAX CHARGE.
2. Enter the value for MAX CHARGE.

GROSS CHARGE

The gross charge parameter (GROSS CHARGE) defines the coloration of the surface by gross atomic charge for the "+,0,-" property (see Section 4.3.7). Surface dots for atoms with a charge less than or equal to $-(\text{GROSS CHARGE})$ are red, those with a charge greater than or equal to $+\text{GROSS CHARGE}$ are blue, and all other dots are aqua.

User Input:

1. Pick GROSS CHARGE.
2. Enter the value for GROSS CHARGE.

Prop defaults:
max elec value
100.000
max charge
0.500
gross charge
0.200
total colors
20
return

Figure 4.10:
PROP DEFAULTS
menu

TOTAL COLORS

The total colors (TOTAL COLORS) toggle is used to set the total number of colors which are used in the electrostatic potential and atomic charge properties (see Section 4.3.7). The minimum is 3 and the maximum is 80. Increasing the number of colors creates smoother transitions between positive and negative regions, but requires more time to compute and display. Decreasing the number of colors emphasizes the demarcation between positive and negative regions.

User Input:

1. Pick TOTAL COLORS.
2. Enter the total number of colors to be used (3 - 80).

4.3.7 PROPERTY

The dot surfaces generated with the CALCULATE operation may be color-coded in order to illustrate certain features. The color-coding may be done by group or according to one of five properties. The PROPERTY option allows the user to select the property to be illustrated from the Properties menu which appears as shown in Figure 4.11. Five properties are available: ATOM TYPE, electrostatic potential (ELECTROSTATIC), atomic charge (CHARGE), gross atomic charge (-,0,+), and hydrogen bond donor/acceptor (H BOND DON/ACC). A CURRENT selection is also available which uses the property value currently stored in memory. This option would be used when you want to create a new surface group with the same property values as the most recently generated group or when a dot surface file containing the property colors to be displayed has been read from disk. These five properties are described in more detail below.

Properties:
current
atom type
electrostatic
charge
-, 0, +
h bond don/acc
return

The property defaults (Section 4.3.6) are applied if the electrostatic potential or charge properties are selected.

Figure 4.11:
Properties menu

User Input:

1. Pick PROPERTY.
2. Select the property to be illustrated from the PROPERTY menu. Pick RETURN to complete the selection.

CURRENT

The CURRENT property allows the EXECUTE function to be performed using the property values currently stored in memory. It does not change the current property values. This option is used when the user wishes to create a new surface group with the same property values as the most recently generated group, or when a dot surface file containing property colors has been read from disk.

ATOM TYPE

The ATOM TYPE property colors the dot surface according to the half-bond color of the nearest atom (CPK colors are used except for carbon which is aqua).

ELECTROSTATIC

The electrostatic potential (ELECTROSTATIC) property colors the surface according to the electrostatic potential at each point on the surface, using hues from red to blue. The potential at any point is given by the equation

$$ESP_i = S \times \sum_{j=1}^{N_{atom}} q_j / r_{ij} \quad (4.1)$$

where S is a scaling factor, q_j is the charge on each atom, and R_{ij} is the distance from each atom to the atom that is being surfaced.

CHARGE

The CHARGE property colors the surface according to the atomic charge of the nearest atom, using hues from red to blue.

-,0,+

The gross atomic charge property (+,0,-) colors highly charged sections of the surface. Dots near atoms having a charge greater than the cutoff are blue, and those near atoms less than the negative cutoff are red. All other dots are aqua. The cutoff parameter is defined under GROSS CHARGE in Section 4.3.6.

H BOND DON/ACC

The hydrogen bond donor/acceptor (H BOND DON/ACC) property colors the surface dots of hydrogen bond donors blue, acceptors red, and the remaining dots aqua. Should any surface dots belong to a residue listed as both a donor and an acceptor, they are colored yellow.

4.3.8 EXECUTE

The EXECUTE function computes the selected property at each dot position, creates a group to contain the surface dots and displays the group.

User Input:

1. Pick EXECUTE.

4.3.9 WRITE

The WRITE option creates a file containing the surface coordinates and property currently stored in memory - either generated by the last CALCULATE operation, or read in by a READ operation. The user is prompted for a file name to contain the dot surface. The last property which was computed is stored with the dot coordinates.

User Input:

1. Pick WRITE.
2. Enter a name for the dotted surface file.

4.3.10 WRITE PART

WRITE PART is like WRITE, but allows the reduction of the surface data written to just those atoms contained within certain groups.

User Input:

1. Pick WRITE PART.
2. Pick the group(s) which include the atoms to be selected for the surface data file from the Groups menu. Pick RETURN to complete the selection.
3. Enter the name of the dotted surface file to be output.

4.3.11 Surface File Format

The dotted surface or *dot* files generated under the CALCULATE option have a straightforward format. The first record has the keyword DOTS. The second record has the program file version number, *e.g.*, BIOGV300. The third record has a 20 character label indicating the property calculated on the surface. This label is the first 20 characters of either COLOR BY NO PROPERTY, COLOR BY ATOM TYPE, COLOR BY ELECTROSTATIC ENERGY, COLOR BY CHARGE or COLOR BY DON/ACC. The remaining records are written with the format shown below:

Record:	NATOM		X		Y		Z		NCOLOR
Format:	I5	1X	F8.3	1X	F8.3	1X	F8.3	1X	I3

where NATOM is the number of the atom associated with the dot, (X,Y,Z) are the x,y,z coordinates of the dot, and NCOLOR is the color of the dot. The colors correspond to fully saturated hues on a 360 degree wheel as listed in Table 4.1.

Color	NCOLOR
White	0
Purple	30
Magenta	60
Pink	90
Red	120
Orange	150
Yellow	180
Chartreuse	210
Aqua	270
Cyan	300
Light Blue	330
Blue	360

Table 4.1: Color wheel

Normally NCOLOR=0 would also be blue, but it has been redefined to be white. For dotted surface files with no property calculated, NCOLOR=0 for all dots.

4.4 SET ORIGIN

The options on the SET ORIGIN menu shown in Figure 4.12 are used to reset the center of rotation (FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, and ZERO ORIGIN), to center a selected file on the display (CENTER), to reset the dials or mouse to zero (RESET DIALS), and to reset the rotation/translation transformation matrices (RESET ALL).

4.4.1 FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN

These options are used to reset the center of rotation. FILE ORIGIN sets it to the geometric center of a selected input data file, GROUP ORIGIN sets it to the geometric center of a set of groups selected from the Groups menu, and PICK ORIGIN sets it to an atom picked from the display. ZERO ORIGIN sets the center of rotation to $x, y, z = 0,0,0$. These options are described in Sections 3.9.8 - 3.9.11.

4.4.2 CENTER

Toggling CENTER centers the selected file in the middle of the display by modifying its coordinates so that the average value of x, y , and z each equal 0.

User Input:

1. Pick CENTER.
2. Pick the file to be centered. This step is skipped if there is only one file in the system.

4.4.3 RESET DIALS

Picking RESET DIALS resets the dials or mouse to zero (zero rotation, zero translation, and zero zoom). All groups currently connected to the dials or mouse will move back to the positions they occupied before being last connected, *i.e.*, the last time ROTATE ALL or EXECUTE (under DOCK) was toggled. (**Warning:** The molecule may disappear from the display. If so, zoom out until the molecule appears in view; use the x and y -translate functions to position the molecule in the center of the display, then zoom in to fill the display.)

4.4.4 RESET ALL

Picking RESET ALL resets the dials or mouse to zero (see above) and resets the RT transformation to the unit matrix. All groups currently connected to the dials or mouse will move back to the positions they occupied before the last time MODIFY COORDS was toggled, *i.e.*, to the current atomic coordinates stored on disk. (**Warning:** The molecule may disappear from the display. If so, zoom out until the molecule appears in view; use the x and y -translate functions to position the molecule in the center of the display, then zoom in to fill the display.)

Set origin:
file origin
group origin
pick origin
zero origin
center
reset dials
reset all
return

Figure 4.12:
SET ORIGIN
menu

4.5 CHANGE COLORS

The CHANGE COLORS operations allow the user to change the color of individual groups. The groups must be displayed in a single color, however; CHANGE COLORS will not work on groups displayed with the half-bonded colors (see Section 4.1.10) or on surfaces displaying properties (see Section 4.3). After toggling CHANGE COLORS, the Groups menu appears and the user is asked to pick the group whose color is to be changed. Only one group can be modified at a time. After making the selection, a Colors menu appears as shown in Figure 4.13 with 13 predefined colors: purple, magenta, pink, red, orange, yellow, chartreuse, green, aqua, cyan, light blue, blue, and white. Selecting one of these colors immediately changes the color of the groups selected. The final color selected before toggling RETURN is kept.

4.5.1 RGB VALUE If none of the preselected colors is appropriate, then the user may specify the exact color using the RGB VALUE option. After toggling RGB VALUE a dialog box with several slider bars in it appears (see Figure 4.14). The user moves the sliders with the mouse to specify values for the red (R), green (G), and blue components (B). Changes in the R, G, B sliders are reflected in the hue (H). Hue values range from red through to purple then back to red. The hue specified in the example here (0.167) is yellow corresponding to values of 1, 1, and 0 for the red, green, and blue components respectively. The saturation (S) and value (V) may also be adjusted. The current color is shown in the box at the bottom. When satisfied with the color, the user picks OK.

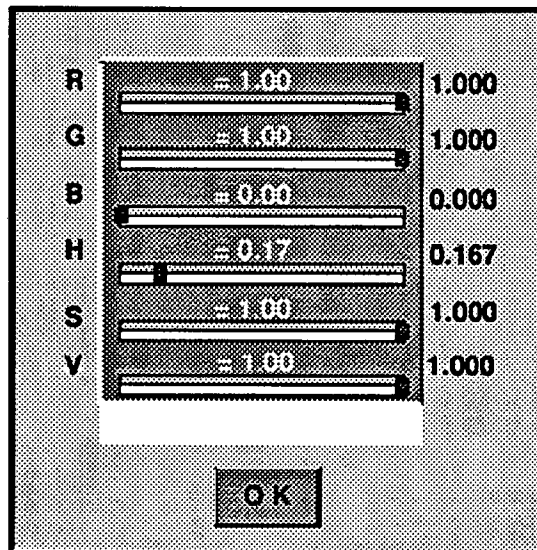


Figure 4.14: Dialog box for specifying exact color (RGB VALUE option)

Colors:
purple
magenta
pink
red
orange
yellow
chartreuse
green
aqua
cyan
light blue
blue
white
RGB value
return

Figure 4.13: Colors menu

User Input:

1. Pick **CHANGE COLORS**.
2. Pick the non-half-bonded group to be colored from the Groups menu.
3. Pick the desired group color from the Colors menu. To specify the exact color, pick **RGB VALUE** and use the sliders to specify values for the red (R), green (G), and blue (B) components; the hue (H), saturation (S), and value (V) may also be adjusted. When the box at the bottom shows the desired color, pick **OK**. (See Figure 4.14).
4. After the desired color has been selected, terminate by picking **RETURN**.

4.6 STEREO

The operations on the STEREO menu shown in Figure 4.15 are used to render the image in the display area as a stereo pair. The stereo pair can be displayed in either distal ("wall-eyed") or proximal ("cross-eyed") mode. Figure 4.16 shows a stereo pair view of the morphine molecule in proximal mode (the default), and Figure 4.17 shows stereo views in distal mode. The two images are rotated slightly in opposite directions about the y axis, so that they can be viewed together as a three-dimensional image by using a stereo viewer. (A trained user can easily see the three-dimensional image with the naked eye.)

Other functions allow the user to change the angle by which the two images are rotated, the separation of the images, and the distance between the edge of the screen and the outer edge of each image.

Stereo:
off
mode\
proximal
angle
2.850
separation
0.010
offset
0.000
return

Figure 4.15:
STEREO menu

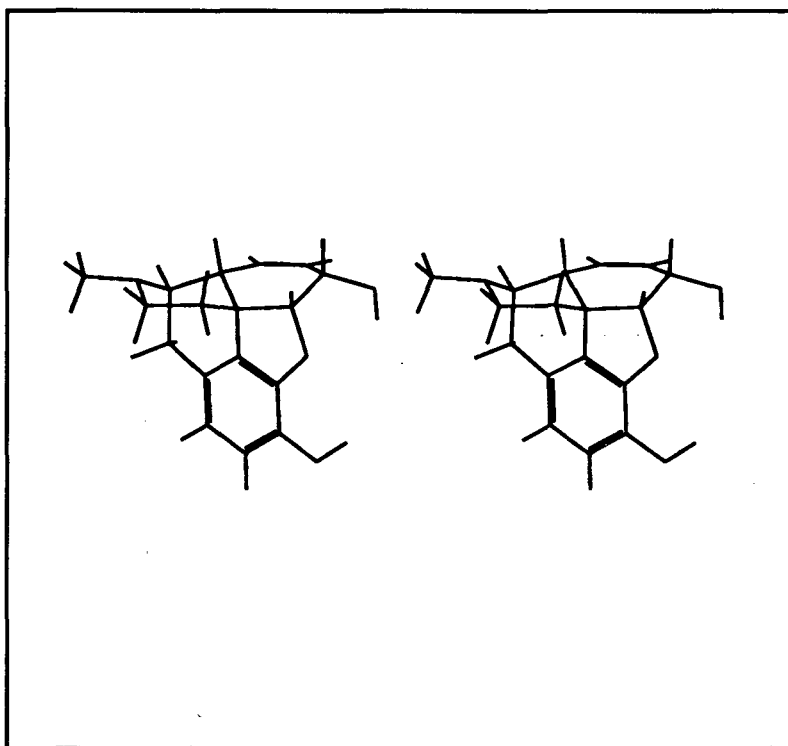


Figure 4.16: Proximal stereo view of morphine

4.6.1 ON/OFF

Toggling ON/OFF renders the image as either a stereo pair (ON) or a single image (OFF).

User Input:

1. Toggle the ON/OFF box to ON to render the image as a stereo pair, or toggle it to OFF to render it as a single image.

4.6.2 MODE

Toggling MODE allows the user to change the representation between distal ("wall-eyed") and proximal ("cross-eyed") mode (see Figures 4.16 and 4.17). The default mode is proximal. See Wood, William B., Wilson, John H., Benbow, Robert M., and Leroy H. Hood, *Biochemistry, A Problems Approach*, 2nd Edition (Menlo Park, California, Benjamin/Cummings, 1981), for a discussion of distal and proximal modes and a tutorial on viewing stereo pairs with the naked eye.

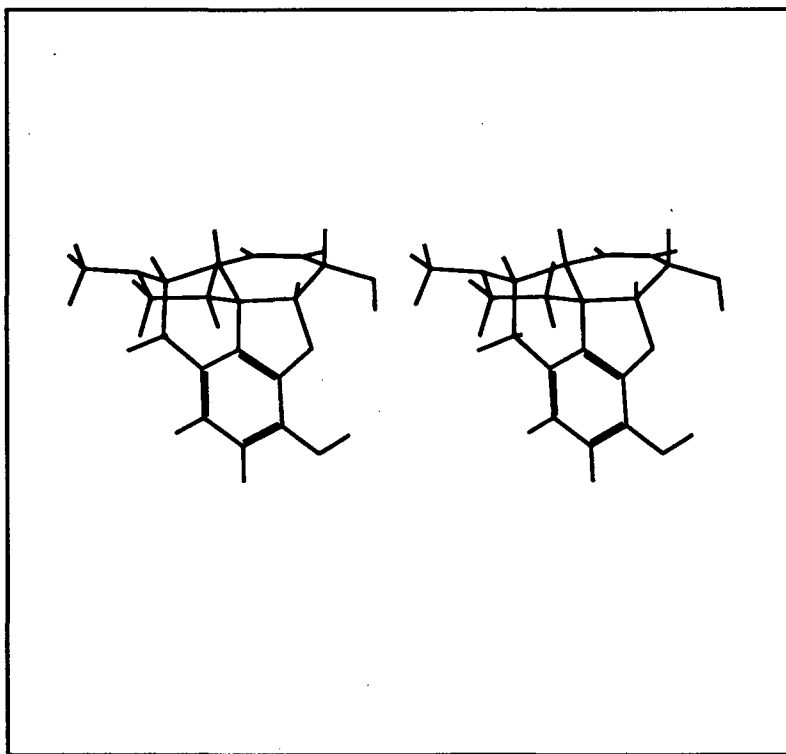


Figure 4.17: Distal stereo view of morphine

User Input:

1. Toggle MODE until the box beneath it shows the representation desired (PROXIMAL or DISTAL).

4.6.3 ANGLE

By default each stereo view is rotated 2.85° in opposite directions about the y axis. This angle of rotation gives the strongest stereo effect for people with average eye separation. (See Rogers, David F. and J. Alan Adams, *Mathematical Elements for Computer Graphics* (New York, McGraw-Hill, 1976), for a discussion of the mathematics of stereo pairs.) The angle of rotation can be reset by toggling the ANGLE function.

User Input:

1. Toggle ANGLE.
2. Enter the angle of rotation about the Y axis.

4.6.4 SEPARATION

The separation is the distance in the x direction between the inner edges of the two stereo views in screen coordinates (see Figure 4.18). The default value for the separation is 0.10, where the left edge of the display window has an x coordinate of -1.0 and the right edge of the display window has an x coordinate of +1.0. The separation value can be reset by toggling the SEPARATION function.

User Input:

1. Toggle SEPARATION.
2. Enter the separation between the two stereo views in screen coordinates.

4.6.5 OFFSET

The offset is the distance in screen coordinates from the outer edge of each stereo view to the nearest inner edge of the screen (see Figure 4.18). The default value for the offset is 0.000, meaning that each stereo view extends to the edge of the display window. The offset can be changed by toggling the OFFSET function.

User Input:

1. Toggle OFFSET.
2. Enter the offset of each stereo view from the edge of the display window in screen coordinates.

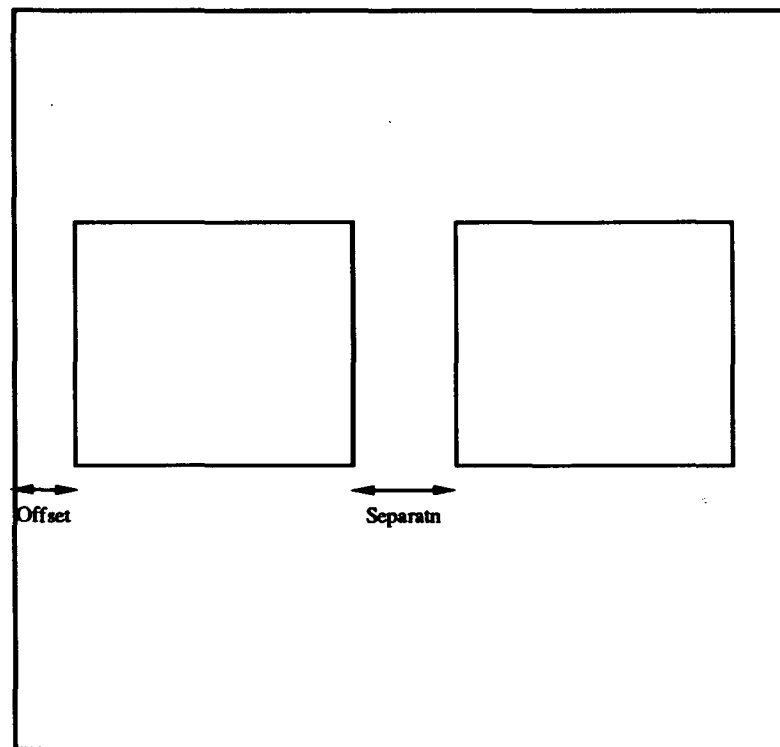


Figure 4.18: Separation and offset of stereo pairs

4.7 ROCK

Picking ROCK causes objects on the display to rock automatically about the y axis.

Warning: Extensive use of ROCK may cause the displayed groups to distort due to accumulated errors in rotation. Because of this, the dials are automatically reset when the function is terminated.

User Input:

1. Pick ROCK.
2. Pick anywhere on the display to terminate rocking.

4.8 AUTO ROTATE

Picking AUTO ROTATE causes objects on the display to rotate automatically about the y axis.

Warning: Extensive use of AUTO ROTATE may cause the displayed groups to distort due to accumulated errors in rotation. Because of this, the dials are automatically reset when the function is terminated.

User Input:

1. Pick AUTO ROTATE.
2. Pick anywhere on the display to terminate rotation.

Chapter 5

Simulate Operations

The operations on the **SIMULATE** menu shown in Figure 5.1 allow the user to perform energy optimizations and molecular dynamics calculations on structures, to generate new structures using conformational searches, and to perform constrained minimizations. One may also animate the sequences of conformations generated and perform statistical analysis on them. The calculations used are based on molecular mechanics, which describes atomic interactions in terms of classical force fields. A detailed description of the available force fields is presented in Appendix G. The default parameters used in the force fields are set when **COMSTRUCT** is run (see the *Molecular Simulations Installation Guide*). Appendix I describes the parameter file format. The user may edit the parameter files directly or modify them interactively (see Section 5.13). The Dreiding force field parameters are designed to give reasonable results on a wide variety of chemical systems.

The energy expression, which is the basis of all structure and dynamics calculations, must be set up first using **SETUP EEX**. Setting up the energy expression entails specifying which groups are to be included in the calculation and the extent of their inclusion. Atoms can have one of three dispositions.

Movable: atoms which contribute to the energy expression and can move during the calculation in response to the forces acting on them.

Fixed: atoms which contribute to the energy expression but cannot move during the calculation.

Excluded: atoms which do not contribute to the energy expression and do not move during the calculation.

Defaults are available which affect the way the energy expression is set up. If the **FAST SETUP** switch is on, the energy expression is automatically set up with all atoms considered movable. A **RIGID MOLECULE** switch is also available. When on, each molecule is treated as a rigid molecule. These switches are found under **EEX OPTIONS**. The energy expression can also be modified after the initial setup using **MODIFY EEX**. The operations under **MODIFY EEX** allow the user to

Simulate:
eex options
setup eex
modify eex
show eex
constraints
charges
mechanics
dynamics
constrained min
grid search
M-Carlo search
defaults
modify params
animate
return

Figure 5.1:
SIMULATE menu

reassign the disposition of atoms on an atom by atom basis (rather than a group by group basis as in SETUP EEX). SHOW EEX shows how the energy expression has been set up: the movable atoms are marked with green crosses and the fixed atoms are marked with blue crosses. CONSTRAINTS is used to set harmonic constraints on the interatomic distances and angles of the structure, and CHARGES is used to reassign atomic charges.

Once the energy expression has been set up, standard energy optimizations and simple molecular dynamics calculations can be performed under MECHANICS. Extended molecular dynamics calculations can be performed under DYNAMICS. The CONSTRAINED MIN operation allows the user to perform constrained minimizations which simulate reaction pathways. The GRID SEARCH and M-CARLO SEARCH options allow the user to do conformational searches.

The DEFAULTS menu allows the user to specify the parameters that govern energy calculations, such as the choice of which energy terms to calculate (bonds, angles, *etc.*) and the simulation temperature for molecular dynamics trajectories. The MODIFY PARAMS menu allows the user to examine and modify the force field parameters used to calculate the energy of each bond, angle, torsion, inversion, nonbonding interaction, and hydrogen bond in the structure. The ANIMATE option allows the user to display, animate and energy analyze dynamics trajectories, conformational searches and constrained minimizations.

5.1 EEX OPTIONS

The EEX OPTIONS menu shown in Figure 5.2 allows the user to set the defaults which will be used when setting up the energy expression and performing the energy calculations. FAST SETUP is an on/off toggle. When on, the energy expression for the structure is set up quickly (all atoms are movable). When off, the user must select which groups are to be movable and which are to be fixed. RIGID MOLECULE is also an on/off toggle. When on, each molecule will be treated as a rigid body during the energy calculations. This is useful as a first step in minimizing the structure for a collection of molecules (*e.g.*, liquid crystals). This should be followed by a full minimization (with all atoms movable).

EEX options:
fast setup
rigid molecule
return

Figure 5.2: EEX
OPTIONS menu

5.1.1 FAST SETUP

If FAST SETUP is on, the SETUP EEX operation will be done quickly. That is, all atoms in the structure will be considered movable. If FAST SETUP is off, then the user must select the groups which are movable and the groups which are fixed. See SETUP EEX, Section 5.2. The default is off.

User Input:

1. Toggle FAST SETUP on to have all atoms automatically be considered movable when the energy expression is setup.

5.1.2 RIGID MOLECULE

If RIGID MOLECULE is on, each molecule is treated as a rigid body during the energy calculations. The atoms within the molecule are not allowed to move with respect to each other. Thus the molecule is defined by specifying that only three translations and three rotations be allowed (about the x, y, and z axes). This is useful for considering a large number of solvent molecules and for determining packing of molecules. It is also useful as a first step in minimizing the structure for a collection of molecules (*e.g.*, liquid crystals), but should be followed by a full minimization (with all atoms movable). The default is off.

In version 3.0, this option is available only for minimizations, not for dynamics.

User Input:

1. Toggle RIGID MOLECULE on to have each molecule be treated as a rigid body.

5.2 SETUP EEX

The SETUP EEX option allows the user to set up the energy expression by choosing groups containing atoms that are to be movable and groups containing atoms that are to be fixed. All other atoms will be excluded. If an atom is in both a group designated as movable and a group designated as fixed, it is considered a movable atom. That is, groups containing movable atoms have priority over groups containing fixed atoms (this priority can be changed using the options under MODIFY EEX, Section 5.3). Groups picked to be updated during the calculation will have their display positions modified as the calculation proceeds.

If there is only one group in the system or if FAST SETUP is on (under EEX OPTIONS) then the energy expression is automatically set up with all atoms movable.

Note: During setup, the program checks that force field parameters have been defined for torsions and inversions. If there are no defined parameters for *any* torsions (IJKL) involving the central pair (JK) the program gives a warning message. If the USE ALL TORSION switch under ENERGY VAR is on, the program will insist upon all parameters for the torsion being defined.

Similarly, if there are no defined parameters for *any* inversions (IJKL) where I is the central atom the program gives a warning message. If the USE ALL INVERS switch under ENERGY VAR is on, the program will insist upon all parameters for the inversion being defined.

User Input:

1. Pick SETUP EEX.
If there is only one group in the system or if FAST SETUP is on, steps 2-4 are skipped.
2. Pick the group(s) containing movable atoms from the Groups menu. Pick RETURN to complete the selection.
3. Pick the group(s) containing fixed atoms from the Groups menu. Pick RETURN to complete the selection.
4. Pick the group(s) to be updated during calculation from the Groups menu. Pick RETURN to complete the selection.

5.3 MODIFY EEX

The MODIFY EEX operations shown in Figure 5.3 allow the user to modify the energy expression set up under SETUP EEX by changing the disposition of atoms to movable, fixed or excluded. MOVE ALL makes all atoms movable, MOVE RANGE changes the disposition of a range of atoms to movable, and MOVE SINGLE changes it for selected single atoms. FIX ALL fixes all atoms, FIX RANGE makes a range of atoms fixed, FIX SINGLE fixes selected single atoms, and FIX GROUPS fixes selected groups. EXCLUDE RANGE changes the disposition of a range of atoms to be excluded, and EXCLUDE SINGLE does the same for selected single atoms.

The energy expression is automatically updated when the user exits the MODIFY EEX menu. The disposition of the atoms can be displayed by toggling SHOW EEX (see Section 5.4). The movable atoms are marked with green crosses, while the fixed atoms are marked with blue crosses.

Note: When some of the atoms are fixed, these fixed atoms act as an external force on the movable atoms in the system. Thus, the net force and torque are not zero and the total linear momentum and angular momentum are not conserved.

Modify eex:
move all
move range
move single
fix all
fix range
fix single
fix groups
exclude range
exclude single
return

Figure 5.3:
MODIFY EEX
menu

5.3.1 MOVE ALL, FIX ALL

Toggling MOVE ALL makes *all* atoms movable. Toggling FIX ALL makes *all* atoms fixed.

5.3.2 MOVE RANGE, FIX RANGE, EXCLUDE RANGE

The MOVE RANGE, FIX RANGE and EXCLUDE RANGE options enable the user to change the disposition of a range of atoms to movable, fixed and excluded, respectively. (See Section 1.6 for information on picking a range of atoms.)

User Input:

1. Pick MOVE RANGE, FIX RANGE, or EXCLUDE RANGE.
2. Pick a range of atoms from the display as described in Section 1.6.

5.3.3 MOVE SINGLE, FIX SINGLE, EXCLUDE SINGLE

The MOVE SINGLE, FIX SINGLE and EXCLUDE SINGLE options enable the user to change the disposition of single atoms to movable, fixed and excluded, respectively.

User Input:

1. Pick MOVE SINGLE, FIX SINGLE, or EXCLUDE SINGLE.
2. Pick a single atom twice from the display.

5.3.4 FIX GROUPS

FIX GROUPS fixes selected groups.

User Input:

1. Pick **FIX GROUPS**.
2. Pick the groups to be fixed from the Groups menu. Pick **RETURN** to complete the operation.

5.4 **SHOW EEX**

Toggling **SHOW EEX** shows how the energy expression was set up. Atoms selected as movable are indicated with green crosses, while fixed atoms are marked with blue crosses. Other relevant information is also presented in the text window.

User Input:

1. Pick **SHOW EEX** to display the movable and fixed atoms and to display other relevant information.

5.5 CONSTRAINTS

The CONSTRAINTS menu shown in Figure 5.4 allows the user to enter constraints that force particular bonds, angles, torsions, and/or inversions to remain at or near their constrained values. The user can constrain to the equilibrium values given by the force field parameters and current molecular structure or enter new values. For proteins, the peptide bonds may be identified and all the ω angles may be constrained to be all *trans* or all *cis*; similarly, all the amino acids may be constrained so that they are in the D- or L-form. Before entering constraints, the energy expression must be set up (see Section 5.2). Constraints do not replace terms in the energy expression.

Constraint entry is cumulative: when new constraints are added, the previous constraints are not removed. For angle, torsion, and inversion constraints, the order in which the atoms are picked is significant. Equivalent (but not symmetric) angles are detected. To change a constraint, reenter that constraint. To delete a single constraint, set its force constant to zero (or RESET the whole list).

The user may save the data describing the constraints with the WRITE CONSTRAINT option. This constraint data may then be read in later using READ CONSTRAINT. A special constraints file format has been developed for storing this information (see Appendix D.6).

Energy optimization using constraints can be used to find the closest match between the conformations of two related structures. To perform optimized template matching, define one structure as movable and the other as fixed, and define distance constraints between key atoms of the two structures. When the energy of the system is minimized, the conformation of the movable structure will be modified to get the best possible fit to the fixed (template) structure.

5.5.1 DISTANCE

The DISTANCE option allows the user to define constraints based on the distance between atoms. The atoms do not have to be bonded. The current distance constraint force constant is automatically used for each new constraint. Toggle K(DISTANCE) to define a new distance constraint force constant (see Section 5.5.5).

User Input:

1. Pick the first atom of the constrained distance.
2. Pick the second atom of the constrained distance.
3. Enter the equilibrium distance, or type *cur* to use the current distance.

Constraints:
distance
k(distance)
10000.000
angle
k(angle)
1000.000
torsion
k(torsion)
1000.000
inversion
k(inversion)
1000.000
reset
read constraint
write constraint
show all
protein constrnt
return

Figure 5.4:
CONSTRAINTS
menu

4. Pick the first atom of the next constrained distance, or toggle off the molecule to terminate.

5.5.2 ANGLE

Toggle ANGLE to define constraints based on angles between atoms. The atoms do not have to be bonded.

User Input:

1. Pick the first atom of the constrained angle.
2. Pick the second (middle) atom of the constrained angle.
3. Pick the third atom of the constrained angle.
4. Enter the constraint equilibrium angle, or type *cur* to use the current value.
5. Pick the first atom of the next constrained angle, or toggle off the molecule to terminate.

5.5.3 TORSION

Toggle TORSION to define constraints based on torsion angles between atoms. The atoms do not have to be bonded. A torsion angle of 0° constrains atoms 1 and 4 of the torsion angle to be in the *cis* conformation, while an angle of 180° constrains them in the *trans* conformation.

User Input:

1. Pick the first atom of the constrained torsion angle.
2. Pick the second atom of the constrained torsion angle.
3. Pick the third atom of the constrained torsion angle.
4. Pick the fourth atom of the constrained torsion angle.
5. Enter the constraint torsion angle, or type *cur* to use the current value.
6. Pick the first atom of the next constrained torsion, or toggle off the molecule to terminate.

5.5.4 INVERSION

Toggle INVERSION to define constraints based on inversion angles between atoms (improper torsions). The atoms do not have to be bonded.

The inversion angle is defined as the angle between the plane of picked atoms IJK and the plane of atoms IJL, where I is the central atom bonded to atoms J, K and L, and J is the special outer atom (see Figure G.7 in Appendix G). Thus, planar structures have an inversion angle of 0° and tetrahedral structures have angles of 60° or -60°.

User Input:

1. Pick the first (center) atom of the constrained inversion.
2. Pick the second (special outer) atom of the constrained inversion.
3. Pick the third atom of the constrained inversion.
4. Pick the fourth atom of the constrained inversion.
5. Enter the constraint inversion angle, or type *cur* to use the current value.

6. Pick the first atom of the next constrained inversion or toggle off the molecule to terminate.

5.5.5 K(DISTANCE), K(ANGLE), K(TORSION), K(INVERSION)

All constraints use a harmonic force field and each is defined with an individual force constant. Therefore the rigidity of each constraint can be controlled by changing the corresponding force constant before defining the constraint. The default force constants produce very rigid structures. The default values are as follows in (kcal/mol)/Å²: distance = 10,000, angle = 1000, torsion = 1000, and inversion = 1000. (The default values were chosen to be about one order of magnitude larger than the normal values.) Changing a force constant applies to the next constraints of that type and does not change the force field of constraints already defined.

User Input:

1. Pick K(DISTANCE), K(ANGLE), K(TORSION), or K(INVERSION).
2. Enter the harmonic force constant to be applied to new constraints.

5.5.6 RESET

The RESET option removes all constraints.

User Input:

1. Pick RESET to remove all constraints.
2. Pick YES or NO to confirm the RESET operation.

5.5.7 READ CONSTRAINT

The READ CONSTRAINT option allows the user to restore the constraints to the values previously saved for the structure (with the WRITE CONSTRAINT option). The same files must be in the system and they must be in the same order as when the constraints file was written.

1. Pick READ CONSTRAINT.
2. Enter the name of the constraints file used to store the constraint values for this structure.

5.5.8 WRITE CONSTRAINT

The WRITE CONSTRAINT option allows the user to save the constraints data describing the distance, angle, torsion, and inversion constraints to be imposed on a structure for an energy calculation. This data is stored in a constraints (or *con*) file which may be read back in later using the READ CONSTRAINT option. This eliminates the need to redefine the constraints for this structure each time the program is started.

User Input:

1. Pick WRITE CONSTRAINT.
2. Enter a name for the constraints file being saved. Use *con* for the file name extension.

5.5.9 SHOW ALL/SHOW CURRENT

When this switch is toggled to read **SHOW ALL**, the default values of all existing constraints of the current constraint type are shown on the screen while the atoms for a constraint are being picked. When the switch is toggled to **SHOW CURRENT**, previously entered constraints of the same type are not shown. In either case, after all of the atoms for a constraint have been picked, the old constraints are no longer displayed on the screen and the current value of the selected distance or angle is displayed.

5.5.10 PROTEIN CONSTRNT

The **PROTEIN CONSTRNT** option allows the user to set torsion and inversion angle constraints about the peptide bonds in a molecule. The **PROTEIN CONSTRNT** menu is shown in Figure 5.5. The ω angles may be constrained to their current values, or they may be constrained to be all *trans* (180°) or all *cis* (0°). Inversion angle constraints may be set for all C_α atoms that have a side chain: they may be constrained to their current values or constrained so that all the amino acids are either in the D- or L-form. As in the main **CONSTRAINTS** menu, the rigidity of each constraint can be controlled by changing the corresponding force constant before setting the constraint. Before setting constraints, the energy expression must be set up (see Section 5.2).

RESET TORSION

The **RESET TORSION** option removes all torsion angle constraints; **RESET INVERSION** removes all inversion angle constraints.

User Input:

1. Pick **RESET TORSION** to remove all torsion angle constraints; pick **RESET INVERSION** to remove all inversion angle constraints.

K(TORSION)

K(TORSION) is the harmonic force constant for torsions which is to be applied to the new constraints (see Section 5.5.5 above).

ALL TRANS, ALL CIS, KEEP TRANS/CIS

The **ALL TRANS**, **ALL CIS** and **KEEP TRANS/CIS** options apply to all movable groups. They identify all the peptide groups in the structure and add torsion angle constraints to the four main chain atoms about the peptide bond (*i.e.*, $C_\alpha-C-N-C_\alpha$). **ALL TRANS** constrains the ω angles to 180°; **ALL CIS** constrains them to 0°; **KEEP TRANS/CIS** constrains the ω angles to their current values.

Protein constr nt:
reset torsion
k(torsion)
1000.000
all trans
all cis
keep trans/cis
reset inversion
k(inversion)
1000.000
all l
all d
keep l/d
return

Figure 5.5:
PROTEIN
CONSTRNT
menu

User Input:

1. Pick ALL TRANS or ALL CIS to constrain the ω angles to 180° or 0° respectively; pick KEEP TRANS/CIS to constrain the ω angles to their current values.

RESET INVERSION

RESET INVERSION removes all inversion angle constraints.

User Input:

1. Pick RESET INVERSION to remove all inversion angle constraints.

K(INVERSION)

K(INVERSION) is the harmonic force constant for inversions which is to be applied to the new constraints (see Section 5.5.5 above).

ALL L, ALL D, KEEP L/D

The ALL L, ALL D, and KEEP L/D options all set inversion angle constraints about the C_α atoms that have a side chain (*i.e.*, no constraints are set for glycine). The inversion angle which is constrained is $C_\alpha-N-C_\beta-C$. ALL L keeps all the amino acids in the L-form (-35°); ALL D keeps them in the D-form (+35°). KEEP L/D constrains these inversion angles to their current values.

User Input:

1. Pick ALL L or ALL D to constrain the amino acids to the L-form or D-form respectively; pick KEEP TRANS/CIS to constrain them to their current values.

5.6 CHARGES

The CHARGES option is described under the Organic Builder's EDIT ID option. See Section 3.1.13.

5.7 MECHANICS

The MECHANICS operations shown in Figure 5.6 allow the user to calculate the energy of the current geometry (ONE ENERGY), to minimize the energy of the structure (MINIMIZE), and to run simple molecular dynamics (SIMPLE DYNAMICS). He may also change the number of steps to be performed in a minimization or dynamics calculation (# OF STEPS), reset the convergence criteria (RMS FORCE CVRG), or modify other simulation defaults (DEFAULTS). Options are also available under GEOMETRY PROP which allow the user to analyze the geometric properties of the structure. The XYZ TEMP SAVE and XYZ RESTORE options are provided as a safeguard allowing the user to internally save the current coordinates for all the atoms in all the files and then restore them later.

Note: The energy expression must be set up before any calculations can be performed. This is done using SETUP EEX (under SIMULATE). If the MECHANICS menu is entered from one of the builders, however, the energy expression is automatically setup with all atoms movable.

The most time consuming portion of a minimization or dynamics calculation based on molecular mechanics is the calculation of the long range nonbond and hydrogen bond interactions. For the nonbonded interactions (van der Waals and electrostatic) the order of the number of terms increases as the square of the number of atoms. Distance cutoffs are used to reduce the number of terms. The program calculates a list of all the nonbond and hydrogen bond interactions within the cutoffs. During a minimization or dynamics run, the nonbond and hydrogen bond lists need to be updated as atoms move. The nonbond and hydrogen bond lists are automatically updated on a regular basis determined by the update frequency parameter (UPDATE FREQ), typically every 50 steps. It is wise to initially keep the update frequency small so that these lists are updated frequently at the beginning when atoms are moving around a lot. Once the structure is closer to a local minimum and very little change in energy is observed, then longer minimization runs are appropriate.

For single energy calculations and energy minimizations, the Simulation results menu (see Figure 5.7) lists several key quantities, such as the current total energy, the change in energy for the last step of minimization (delta energy), the root-mean-square energy derivative of the atomic coordinates (RMS force), and the breakdown of the total energy into bonds, angles, torsions, inversions, van der Waals, electrostatic, hydrogen bonds, constraints, and user energy. The energies are given in kcal/mol. Similar information is displayed in the text window at each step for energy minimizations. Energy information is also displayed in the text window at every step for calculations of a dynamics trajectory, but the Simulation results menu is not updated.

Mechanics:
one energy
minimize
simple dynamics
of steps
50
rms force cvrg
0.100
defaults
geometry prop
xyz temp save
xyz restore
return

Figure 5.6:
MECHANICS
menu

5.7.1 ONE ENERGY

Toggling ONE ENERGY calculates the energy of the current conformation and displays it on the Simulation results menu (see Figure 5.7). Key quantities listed are: the current total energy, the change in energy for the last step of minimization (delta energy), the root-mean-square energy derivative of the atomic coordinates (RMS force), and the breakdown of the total energy into bonds, angles, torsions, inversions, van der Waals, electrostatic, hydrogen bonds, constraints, and user energy. The energies are given in kcal/mol.

5.7.2 MINIMIZE

Toggling MINIMIZE initiates # OF STEPS steps of energy minimization. The energies are updated on the Simulation results menu during the calculation. The minimization can be aborted at any time with the abort key (F10 or F12); it will stop after the current step is completed. The energy can be minimized using the conjugate gradient, Fletcher-Powell, or steepest descents algorithms set under the MINIMIZE VAR menu (see Section 5.12.2). The default is conjugate gradient.

Simulation results:	
total energy	- - -
delta energy	- - -
RMS force	- - -
bonds	- - -
angles	- - -
torsions	- - -
inversions	- - -
van der Waals	- - -
electrostatic	- - -
H bonds	- - -
constraints	- - -
user	- - -

Figure 5.7: Simulation results menu

5.7.3 SIMPLE DYNAMICS

Toggling SIMPLE DYNAMICS initiates # OF STEPS steps of molecular dynamics. The function is called SIMPLE DYNAMICS because it does dynamics without quenching, annealing, or an impulse direction. This is the same as the ADIABATIC dynamics option on the DYNAMICS menu (see Section 5.8.2).

5.7.4 # OF STEPS

The # OF STEPS option allows the user to change the number of steps performed in an energy minimization or simple molecular dynamics calculation.

User Input:

1. Pick # OF STEPS.
2. Enter the new value for # OF STEPS (default is 50).

5.7.5 RMS FORCE CVRG

The RMS FORCE CVRG option allows the user to reset the criterion for convergence that is used in minimizations. If the residual total force is less than RMS FORCE CVRG, then the program will print a message in the output listing and stop the minimization. The default is 0.1 (kcal/mol)/Å.

User Input:

1. Toggle RMS FORCE CVRG.
2. Enter the new convergence criterion for residual total force (default is 0.1).

5.7.6 DEFAULTS

Toggling DEFAULTS displays the Simulation DEFAULTS menu. See Section 5.12 for more information.

5.7.7 GEOMETRY PROP

The options on the Geometry properties menu shown in Figure 5.8 are on/off toggles which allow the user to specify whether geometric analysis will be performed (ANALYZE BY GEOM), and to specify which parameters will be included in the analysis (BONDS, ANGLES, TORSIONS, INVERSIONS, HYDROGEN BONDS, and NONBONDS). Analysis results are shown in the text window during energy calculations (for minimizations, the analysis is made at the last step). If the ANALYZE BY FF switch is on, all bonds or angles are grouped together by force field type (*e.g.*, all C_3-N_3 bond distances are averaged together). Cutoff distances (RCUTOFF) may also be specified to limit the hydrogen bond and nonbond interactions which are included. Default values for RCUTOFF are 3.25 Å for HYDROGEN BONDS and 2.75 Å for NONBONDS.

ANALYZE BY GEOM

Toggle ANALYZE BY GEOM on to do geometry analysis. Otherwise toggle it off.

ANALYZE BY FF

Toggle ANALYZE BY FF on to have all bonds or angles grouped by force field type (*e.g.*, all C_3-N_3 bond distances are averaged together). Otherwise toggle it off.

Geometry prop:
analyze by geom
analyze by FF
bonds
angles
torsions
inversions
hydrogen bonds
rcutoff=3.250
nonbonds
rcutoff=2.750
return

Figure 5.8:
Geometry
properties menu

BONDS, ANGLES, TORSIONS, INVERSIONS

Toggle the parameter(s) to be included in the geometry analysis on. Values listed are bond lengths for BONDS and angles for the other three parameters. Note that in order for ANGLES to be operative, BONDS must also be analyzed (toggled on). Similarly, TORSIONS and INVERSIONS require that BONDS and ANGLES be on.

HYDROGEN BONDS

Toggle HYDROGEN BONDS on to have all hydrogen bonds up to RCUTOFF included in the analysis (atom identification numbers and distance are listed for each pair). Otherwise, toggle it off. If ANALYZE BY FF is on, the program lists only the smallest distance for each pair of force field types. The default RCUTOFF value (shown in the menu box directly below HYDROGEN BONDS) is 3.25 Å and can be changed if desired (see below).

RCUTOFF

RCUTOFF is the cutoff distance used for HYDROGEN BONDS. It is the item directly beneath HYDROGEN BONDS on the menu. Hydrogen bonds more than RCUTOFF apart are not included in the analysis. The default is 3.25 Å.

User Input:

1. Pick RCUTOFF.
2. Enter the new cutoff distance for hydrogen bond analysis (in Å).

NONBONDS

Toggle NONBONDS on to have all nonbond interactions up to RCUTOFF distance apart included in the analysis (atom identification numbers and distance are listed for each pair). Otherwise, toggle it off. If ANALYZE BY FF is on, the program lists only the smallest distance for each pair of force field types. The default RCUTOFF value (shown in the menu box directly below NONBONDS) is 2.75 Å and can be changed if desired (see below).

RCUTOFF

RCUTOFF is the cutoff distance used for NONBONDS. It is the item directly beneath NONBONDS on the menu. Nonbond interactions more than RCUTOFF distance apart are not included in the analysis. The default is 2.75 Å.

User Input:

1. Pick RCUTOFF.
2. Enter the new cutoff distance for nonbond analysis (in Å).

5.7.8 XYZ TEMP SAVE, XYZ RESTORE

The XYZ TEMP SAVE option provides a safeguard allowing the user to internally save the current coordinates for all the atoms in all the files. Minimizations, dynamics or other operations can be done on the structure. If the user then wishes to return to the saved conformation, he simply picks the XYZ RESTORE option and the previously saved coordinates are restored. Any subsequent operation that redisplay the molecule will then display the saved conformation.

User Input:

1. Pick XYZ TEMP SAVE to save the current coordinates for all the atoms in all the files.
2. Pick XYZ RESTORE to restore the previously saved coordinates.

5.8 DYNAMICS

In contrast to molecular mechanics, molecular dynamics methods enable us to investigate the behavior and properties of model systems at non-zero temperatures. Molecular mechanics minimizations result in structural configurations which correspond to the minimum energy of the system. In molecular dynamics, however, one can investigate the effect of thermal motion or applied external pressure on the system. Studying structures at elevated temperatures is usually more appropriate since this represents the natural state.

In molecular dynamics one generally uses an iterative procedure to solve the Newtonian equations of motion ($F = ma$) to obtain the velocities and positions of the atoms as a function of time. If no external forces are applied, the total energy of the system remains constant. Most natural phenomena occur under conditions where a system is exposed to external pressure and/or exchanges heat with the environment. Under these conditions, the total energy of the system is no longer conserved and extended forms of molecular dynamics are required. The program provides several different types of equilibrium molecular dynamics (see Table 5.1). The constant pressure forms (types 2 and 4) are not available in BIOGRAF since they apply only to periodic systems.

Dynamics:
adiabatic
quenched
annealed
impulse
canonical (TVN)
temp damp (TVN)
return

Figure 5.9:
DYNAMICS
menu

Type	Options Used			
1	Constant energy, constant volume	(EVN)	adiabatic	ADIABATIC
2	Constant enthalpy, constant pressure	(HPN)	adiabatic	
3	Constant temperature, constant volume	(TVN)	isothermal	CANONICAL (TVN) TEMPDAMP (TVN)
4	constant temperature, constant pressure	(TPN)	isothermal	
E = Energy, V = volume, N = number of atoms, H = enthalpy ($E + PV$) P = pressure, T = temperature				

Table 5.1: Types of Equilibrium Molecular Dynamics

The first two types of dynamics are adiabatic (do not exchange heat), while the third and fourth types are isothermal (exchange heat with a thermal bath). Standard Newtonian equations are used only for the first type. The last three types are called extended molecular dynamics methods since they include additional dynamic variables. For example, for the constant pressure types of dynamics, the additional variables are volume (V) and dV/dt . The forms of molecular dynamics are de-

scribed in more detail in Appendix P. In addition to the methods listed above, the program also provides annealed, quenched, and impulse dynamics. These different dynamics methods are implemented via the options on the DYNAMICS menu shown in Figure 5.9.

The ADIABATIC option is used to perform constant energy, constant volume dynamics (type 1). The forces on each atom are calculated from the strengths of its interactions with the other atoms of the structure. A new set of coordinates and velocities for each atom at a time Δt later are calculated from the forces, current positions and current velocities. This procedure is repeated for each new time increment Δt . While minimization computes the forces on the atoms and changes their positions to minimize the interaction energies, dynamics computes the forces and moves the atoms in response to the forces.

The time step for the calculation is typically in the range of 0.001 picoseconds. The user may choose to write the coordinates and velocities of the structure at intervals during the dynamics run to a trajectory file. This trajectory file can later be analyzed using the TRAJECTORY option under ANALYZE (see Section 6.7) or the conformations can be displayed or extracted using the functions on the ANIMATE menu (see Section 5.14.)

Two options are available for running isothermal dynamics: CANONICAL (TVN) and TEMPDAMP (TVN). CANONICAL (TVN) uses an implementation of the Nosé-Hoover^{1,2} method. In order to maintain a constant thermodynamic temperature (not kinetic temperature) the velocities are scaled by an additional variable, s , which represents the interaction of the system with a heat bath. The rate of change of s is determined by the difference between the thermodynamic and the kinetic temperature at each step. TEMPDAMP (TVN) maintains a constant thermodynamic temperature through weak coupling to a thermal bath via the Berendsen *et al*³ approach.

In QUENCHED dynamics, after each period of adiabatic dynamics the energy of the structure is minimized (quenched) for a number of steps. In ANNEALED dynamics, the temperature of the structure is adjusted progressively from the starting temperature to the final temperature. The temperature is changed by adjusting the kinetic energy of the structure, *i.e.*, the velocities of the atoms are rescaled. A number of steps of dynamics is then run starting at the new temperature. When the final temperature has been reached, the temperature is progressively adjusted back to the starting temperature of the run with a number of steps of dynamics at each incremental temperature.

The ANNEALED option allows the energy of the structure to be minimized gradually without trapping the structure in a conformation that has a lower energy than nearby conformations but higher energy than more distant conformations (*i.e.* a local energy minimum). The structure with the lowest potential energy from the annealing cycle is minimized after the end of the cycle, and then written to the trajectory file.

At the beginning of an IMPULSE dynamics run, velocities are assigned to each atom in the direction vector that the user has set. The velocities are assigned so as to give the structure a tempera-

¹ S. Nose, *J. Chem. Phys.*, **91** (1984) 511, and *Mol. Phys.*, **52** (1984) 255.

² W. H. Hoover, *Phys. Rev. A.*, **31** (1985) 1695.

³ Berendsen, H. J., Postma, J. P. O., van Gunsteren, W. I., Di Niola, A. and Haak, J. R., *J. Chem. Phys.*, **81** (1984), 3684.

ture equal to a translational temperature **TRANSLAT TEMPER**, a temperature the user can set under the **IMPULSE** menu. This gives the structure an impulse along the direction vector. Then simple dynamics calculations are performed on the structure for the rest of the run without velocity adjustments. **IMPULSE** dynamics runs are often used to push interacting molecules over transition barriers while allowing energy relaxation corrections to occur.

The dynamics run can be aborted at any time with the abort key (F10 or F12). The run will stop after the current step is completed.

One can restart dynamics calculations and essentially mimic a continuous run by using the **APPEND TRAJECT** option located on the dynamics menus. The last set of velocities from the trajectory file created by the first run are used as starting velocities when dynamics is restarted.

5.8.1 Functions on All Dynamics Menus

Several menu options are present on all or most of the dynamics menus. These common options are described briefly here and in more detail in the following sections.

TEMPERATUR VAR - Calls the **TEMPERATUR VAR** (temperature variables) menu of the **DEFAULTS** menu.

DYNAMICS VAR - Calls the **DYNAMICS VAR** (dynamics variables) menu of the **DEFAULTS** menu.

TIME - Allows the user to set the length of the dynamics run.

USE VELOCITY - Uses the velocities for an extracted conformation for the next dynamics run rather than assigning the velocities randomly (see **EXTRACT** under **ANIMATE**, Section 5.14).

RESET VELOCITY - Assigns random velocities before continuing a dynamics run rather than using the last set of velocities as starting velocities.

WRITE TRAJECT, WRITE VELOCITY - Determines whether a trajectory file containing the atomic coordinates, energies, and velocities for the dynamics run are written to disk.

APPEND TRAJECT - Allows the user to add information to the end of an existing trajectory file and to restart dynamics using the last set of velocities.

EXECUTE - Begins the dynamics run.

TEMPERATUR VAR

The TEMPERATUR VAR (temperature variables) menu is shown in Figure 5.10. It allows the user to set the initial and final temperatures, heating increment, allowed temperature window, and annealing temperature increment. Section 5.12.3 gives instructions for this menu.

At the start of a dynamics run velocities are assigned to give the structure a temperature equal to the initial temperature (INITIAL TEMPER). The velocities are raised (or lowered) in steps during the run so that the final temperature of the run is equal to the value under FINAL TEMPER. The default value for both the initial and final temperatures is 300° K. If the values for the initial and final temperature differ, then the temperature is adjusted after each time period as specified by the heating/equilibration frequency (HEAT EQUIL FREQ, default is 0.1 ps) on the DYNAMICS VAR (dynamics variables) menu (see Figure 5.11). The temperature is incremented by the value that the user enters for HEAT INCREMENT (heating increment).

After HEAT EQUIL FREQ period of dynamics the program compares the average temperature over the period to the final temperature desired. If the difference between the average temperature and the desired temperature is greater than the absolute value of TEMPER WINDOW, velocities are reassigned to bring the temperature of the structure back to the desired temperature. If the average temperature falls within the TEMPER WINDOW no temperature readjustment is performed. The ANNEAL INCREMENT option allows the user to set the temperature increment for each step of annealing.

User Input:

1. Toggle TEMPERATUR VAR to bring up the TEMPERATUR VAR (temperature variables) menu.

Temperatur var:
initial temper
300.000
final temper
300.000
heat increment
20.000
temper window
50.000
anneal increment
10.000
return

Figure 5.10:
TEMPERATUR
VAR menu

DYNAMICS VAR

The DYNAMICS VAR (dynamics variables) menu shown in Figure 5.11 is used to control the time step for dynamics runs, set several kinds of update frequencies, and set the level of information written out during batch dynamics runs. More information is given in Section 5.12.4.

The DRAW FREQUENCY option allows the user to set the number of steps of dynamics to perform before updating the positions of the atoms on the screen. The default value is 5, causing an update of the screen after every five steps of dynamics. The TIME STEP option determines the time step for each time integration cycle in picoseconds. The default value is 0.001 picoseconds. As Section 5.10.3 explains, the time step must be kept fairly small to generate a reliable dynamics trajectory. The ANNEAL FREQ option is used in annealed dynamics runs to set the interval between temperature resets. The UPDATE FREQ option allows the user to change the interval between updates of the nonbond and hydrogen bond interactions lists. The default value is 0.1 picoseconds. The HEAT EQUIL FREQ option allows the user to change the interval between temperature adjustments in dynamics runs. The default value is 0.1 picoseconds. The WRITE TRJ FREQ option allows the user to change the interval at which the current coordinates are written to the trajectory file. The default value is 0.1 picoseconds. The INFO LEVEL option allows the user to determine the amount of information that is written in the batch output file as dynamics is run. SAVE AVG COORD saves the structure's average coordinates during the course of the dynamics run, and the RESET DYNAMICS option allows the user to reset the dynamics time to zero.

User Input:

1. Toggle DYNAMICS VAR to bring up the DYNAMICS VAR (dynamics variables) menu.

TIME

The TIME option on each dynamics menu determines the duration of the dynamics run (in picoseconds). The number of steps of dynamics is equal to the duration divided by the time step. The TIME option is not present on the ANNEALED dynamics menu. The duration of the run is determined by the number of annealing cycles (see Section 5.8.4).

User Input:

1. Toggle TIME.
2. Enter the duration of the dynamics run in picoseconds (default is 0.5 picoseconds).

Dynamics var:
draw frequency
5
time step
0.001
anneal freq
0.100
update freq
0.100
heat equil freq
0.100
write trj freq
0.100
info level
1
save avg coord
reset dynamics
return

Figure 5.11:
DYNAMICS VAR
menu

USE VELOCITY

The USE VELOCITY option allows the user to assign the velocities for an extracted conformation to the next dynamics run, rather than having the velocities assigned in the usual manner (randomly). The EXTRACT option under ANIMATE is first used to specify the conformation whose velocities are to be saved. This is described in Section 5.14.5. The EXTRACT operation automatically creates a file for the extracted conformation (named CONF n where n is the sequence number). The energy expression must be set up for this file (see SETUP EEX, Section 5.2). USE VELOCITY is toggled on prior to running dynamics. USE VELOCITY applies to all types of dynamics except impulse.

User Input:

1. Use the EXTRACT option under ANIMATE to specify the conformation whose velocities are to be used (see Section 5.14.5).
2. Use SETUP EEX to set up the energy expression for the extracted file (see Section 5.2).
3. Toggle on the appropriate dynamics option under DYNAMICS to bring up the dynamics menu desired (see Sections 5.8.2 - 5.8.7).
4. Toggle on USE VELOCITY and set all other parameters to the desired values before executing the dynamics run.

RESET VELOCITY

The RESET VELOCITY option is only functional if the user has just run dynamics and has not left the Dynamics menu. This option allows the user to reset the velocities so that when the dynamics run is continued the velocities will be assigned randomly rather than using the last set of velocities as starting velocities. RESET VELOCITY applies to all types of dynamics except impulse dynamics.

User Input:

1. Toggle on RESET VELOCITY to have the velocities assigned randomly before continuing the dynamics run.

WRITE TRAJECT, WRITE VELOCITY

These toggles determine whether certain types of information about the dynamics run are written to disk. If the WRITE TRAJECT toggle is turned on, the program writes a binary trajectory file containing the atomic coordinates and energies of the structure at intervals determined by the value of WRITE TRJ FREQ on the DYNAMICS VAR menu. If the WRITE VELOCITY toggle is on, the atomic velocities are also written to the .trj file. Section 5.14 (ANIMATE) gives information about reading and displaying trajectory files.

User Input:

1. Toggle on WRITE TRAJECT to write a trajectory file during the dynamics run, or toggle it off to perform dynamics without saving the trajectory (default is off).
2. Toggle on WRITE VELOCITY to write the velocities in the trajectory file (default is off).

APPEND TRAJECT

The APPEND TRAJECT option allows the user to: 1) add information to the end of an existing trajectory file, and 2) restart dynamics calculations for a molecule using the last set of velocities as starting velocities rather than having the velocities assigned in the usual manner (randomly). This second feature requires that the existing trajectory file contain velocity information (WRITE VELOCITY must have been toggled on when the trajectory file was created). When the dynamics calculations are started, messages appear in the text window indicating the appending trajectory start time. If the trajectory file specified does not contain velocity information, the starting velocities are assigned randomly.

User Input:

1. Toggle on APPEND TRAJECT.
2. Enter the name of the trajectory file being appended.

EXECUTE

The EXECUTE function starts the dynamics run.

User Input:

1. Toggle EXECUTE to start the dynamics run.

5.8.2 ADIABATIC Dynamics

The ADIABATIC dynamics menu shown in Figure 5.12 allows the user to execute constant energy molecular dynamics. This menu has no special options; the options are common to all the dynamics menus.

Adiabatic(EVN) :
temperatur var
dynamics var
time
0.500
use velocity
reset velocity
write traject
write velocity
append traject
execute
return

Figure 5.12:
ADIABATIC
dynamics menu

5.8.3 QUENCHED Dynamics

The QUENCHED dynamics menu shown in Figure 5.13 runs dynamics with a number of steps of minimization after each period of molecular dynamics. The dynamics trajectory is not disrupted by this quenching process. Quenched dynamics is a way to search conformational phase space for low-energy structures. This menu has one special option: QUENCH STEP.

QUENCH STEP

The QUENCH STEP option allows the user to specify the number of steps of energy minimization that will be performed after each cycle of straight dynamics. The minimized coordinates will be written to the trajectory file.

User Input:

1. Pick QUENCH STEP.
2. Enter the number of steps of energy minimization after each dynamics period (default is 10).

Quenched:
temperatur var
dynamics var
quench step
10
time
0.500
use velocity
reset velocity
write traject
write velocity
append traject
execute
return

Figure 5.13:
QUENCHED
dynamics menu

5.8.4 ANNEALED Dynamics

The ANNEALED dynamics menu shown in Figure 5.14 allows the user to perform annealed dynamics runs. This menu does not use the TIME function. The length of the run is determined by the value of ANNEAL CYCLE on this menu, the value of ANNEAL FREQ on the DYNAMICS VAR menu, and the INITIAL TEMPER, FINAL TEMPER, and ANNEAL INCREMENT values on the TEMPERATUR VAR menu.

A single cycle of annealed dynamics begins at the initial temperature. After running dynamics for the time given by ANNEAL FREQ, the temperature is raised or lowered by ANNEAL INCREMENT, and another cycle of dynamics is run. This continues until the structure reaches the final temperature. Then the structure returns to the initial temperature by the same process. At this point one cycle of annealed dynamics is complete. This process will occur ANNEAL CYCLE times.

User Input:

1. Toggle ANNEAL CYCLE.
2. Enter the number of cycles of annealed dynamics to be performed (default is 1).

Annealed:
temperatur var
dynamics var
anneal cycle
1
use velocity
reset velocity
write traject
write velocity
append traject
execute
return

Figure 5.14:
ANNEALED
dynamics menu

5.8.5 IMPULSE Dynamics

The operations on the IMPULSE dynamics menu shown in Figure 5.15 allow the user to impart translational kinetic energy in a specified direction to a set of movable atoms and to follow the dynamics trajectory. The DIRECTION START and DIRECTION STOP options are used to define the direction vector, while TIME (default is 0.500 picoseconds) sets the duration of the dynamics trajectory. The translational kinetic energy imparted to the movable atoms is determined by TRANSLAT TEMPER. The velocity imparted to each atom along the direction vector is given by

$$v = \sqrt{(NkT) / M}$$

where N is the number of movable atoms and M is the mass of the movable atoms. If $M/N \approx 10$, which is typical for organic and biological compounds, then the atoms will translate about 0.01 Å in 1 femtosecond (default time step is 1 fs) at 1000° K.

TRANSLAT TEMPER

The TRANSLAT TEMPER option allows the user to set the temperature that defines translational kinetic energy parallel to the direction vector. The initial atomic velocities will be assigned to give the structure the correct amount of kinetic energy for the translational temperature.

User Input:

1. Pick TRANSLAT TEMPER.
2. Enter the temperature (default is 1000° K).

DIRECTION START

The DIRECTION START option allows the user to pick the starting atom of the direction vector for impulse dynamics or constrained minimization. Since only a direction vector is being defined, any pair of atoms specifying the appropriate direction can be used. They do not need to be movable atoms. The center of mass of the movable atoms will move in the direction specified from DIRECTION START toward DIRECTION STOP.

User Input:

1. Pick DIRECTION START.
2. Pick an atom to define the tail of the direction vector.

DIRECTION STOP

The DIRECTION STOP option allows the user to pick the stopping atom of the direction vector for impulse dynamics or constrained minimization. The center of mass of the movable atoms will move in the direction specified from DIRECTION START toward DIRECTION STOP.

Impulse:
temperatur var
dynamics var
translat temper
1000.000
time
0.500
direction start
direction stop
write traject
write velocity
append traject
execute
return

Figure 5.15:
IMPULSE
dynamics menu

User Input:

1. Pick DIRECTION STOP.
2. Pick an atom to define the head of the direction vector.

5.8.6 CANONICAL (TVN) Dynamics

The operations on the CANONICAL (TVN) menu shown in Figure 5.16 allow the user to run constant temperature/constant volume canonical dynamics. Special options are available for specifying the relaxation time constant, TAU(CAN), and choosing the algorithm which will be used (HOOVER or NOSE). See Appendix P for more details on canonical dynamics methods.

TAU (CAN)

The TAU (CAN) option allows the user to specify the relaxation time constant which will be used for canonical dynamics. The default value is 0.100 ps.

User Input:

1. Pick TAU (CAN).
2. Enter the relaxation time constant in picoseconds.

HOOVER/NOSE

This option allows the user to choose whether the Hoover or the Nose formalism will be used.

User Input:

1. Toggle the menu item until the desired method is displayed (HOOVER or NOSE).

Canonical(TVN) :
temperatur var
dynamics var
tau(can)=0.100
time
0.500
temp=300.000
use velocity
reset velocity
write traject
write velocity
append traject
Hoover
execute
return

Figure 5.16:
CANONICAL
(TVN) menu

5.8.7 TEMPDAMP (TVN) Dynamics

The operations on the TEMPDAMP (TVN) dynamics menu shown in Figures 5.17 allow the user to perform constant temperature/constant volume dynamics through weak coupling to a thermal bath via the Berendsen *et al* approach. The strength of coupling may be specified through a damping time constant, TAU(DMP). See Appendix P for more details.

TAU (DMP)

This option allows the user to control the strength of coupling used for constant volume isothermal dynamics by specifying the damping time constant, TAU (DMP). The default value is 0.100 ps.

User Input:

1. Pick TAU (DMP).
2. Enter the damping time constant in picoseconds.

Tempdamp (TVN) :
temperatur var
dynamics var
tau(dmp)=0.100
time
0.500
temp=300.000
use velocity
reset velocity
write traject
write velocity
append traject
execute
return

Figure 5.17:
TEMPDAMP
(TVN) menu

5.9 CONSTRAINED MIN

The operations on the CONSTRAINED MIN (constrained minimization) menu shown in Figure 5.18 allow the user to perform a series of constrained minimizations in which the center of mass of the movable atoms is moved incrementally in a specified direction. The DIRECTION START and DIRECTION STOP options are used to define the direction vector. Three types of constrained minimizations are allowed: PLANE (default), LINE and POINT. Toggling PLANE cycles through the options. The options correspond to constraining the center of mass to be (1) in the plane perpendicular to the direction vector (PLANE), (2) along a line parallel to the direction vector (LINE), and (3) fixed at a point (POINT). INCREMENT sets the distance the center of mass is moved between minimizations (default is 0.5 Å), and POINTS sets the number of increments of constrained minimization that are performed (default is 10). Each minimization is controlled by the convergence criterion, E CONVERGE; when the residual total force is below E CONVERGE (default is 0.001 kcal/mol/Å), the minimization stops. The maximum number of steps of minimization per increment is set with # OF STEPS (default is 100).

The results of the constrained minimizations can be saved in a trajectory or *trj* file using WRITE TRAJECT (Section 5.8.1) and displayed using ANIMATE (Section 5.14).

5.9.1 DIRECTION START

See Section 5.8.5.

5.9.2 DIRECTION STOP

See Section 5.8.5.

5.9.3 PLANE/ LINE/ POINT

Toggling PLANE/LINE/POINT cycles through the three types of constrained minimizations, namely, whether the center of mass of the movable atoms is constrained to (1) a plane perpendicular to the direction vector, (2) a line parallel to the direction vector, or (3) a point.

5.9.4 INCREMENT

The INCREMENT option allows the user to set the distance the center of mass is moved between sequential constrained minimizations.

Constrain min:
direction start
direction stop
plane
increment
0.500
update freq
5 0
points
0 / 1 0
e converge
0.001
of steps
100
execute
return

Figure 5.18:
CONSTRAINED
MIN menu

User Input:

1. Pick INCREMENT.
2. Enter the distance to be moved between minimizations (default is 0.5Å).

5.9.5 UPDATE FREQ

The UPDATE FREQ function allows the user to set the number of steps between updates of the nonbond and hydrogen bond interactions list when doing batch energy calculations. It is the same as the UPDATE FREQ function on the DYNAMICS VAR menu described in Section 5.12.4.

5.9.6 POINTS

The POINTS option allows the user to set the number of increments of constrained minimizations to be performed. Two numbers are displayed on the menu entry below POINTS (see Figure 5.18). The first number is a counter tracking the number of constrained minimizations which have been performed. The second number is the number of constrained minimizations to be performed.

User Input:

1. Pick POINTS.
2. Enter the number of increments of constrained minimizations to be performed (default is 10).

5.9.7 E CONVERGE

The E CONVERGE option allows the user to reset the convergence criterion that is used in minimizations. When the residual total force is below E CONVERGE, the minimization stops. The default is 0.001 kcal/mol/Å.

User Input:

1. Pick E CONVERGE.
2. Enter the new convergence criterion for residual total force (default is 0.001).

5.9.8 # OF STEPS

The # OF STEPS option allows the user to specify the maximum number of steps of constrained minimization that will be performed per increment. The number of points converged during the specified number of steps will vary depending on the convergence rate of the system being calculated. The calculation can be aborted using the abort key (F10 or F12).

Note: It is possible to do constrained minimization runs without doing any minimization. If # OF STEPS is set to 0, only the energy of each configuration is evaluated; no minimization is done.

User Input:

1. Pick # OF STEPS.

2. Enter the maximum number of steps of minimization per increment (default is 100).

5.9.9 EXECUTE

Toggling EXECUTE initiates the constrained minimizations. The user is prompted to indicate whether the results should be saved in a trajectory or *trj* file so that they can be displayed later using ANIMATE (see Section 5.14).

5.10 GRID SEARCH

The operations on the GRID SEARCH menu shown in Figure 5.19 allow the user to perform conformational searches on molecules by varying selected dihedral angles and minimizing the energy. The dihedral angles may be varied systematically within a specified range, or a Monte Carlo technique can be used to generate random dihedral angle values. Although the program allows up to one hundred dihedrals to be searched simultaneously, this number would not be computationally feasible. The conformations are ranked by energy and up to one hundred 'best' conformations can be saved and viewed sequentially using the BEST options (SET BEST, RESET BEST, BEST=). The conformations can also be displayed under ANIMATE (Section 5.14). The potential energy surface for two dihedral searches (*i.e.*, a ϕ - ψ plot) can be generated and displayed as a colored contour plot (READ PLOT).

The conformational search is set up using SETUP and can be examined before initiating execution by toggling EXAMINE. The # OF STEPS option sets the number of steps of energy minimization to be performed on each conformation generated by the dihedral angles. The CONSTRAINTS ON/OFF switch determines whether dihedral angle constraints are imposed on the selected dihedrals during energy minimization when doing a sequential search. When constraints are on, the selected dihedrals are forced to keep the angle specified for the given conformation throughout the minimization, while the rest of the molecule is allowed to relax. The DRAW ON/OFF switch determines whether the conformations generated are drawn on the screen during the search. The NBOND ON/OFF and H BOND ON/OFF options control whether the nonbond and hydrogen bond interaction lists are updated for each conformation. The distance cutoffs used for calculating the nonbond and hydrogen bond lists can be changed under the DEFAULTS menu.

The conformational energies are based on the energy expression set up under SETUP EEX (Section 5.2). The atoms which are to be moved in the conformational search must be included in the movable group.

Note: An M-CARLO SEARCH option is available which also provides Monte Carlo searches (see Section 5.11). The search used with M-CARLO SEARCH differs in that *all* the rotatable bonds in the molecule are varied instead of selected dihedrals. It also differs from the search described here in that torsions contained within rings can be searched and bump checking is used.

Grid search:
setup
ntorsions=0
examine
of steps
0
constraints on
draw on
nbond on
H bond on
defaults
execute
set best
reset best
best = 0
energy= 0.00
read plot
return

Figure 5.19:
GRID SEARCH
menu

5.10.1 SETUP

The **SETUP** option allows the user to select dihedral angles and to generate new conformations by varying these angles. Two methods are available: sequential and Monte Carlo. With the sequential method, the user specifies the total number of conformations to be generated and a range of angles to be explored; the program generates the conformations sequentially from the starting angle to the stopping angle. The Monte Carlo method generates random angles (0° to 360°) for all the dihedrals selected. This method is useful for getting a rough idea of the optimal angles. The sequential method could then be used to do a sequential search centered about the best angles as defined by the Monte Carlo method.

Selections:
sequential
Monte Carlo
return

Figure 5.20:
Search methods
menu

The number of dihedrals to be varied is indicated in the **NTORSIONS** box below **SETUP** on the **GRID SEARCH** menu.

User Input:

1. Toggle **SETUP**.
2. Enter the total number of conformations and energies to be saved (maximum is 100).
3. Pick the method to be used in generating the conformations (**SEQUENTIAL** or **MONTE CARLO**). Pick **RETURN** to complete the selection.
4. If **MONTE CARLO** was picked, enter the number of conformations to be generated.
5. Pick the first, second, third, and fourth atoms of the first dihedral. (A cross appears on the selected atoms; the crosses are successively colored red, yellow, blue, then green to indicate the order selected).
6. If using the **SEQUENTIAL** method:
 - a) Enter the starting angle for the search (the dihedral angle value of the current conformation is displayed in the text window).
 - b) Enter the stopping angle for the search.
 - c) Enter the number of divisions between the starting and stopping angles for the search. (The number of conformations calculated equals the number of divisions plus one.)
7. Repeat steps 5 and 6 for successive dihedrals (maximum total is 100). The number of dihedrals defined is indicated in the **NTORSIONS** box on the **GRID SEARCH** menu.
8. Pick any non-atom to terminate.

5.10.2 NTORSIONS

The **NTORSIONS** box displays the number of dihedrals to be varied. The user specifies this number when setting up the search with **SETUP** (see Section 5.10.1). The default is 0.

5.10.3 EXAMINE

The **EXAMINE** option allows the user to check the conformation search which has been set up. All of the conformations specified in the setup will be generated and displayed without calculating any energies.

User Input:

1. Pick the group(s) to be updated during conformation generation from the Groups menu. This step is skipped if only one group is in the system.

5.10.4 # OF STEPS

The # OF STEPS option allows the user to set the number of steps of energy minimization performed at each conformation generated by the dihedral search. This minimization process may modify the value of the dihedral angle being searched.

User Input:

1. Pick # OF STEPS.
2. Enter the number of steps of energy minimization (default is 0).

5.10.5 CONSTRAINTS ON/OFF

Toggling the CONSTRAINTS ON/OFF switch determines whether dihedral angle constraints are imposed on the selected dihedrals during energy minimization when doing a sequential search. If the status is ON, the selected dihedrals are forced to keep the angle specified for the given conformation throughout the minimization, while the rest of the molecule is allowed to relax. If the status is OFF, the whole molecule is allowed to relax during energy minimization. This option is only applicable when doing energy minimization (# OF STEPS is not set to zero) with a sequential search (see SETUP, Section 5.10.1).

5.10.6 DRAW ON/OFF

Toggling DRAW ON/OFF determines whether each of the conformations generated during the search are shown on the display. If the status is ON, then the display is updated as each conformation is generated. If the status is OFF, then the display is not updated. The default is DRAW ON.

5.10.7 NBOND ON/OFF

Toggling NBOND ON/OFF controls whether the nonbond list is updated for every new conformation. If the status is ON, then the list is recalculated for every conformation. If the status is OFF, then the list is calculated only for the first conformation.

5.10.8 H BOND ON/OFF

Toggling H BOND ON/OFF controls whether the hydrogen bond list is updated for every new conformation. If the status is ON, then the list is recalculated for every conformation and the best hydrogen positions are calculated for those donors which have rotational degrees of freedom. If the status is OFF, then the list is calculated only for the first conformation.

5.10.9 DEFAULTS

The DEFAULTS option is described in Section 5.12.

5.10.10 EXECUTE

Toggling EXECUTE initiates the conformational search. The conformations can be saved in a trajectory or *trj* file for later display and analysis under ANIMATE (Section 5.14). The potential energy surface for two dihedral searches can be saved for later examination using READ PLOT (Section 5.10.12).

User Input:

1. Pick EXECUTE.
2. Pick the groups to be updated during the calculation from the Groups menu. Pick RETURN to terminate the list. (This step is skipped if there is only one group in the system.)
3. Pick YES or NO to write a plot file.
4. Enter the full name of the file with the *plot* extension.
5. Pick YES or NO to write a trajectory or *trj* file.
6. Enter the full name of the file with the *trj* extension.

5.10.11 SET BEST, RESET BEST, BEST

The SET BEST, RESET BEST and BEST options allow the user to display the energetically best conformations generated during the conformational search. The energy (in kcal/mol) of each conformation is indicated on the SEARCH menu in the box beneath BEST.

SET BEST

The SET BEST option allows the user to display the *n*th best conformation.

User Input:

1. Pick SET BEST.
2. Enter the number of the conformation to be displayed.

BEST

Toggling BEST displays the next best conformation, *i.e.*, the one with the next lowest energy.

User Input:

1. Toggle BEST until the desired conformations have been displayed.

RESET BEST

Toggling RESET BEST resets the structure to its starting conformation.

5.10.12 READ PLOT

The READ PLOT option allows the user to read energy plots for two dihedral searches (e.g., ϕ - ψ plots). The plots are 2D contour displays of the two dihedral angle values and the total energy (Total E). The program also asks for the energy contour level and the number of contours to be displayed. This information is used to determine the color coding of the plot. The lowest energy contour is displayed in blue, while the highest energy contour is displayed in red. Intermediate energy contours are appropriately shaded from blue to red.

Please note that this option is only valid for two dihedral SEQUENTIAL searches (see Section 5.10.1 above).

User Input:

1. Pick READ PLOT.
2. Enter the name of the plot (*plot*) file.
3. Pick the group to contain the plot from the Groups menu.
4. Enter the name of the group to contain the plot.
5. Enter the contour level spacing in kcal/mol.
6. Enter the number of contours to be displayed.
7. If another plot is desired, repeat steps 2 - 6.

Because the plot is displayed at $z = 0$, it may be too dim. Use the z translate or depth cueing functions to improve the brightness of the plot.

5.11 M-CARLO SEARCH

The operations on the M-CARLO SEARCH menu shown in Figure 5.21 allow the user to generate random conformations of molecules by applying random rotations to a molecule's rotatable bonds. Torsions contained within rings may also be searched. This is done by selecting intra-ring bonds that can be temporarily disconnected during the search process (RINGS). A VDW SCALE option is available which sets the van der Waals bump check scale factor. Randomly generated conformations with any two atoms closer than their van der Waals distance multiplied by this value are rejected. If desired, energy minimizations may automatically be performed after the generation of each conformation. The number of minimization steps is set with # OF STEPS. Several on/off toggles are available. These determine whether sp²-sp² single bonds will be allowed to vary during the search (SP2-SP2), whether a new nonbond list and a new hydrogen bond list will be calculated before any energy calculations are performed (NBOND UPDATE, H BOND UPDATE), and whether each conformation generated will be repositioned so that its geometric center is coincident with that of the starting molecule (CENTER). WRITE TRAJECT determines whether a trajectory file is written to disk. RESET resets the lists containing the ring bonds, and EXECUTE initiates the search process.

Note: The M-CARLO SEARCH feature is limited to a single data file with all atoms movable. If there is more than one file in the system, use PROGRAM RESET on the UTILITIES menu to delete all files, then read in the file desired. Make sure the energy expression is set up with all atoms movable (see SETUP EEX, Section 5.2).

Note: A GRID SEARCH option is available under SIMULATE which provides both Monte Carlo and sequential searches (see Section 5.11). GRID SEARCH differs from the M-CARLO SEARCH method described here in that it doesn't vary all the rotatable bonds in the molecule, but requires the user to select the dihedrals which are to be varied. Also, torsions contained within rings cannot be searched and bump checking is not used.

M-Carlo search:	
defaults	
rings	
sp ² -sp ²	
center	
nbond update	
H bond update	
vdw scale	
0.300	
# of steps	
0	
write	traject
execute	
reset	
return	

Figure 5.21:
M-CARLO
SEARCH menu

5.11.1 DEFAULTS

The DEFAULTS option allows the user to modify the simulation parameters used in energy calculations, minimization and dynamics runs. One may also change the way nonbonding and hydrogen bonding interactions are computed and modify the temperature variables used in dynamics runs. It is identical to the DEFAULTS option described under SIMULATE (see Section 5.12).

5.11.2 RINGS

The RINGS option allows the user to indicate the ring bonds that should be disconnected during the search process. After selecting each bond, the program prompts the user to enter a tolerance distance that will be used to determine whether a particular conformation achieves "ring closure." Conformations with a bond distance greater than the value specified for the indicated ring bond are not accepted. The RESET option may be used to reset the ring bond lists so that they may be re-specified (see RESET below).

User Input:

1. Pick RINGS.
2. Select the ring bond to be broken (pick the two atoms defining the bond on the display).
3. Enter the tolerance for this bond (in Å).
4. Repeat steps 2 and 3 as required.
5. Pick any non-atom to terminate the operation.

5.11.3 SP2-SP2

The SP2-SP2 option is a switch that determines whether sp^2 - sp^2 single bonds will be allowed to vary during the search procedure (*e.g.*, the nitrogen-carbon bond of the peptide linkage). The default setting is off (sp^2 - sp^2 single bonds are *not* varied).

User Input:

1. Toggle SP2-SP2 on to allow rotation around sp^2 - sp^2 single bonds during the search; otherwise, toggle it off.

5.11.4 CENTER

The CENTER option is a switch that determines whether each randomly generated conformation will be translated so that its geometric center is coincident with that of the starting molecule. The default is on (centered).

User Input:

1. Toggle CENTER on to have each conformation's geometric center aligned with that of the starting molecule; otherwise, toggle CENTER off.

5.11.5 NBOND UPDATE

The NBOND UPDATE option is a switch that indicates whether a new nonbond list will be calculated for each randomly generated conformation before any energy calculations are performed. The default setting is on (a new nonbond list is calculated).

User Input:

1. Toggle NBOND UPDATE on to have a new nonbond list calculated; otherwise, toggle it off.

5.11.6 H BOND UPDATE

The H BOND UPDATE option is a switch that indicates whether a new hydrogen bond list will be calculated for each randomly generated conformation before any energy calculations are performed. In addition, if this switch is on, all H-bond hydrogens on rotatable groups (such as hydroxyls) will be positioned as to optimize hydrogen bonding in each conformation. The default setting is on (a new hydrogen bond list is calculated).

User Input:

1. Toggle H BOND UPDATE on to have a new hydrogen bond list calculated; otherwise, toggle it off.

5.11.7 VDW SCALE

The VDW SCALE option allows the user to change the van der Waals bump check scale factor. The default for this value is 0.300. That is, a randomly generated conformation that has any two atoms (that are not connected in a 1-2, 1-3, or 1-4 fashion) closer than 30% of their van der Waals distance is rejected.

User Input:

1. Pick VDW SCALE.
2. Enter the new scale factor.

5.11.8 # OF STEPS

The # OF STEPS option sets the number of energy minimization steps that will be performed after each conformation is generated. The default value is zero.

User Input:

1. Pick # OF STEPS.
2. Enter the number of minimization steps desired.

5.11.9 WRITE TRAJECT

If the WRITE TRAJECT toggle is turned on, a binary trajectory file containing the atomic coordinates and energies of the structures generated is written to disk.

5.11.10 EXECUTE

The EXECUTE option starts the search procedure. The user is prompted for the total number of random conformations to be generated. If the WRITE TRAJECT switch is on, a name for the trajectory file is also entered. Each conformation appears on the display as it is generated and its energy values are listed in the Simulation results menu. The conformation number currently being processed and the total requested are shown (*e.g.*, 16/65). The abort key can be used at any time to abort the search process.

User Input:

1. Toggle EXECUTE to initiate the search procedure.
2. Enter the number of conformations to be generated.
3. If WRITE TRAJECT is on, enter a name for the trajectory file. Include disk and directory specifications if needed (*trj* is used for the extension).

5.11.11 RESET

The RESET option is used to reset the ring bond lists. The RINGS option may then be used to respecify the ring bonds to be disconnected during the search process (see RINGS, above).

User Input:

1. Toggle RESET to reset the ring bond lists previously specified with the RINGS option.
2. Use the RINGS option to respecify the ring bonds to be disconnected during the search (see RINGS above).

5.12 DEFAULTS

The operations on the DEFAULTS menu shown in Figure 5.22 allow the user to specify the simulation parameters included in energy calculations, to set the variables used in minimization and dynamics runs, to change the temperature defaults used in dynamics runs, and to control the way nonbonding and hydrogen bonding interactions are computed. Options are also available which allow the user to update the nonbond and hydrogen bond lists and to set other miscellaneous variables.

The ENERGY VAR (energy variables) option is used to turn the basic components of the energy expression on and off, allowing energy calculations to include or omit bond, angle, torsion, inversion, hydrogen bonding, van der Waals, electrostatic interactions, and constraints. A switch is also available which determines whether or not charge equilibration is done each time the nonbond list is updated.

The MINIMIZE VAR (minimization variables) option allows the user to change the screen update frequency, minimization algorithm, convergence criteria, number of minimization steps, maximum displacement allowed, and amount of information output.

The TEMPERATUR VAR (temperature variables) option allows the user to set the temperature defaults for dynamics runs. The user can reset the initial and final temperatures of the run, change the annealing temperature increment and heating increment, and set the allowed temperature window for the temperature equilibration process.

The DYNAMICS VAR (dynamics variables) option is used to change the screen update frequency, dynamics time step, temperature adjustment frequency for annealing, frequency for updating the nonbond and hydrogen bond interaction lists, frequency of temperature adjustments, frequency for writing the atomic coordinates to a trajectory file, and the level of information written to the print file.

The NONBOND VAR (nonbond variables) option allows the user to specify the method used in determining the nonbond interactions to be included in the energy calculations, set the cutoff values for nonbonding interactions, and specify the combination rule used when calculating off diagonal van der Waals parameters (geometric or arithmetic mean). Options are also provided which allow the user to include or exclude electrostatic interactions, reset the dielectric constant, and turn the distance-dependent dielectric switch on or off. The user may also toggle the switches to include or exclude bond and/or angle interactions from the nonbonding interactions list and set the scaling factor for 1-4 nonbonding interactions.

The H BOND VAR (hydrogen bond variables) option allows the user to set the cutoff distances and angles for hydrogen bond energy calculations.

Defaults:
energy var
minimize var
temperatur var
dynamics var
nonbond var
H bond var
misc var
return

Figure 5.22:
DEFAULTS menu

The MISC VAR (miscellaneous variables) option allows the user to set other miscellaneous variables. These include a scale factor for sp^2 - sp^2 single bond torsions exocyclic to aromatic rings and a polyene switch for conjugated systems which allows the bond and angle force constants to depend on bond order. Also provided is a switch which determines how the angle-stretch terms are to be input, and an option which specifies how rigid minimizations are done.

The current parameter values are displayed on the menus of the different options. Defaults can be changed by toggling the desired parameter and entering the new value.

5.12.1 ENERGY VAR

The ENERGY VAR (energy variables) menu shown in Figure 5.23 allows the user to turn the basic components of the energy calculation on or off. When any of these switches is toggled on, its energy term is calculated and added to the total energy of the structure. The Q EQUIL/UPDATE switch is different in that it is not an energy parameter. When on, charge equilibration is done each time the nonbond list is updated (see Q EQUILIBRATE, under EDIT ID, CHARGES, Section 3.1.13). For the Dreiding force field, BONDS, ANGLES, INVERSIONS, TORSIONS, NONBONDS, COULOMB, and H BONDS are on by default, and the other switches are off.

BONDS	Bond stretches (default is on)
ANGLES	Angle bends (default is on)
INVERSIONS	Inversions defined by the force field (default is on)
USE ALL INVERS	All inversions, defined or not (default is on)
TORSIONS	Torsions defined by the force field (default is on)
USE ALL TORSION	All torsions, defined or not (default is on)
CONSTRAINTS	Harmonic constraints (default is off)
NONBONDS	van der Waals interactions (default is on)
COULOMB	Electrostatic interactions (default is on)
UP CHRGE EACH ITR	Charge equilibration done each time nonbond list is updated (default is off)
H BONDS	Hydrogen bonds (default is on for Dreiding, off for MM2 and Amber)
USER ENERGY	User supplied energy terms (default is off)

User Input:

1. Toggle the components of the energy expression on or off.

Energy var:
bonds
angles
inversions
use all invers
torsions
use all torsion
constraints
nonbonds
coulomb
up chrg each itr
H bonds
user energy
return

Figure 5.23:
ENERGY VAR
menu (Dreiding
force field)

User Input:

1. Toggle the components of the energy expression on or off.

5.12.2 MINIMIZE VAR

The operations on the MINIMIZE VAR (minimize variables) menu shown in Figure 5.24 allow the user to change the screen update frequency, minimization algorithm, convergence criteria, number of minimization steps, maximum displacement allowed, and the amount of information output.

DRAW FREQUENCY

The DRAW FREQUENCY option allows the user to change the frequency with which the structure is updated or redrawn on the display. The default is to update the structure after every 5 step of minimization or dynamics (DRAW FREQUENCY = 5). In some cases the time to redraw the structure can be significant compared to the time to calculate a step of minimization or dynamics, so that decreasing the frequency with which a structure is redrawn can speed a calculation considerably.

User Input:

1. Pick DRAW FREQUENCY.
2. Enter the new value for draw frequency (default is 5).

METHOD (ATOMS)

Toggling METHOD (ATOMS) changes the algorithm used for minimizations of atoms. [For periodic systems, the method used for minimizations of the unit cells may be specified independently using the METHOD (CELLS) option described below]. Five methods are available: two conjugate gradient algorithms named according to the version in which they became available (CONJ GRAD/V200 and CONJ GRAD/V145), two steepest descents algorithms similarly named (STP DSCNT/V200 and STP DSCNT/V145), and a Fletcher-Powell algorithm (FLCHR-PWL/V200). The default algorithm is CONJ GRAD/V200.

The conjugate gradient algorithm makes use of the previous minimization steps as well as the current gradient to determine the next step. This makes it converge very rapidly, comparing favorably with other algorithms in common use. In fact, the method converges to the minimum energy in the order of N steps where N is the number of degrees of freedom (see R. Fletcher and C. M. Reeves, *Comput. J.*, 7, 149 (1964) for more details).

Minimize var:
draw frequency
5
method (atoms)
conj grad/v200
rms force cvrg
0.100
of steps
50
maximum dist
1.500
info level
1
return

Figure 5.24:
MINIMIZE VAR
menu

In steepest descents, the coordinates at each step are displaced in a direction opposite to the gradient of the potential energy. The step size (see below) is increased if a lower energy is obtained and decreased if a higher energy is obtained. Although the convergence of steepest descents is generally poor because the displacements are small, this feature can be useful in minimizing large changes caused by a few bad contacts (see M. Levitt and S. Lifson, *J. Mol. Biol.*, **46**, 269 (1969) for more details).

Fletcher-Powell is a pseudo-second order method which generates a second order derivative matrix by finite differences. The computation time increases as the cube of the number of atoms. Therefore this method is generally not useful for systems larger than 300 atoms. For more information, see Press, William H., Flannery, Brian P., Teukolsky, Saul A. and Vetterling, William T. 1986, *Numerical Recipes, the Art of Scientific Computing* (Cambridge: Cambridge University Press), Chapter 10.

User Input:

1. Toggle METHOD (ATOMS) to change the minimization algorithm.

RMS FORCE CVRG

The RMS FORCE CVRG option allows the user to reset the criterion for convergence that is used in minimizations. If the residual total force is less than RMS FORCE CVRG, then the program will print a message in the output listing and stop the minimization. The default is 0.1 (kcal/mol)/Å.

User Input:

1. Toggle RMS FORCE CVRG.
2. Enter the new convergence criterion for residual total force (default is 0.1).

OF STEPS

The # OF STEPS option allows the user to change the number of steps performed in an energy minimization or simple molecular dynamics calculation.

User Input:

1. Pick # OF STEPS.
2. Enter the new value for # OF STEPS (default is 50).

MAXIMUM DIST

The MAXIMUM DIST option allows the user to enter a maximum displacement size. The motion of the atoms in the structure is limited so that MAXIMUM DIST is not exceeded. On the first step of conjugant gradient (or steepest descent) minimizations, the maximum change in coordinates (in 3N-dimensional coordinate space) is MAXIMUM DIST. A warning message also appears.

User Input:

1. Toggle MAXIMUM DIST.
2. Enter the new maximum distance (default is 1.5 Å).

INFO LEVEL

The INFO LEVEL option allows the user to control the level of information written to the print file, *[dir]proj.out*. If INFO LEVEL is greater than 1, then the energies are written at each step in the output file.

User Input:

1. Pick INFO LEVEL.
2. Enter the desired information level (default is 1).

5.12.3 TEMPERATUR VAR

The functions on the TEMPERATUR VAR (temperature variables) menu shown in Figure 5.25 allow the user to change the initial and final temperatures used in dynamics runs, to set the rate at which the temperature is raised or lowered when the initial and final temperatures are different, to set the rate of temperature change during annealed dynamics runs, and to reset the temperature window that determines how far the actual temperature can drift from the desired temperature before it must be readjusted.

INITIAL TEMPER

The INITIAL TEMPER option allows the user to set the initial temperature for a dynamics run. Assuming the starting configuration is an energy minimum, the initial velocities are assigned using a Maxwell-Boltzmann distribution:

$$\text{Temperature used in Maxwell-Boltzmann distribution} = 2 \times \text{INITIAL TEMPER}$$

User Input:

1. Pick INITIAL TEMPER.
2. Enter the initial temperature (default is 300° K).

FINAL TEMPER

The FINAL TEMPER option allows the user to set the final temperature for a dynamics run.

Temperatur var :
initial temper
300.000
final temper
300.000
heat increment
20.000
temper window
50.000
anneal increment
10.000
return

Figure 5.25:
TEMPERATUR
VAR menu

User Input:

1. Pick FINAL TEMPER.
2. Enter the final temperature (default is 300° K).

HEAT INCREMENT

The HEAT INCREMENT option is used to change the amount by which the temperature is changed when the structure is equilibrated. If the initial temperature is different from the final temperature, the temperature of the structure is raised or lowered by the value of HEAT INCREMENT at each equilibration cycle until the final temperature is reached.

User Input:

1. Pick HEAT INCREMENT.
2. Enter the increment by which to change the temperature.

TEMPER WINDOW

The temperature of the structure is checked at each equilibration cycle of a dynamics run (except during heating or cooling) to see if it is too far from the desired temperature. If the average temperature of the structure is outside the allowed range, the velocities of the atoms are rescaled to give the structure the desired temperature. The TEMPER WINDOW option is used to set the degree of difference allowed from the desired temperature. The structure's temperature must be TEMPER WINDOW higher or lower than the desired temperature before the velocities are rescaled.

User Input:

1. Pick TEMPER WINDOW.
2. Enter the value for the allowed temperature difference from the desired temperature.

ANNEAL INCREMENT

The ANNEAL INCREMENT option is used to reset the temperature increment for each step of annealing.

User Input:

1. Pick ANNEAL INCREMENT.
2. Enter the annealing temperature increment.

5.12.4 DYNAMICS VAR

The options on the DYNAMICS VAR (dynamics variables) menu shown in Figure 5.26 allow the user to modify the parameters used in dynamics calculations. The user can change the screen update frequency, dynamics time step, temperature adjustment frequency for annealing, frequency for updating the nonbond and hydrogen bond interaction lists, frequency of temperature equilibrations, and the frequency for writing the structural coordinates to a trajectory file. The level of information written during batch runs may also be changed, the structure's average coordinates may be saved, and the dynamics time may be reset to zero. The summed Verlet dynamics algorithm is used (see Appendix J).

There is a limitation in the program which requires that the frequency for writing coordinates to a trajectory file (WRITE TRJ FREQ) must be a multiple of the frequency of temperature equilibrations (HEAT EQUIL FREQ). This in turn must be a multiple of the update frequency (UPDATE FREQ) which must itself be a multiple of the anneal frequency (ANNEAL FREQ). Finally, the anneal frequency must also be a multiple of the time step (TIME STEP). If any of these five parameters is changed so that these conditions no longer hold, the inconsistent parameters are modified and a warning message is displayed indicating that these additional parameters have changed.

DRAW FREQUENCY

The DRAW FREQUENCY option is described in Section 5.12.2.

TIME STEP

The TIME STEP option allows the user to set the time step for the dynamics integration in picoseconds. **Warning:** The time step must be significantly smaller (about a factor of 10) than the fastest local motion for the numerical integration of Newton's equations of motions to be stable, *i.e.*, to generate a reliable molecular dynamics trajectory. Without hydrogens, the fastest motions (carbonyl, alkene, aromatic stretching vibrations at 1600-1700 cm^{-1}) have periods of 0.02 picoseconds. With hydrogens the fastest motions (hydrogen stretches) have periods of 0.01 picoseconds. Therefore, the time step should not be increased above 0.002 and for cases with hydrogens and/or elevated temperatures (say 1000° K) should be reduced to 0.001 or even 0.0005 picoseconds to insure stability.

User Input:

1. Pick TIME STEP.
2. Enter the time step in picoseconds (default is 0.001).

Dynamics var:
draw frequency
5
time step
0.001
anneal freq
0.100
update freq
0.100
heat equil freq
0.100
write trj freq
0.100
info level
1
save avg coord
reset dynamics
return

Figure 5.26:
DYNAMICS VAR
menu

ANNEAL FREQ

The ANNEAL FREQ option allows the user to set the time interval between annealing temperature adjustments. Dynamics will be run for the length of time given by ANNEAL FREQ. The number of steps of dynamics is determined by dividing ANNEAL FREQ by TIME STEP.

User Input:

1. Pick ANNEAL FREQ.
2. Enter the new value for the interval between temperature increments in annealed dynamics.

UPDATE FREQ

The UPDATE FREQ option allows the user to change the interval between updates of the nonbond and hydrogen bond interactions lists when doing batch energy calculations. The default value is 0.1 picoseconds.

User Input:

1. Pick UPDATE FREQ.
2. Enter the new value for the update frequency.

HEAT EQUIL FREQ

The HEAT EQUIL FREQ option allows the user to change the interval between heating/cooling adjustments or temperature equilibration cycles in dynamics runs. The default value is 0.1 picoseconds.

User Input:

1. Pick HEAT EQUIL FREQ.
2. Enter the new value for the equilibration frequency.

WRITE TRJ FREQ

The WRITE TRJ FREQ option allows the user to change the interval at which the current coordinates are written to the trajectory file. The default value is 0.1 picoseconds.

User Input:

1. Pick WRITE TRJ FREQ.
2. Enter the new value for the trajectory write frequency.

INFO LEVEL

The INFO LEVEL option allows the user to control the level of information written to the print file, [dir]proj.out. If INFO LEVEL is greater than 1, then the energies are written at each step in the output file.

User Input:

1. Pick INFO LEVEL.
2. Enter the desired information level (default is 1).

SAVE AVG COORD

SAVE AVG COORD option is an on/off toggle. When on, the structure's average coordinates during the course of the dynamics run are saved. Otherwise, they are not.

User Input:

1. Toggle SAVE AVG COORD on to save the structure's average coordinates.

RESET DYNAMICS

The RESET DYNAMICS option allows the user to reset the dynamics time to zero.

User Input:

1. Toggle RESET DYNAMICS to reset the dynamics time to zero.

5.12.5 NONBOND VAR

The operations on the NONBOND VAR (nonbond variables) menu shown in Figure 5.27 allow the user to specify the method to be used in determining the nonbond interactions to be included in the energy calculations. The program supplies two methods: SPLINE SWITCH and NO LIST. SPLINE SWITCH uses a spline switching function for cutting off the nonbond interactions and NO LIST includes all nonbond interactions but avoids calculating a nonbond list. The user may also write his own routine for calculating nonbond interactions. This routine will be used when the USER NONBOND switch is toggled on. For the SPLINE SWITCH method, nonbond cutoff distances may be specified (NBLST CUT, SPLIN ON, SPLIN OFF). A two-way toggle is also provided which allows the user to determine whether the geometric mean or arithmetic mean combination rule is to be used when calculating off diagonal van der Waals parameters. Other options are provided which allow the user to include or exclude electrostatic interactions, change the dielectric constant, or switch between a distance-dependent and an invariant dielectric constant. Switches are also available that exclude bond and angle interactions from the nonbonding interactions list. The SCALE 1-4 option allows the user to determine whether the energies of 1-4 nonbond interactions are multiplied by a scaling factor.

METHOD

The METHOD option is a 3-way toggle which allows the user to specify the method to be used in determining the nonbond interactions to be included in the energy calculations. The current method is displayed in the box beneath METHOD on the menu. The program supplies two methods: SPLINE SWITCH and NO LIST. SPLINE SWITCH uses a spline switching function for cutting off the nonbond interactions (see Appendix H). With the SPLINE SWITCH method, the nonbond cutoff distance (NBLST CUT) and the minimum and maximum distances used (SPLIN ON, SPLIN OFF) may be specified. NO LIST includes all nonbond interactions but avoids calculating a nonbond list. NO LIST is convenient for smaller molecules. EWALD uses the Ewald summations method and is for periodic systems only. The user may also write his own routine for calculating nonbond interactions. This routine must be properly linked to the program. It will appear as the third option under METHOD and will be used when the USER NONBOND switch is toggled on. Please consult the *Molecular Simulations Programmer's Manual* or contact Customer Support for further information on writing your own routines.

User Input:

1. Toggle METHOD to change the method used to determine the nonbond interactions included.

Nonbond var:
method
spline switch
user nonbond
geometric mean
Nonbond cut:
nblst cut=9.000
splin on =8.000
splin of =8.500
coulomb
epsilon=1.000
r*epsilon
Valence excl:
bond(1-2)
angle(1-3)
dihedral(1-4)
scale 1-4=1.000
return

Figure 5.27:
NONBOND VAR
menu

USER NONBOND

USER NONBOND is an on/off switch which determines whether a user-written method for calculating nonbond interactions is used. This routine must be properly linked to the program for proper execution. The name of the routine will appear as one of the options under METHOD and must also be selected here for proper execution. Please consult the *Molecular Simulations Programmer's Manual* or contact Customer Support for further information.

User Input:

1. Toggle USER NONBOND on to activate the user-written nonbond routine. Make sure that the name of the user method also appears under METHOD.

GEOMETRIC MEAN/ARITHMETIC MEAN

The GEOMETRIC MEAN/ARITHMETIC MEAN option is a two-way toggle which applies to the combination rules used in calculating off diagonal van der Waals parameters. When ARITHMETIC MEAN is displayed, the combination rule used is the arithmetic mean; when GEOMETRIC MEAN (the default) is displayed, the geometric mean combination rule is used (see Appendix G.8).

User Input:

1. Toggle the option to display the combination rule desired.

NBLST CUT

The NBLST CUT option allows the user to change the cutoff distance used for cutting off the nonbond (van der Waals and electrostatic) interactions which are included in an energy minimization or dynamics calculation. A list of possible nonbond interactions is constructed which includes all pairs of atoms within NBLST CUT of each other. The default NBLST CUT value is 9.0 Å.

When using the NO LIST method, the NBLST CUT option does not apply (the value changes to a very large number).

User Input:

1. Pick NBLST CUT.
2. Enter the nonbond list cutoff distance.

SPLIN ON, SPLIN OFF

The SPLIN ON and SPLIN OFF options allow the user to change the minimum (SPLIN ON) and maximum (SPLIN OFF) distances used by the nonbond switching function. The default values are 8.0 Å for SPLIN ON and 8.5 Å for SPLIN OFF. All nonbond interactions at distances less than SPLIN ON are included. Between SPLIN ON and SPLIN OFF, the calculated forces are

scaled smoothly to zero. The switching functions used by the program are presented in Appendix H.

When using the NO LIST or EWALD methods, the SPLIN ON and SPLIN OFF options do not apply (the values change to a very large number).

User Input:

1. Pick SPLIN OFF or SPLIN ON.
2. Enter the new nonbond switching distance.

COULOMB

If COULOMB is toggled on, electrostatic interactions are included in the energy calculations. Otherwise, they are not. It is the same as the COULOMB switch under ENERGY VAR (Section 5.12.1). The default is on.

User Input:

1. Toggle COULOMB on or off.

EPSILON

This option allows the user to reset the value of the dielectric constant. The default value for the dielectric constant is unity, *i.e.*, vacuum. A value of zero for EPSILON will cause the program to ignore electrostatic interactions.

User Input:

1. Toggle EPSILON.
2. Enter the new value for the dielectric constant.

R*EPSILON

If R*EPSILON is toggled on (default), then the dielectric constant is multiplied by the interatomic distance, R. If R*EPSILON is toggled off, the dielectric constant is independent of the interatomic distance.

BOND (1-2)

If BOND (1-2) is toggled on (the default), then atoms involved in bond interactions ("1-2" interactions) are excluded from the list of nonbonding interactions used in energy calculations. If BOND (1-2) is toggled off, atoms in bond and angle interactions are used in the nonbonding list.

ANGLE (1-3)

If ANGLE (1-3) is toggled on (the default), then atoms involved in angle interactions ("1-3" interactions) are excluded from the list of nonbonding interactions used in energy calculations. If ANGLE (1-3) is toggled off, atoms in angle interactions are used in the nonbonding list.

DIHEDRAL (1-4)

If DIHEDRAL (1-4) is toggled off (the default), then pairs of atoms that constitute the first and fourth atoms of torsion angles, *i.e.* atoms that are connected by two covalent bonds ("1-4" interactions), are included in the list of nonbonding interactions used in energy calculations. If DIHEDRAL (1-4) is toggled on, atoms that have 1-4 interactions are excluded from the nonbonding list. The default is to include atoms involved in 1-4 interactions in the nonbonding list.

SCALE 1-4

When this option is toggled, a message appears explaining what the option means and when it is used. The user is asked to enter a new scaling factor for 1-4 nonbond interactions. This option is not useful unless the AMBER force field is used for energy calculations. The default when other force fields are used is a scaling factor of 1.0, which means that 1-4 nonbond interactions (van der Waals and electrostatic interactions between atoms that are the first and fourth parts of a torsion angle, with three covalent bonds between them) are treated just like nonbond interactions between unconnected atoms. The default when the AMBER force field is used is to scale the energies and forces of 1-4 nonbond interactions by a factor of 0.5.

User Input:

1. Pick SCALE 1-4.
2. Enter the new scaling factor.

5.12.6 H BOND VAR

The operations on the H BOND VAR (hydrogen bond variables) menu shown in Figure 5.28 allow the user to change the cutoff distances and angles for hydrogen bond interactions.

HB DIST ON, HB DIST OFF, HB DIST CUT, HB ANGLE ON, HB ANGLE OFF, HB ANGLE CUT

The HB DIST ON, HB DIST OFF, HB DIST CUT, HB ANGLE ON, HB ANGLE OFF, and HB ANGLE CUT options allow the user to change the distance and angles used for cutting off the hydrogen bond interactions which are included in an energy minimization or dynamics calculation. The same approach described in the previous section to cut off nonbond interactions is used to cut off hydrogen bond interactions. The default distances for the hydrogen bond cutoffs are given below.

HB DIST ON	Default is 4.0 Å.
HB DIST OFF	Default is 4.5 Å.
HB DIST CUT	Default is 5.0 Å.

The default angles for the hydrogen bond cutoffs are given below.

HB ANGLE ON	Default is 65.0°.
HB ANGLE OFF	Default is 75.0°.
HB ANGLE CUT	Default is 90.0°.

User Input:

1. Pick HB DIST ON, HB DIST OFF, HB DIST CUT, HB ANGLE ON, HB ANGLE OFF or HB ANGLE CUT.
2. Enter the new hydrogen bond cutoff distance or angle.

H bond var:
Hb dist on
4.000
Hb dist off
4.500
Hb dist cut
5.000
Hb angle on
65.000
Hb angle off
75.000
Hb angle cut
90.000
return

Figure 5.28:
H BOND VAR
menu

5.12.7 MISC VAR

The MISC VAR (miscellaneous variables) option allows the user to change other miscellaneous energy variables. EXOCYC TORSION sets the scale factor for sp^2 - sp^2 single bond torsions exocyclic to aromatic rings, and the POLYENE switch allows the bond and angle force constants to depend on bond order. ANGLE-STRET 2 K determines how the angle-stretch terms are to be input. The user may also specify whether rigid minimizations are done using finite rotations through quaternions (RIGID/QUATERN) or small rotation operators [RIGID/EXP(ROT)]

EXOCYC TORSION

The EXOCYC TORSION option allows the user to set the scale factor for sp^2 - sp^2 single bond torsions exocyclic to aromatic rings. The default value is 1.000.

User Input:

1. Pick EXOCYC TORSION from the MISC VAR menu.
2. Enter the new scale factor.
3. Update the energy expression using the UPDATE EE toggle on the SIMULATE menu.

POLYENE

For conjugated molecules such as polyenes (*e.g.*, 1,3-butadiene), the four carbons are all sp^2 and would be considered as C_2 in Dreiding. However, the bond lengths are clearly quite different. In order to handle such cases, the program provides a POLYENE switch. When on, the bond and angle force constants depend on bond order. Thus, for 1,3 butadiene, the common bond is longer, and has a smaller force constant. A more detailed description is given in Appendix G.3.3.

User Input:

1. Toggle POLYENE on to have the bond and angle force constants depend on bond order. Otherwise, toggle it off.

ANGLE-STRET 2 K

ANGLE-STRET 2 K is an on/off toggle which determines how the angle-stretch terms are to be input. When on, the angle stretch terms must be input as K_{θ,R_1} and K_{θ,R_2} ; when off, they must be input as D and E (see Equations 14a and 14b in Appendix G.3.1). The default is off.

Misc var:
exocyc torsion
1.000
polyene
angle-stret 2 k
rigid/quatern
return

Figure 5.29:
MISC VAR menu

User Input:

1. Toggle ANGLE-STRET 2 K on or off.

RIGID/QUATERN, RIGID/EXP(ROT)

This option is a two-way toggle which allows the user to specify how rigid minimizations are done. Finite rotations through quaternions may be used (RIGID/QUATERN) or small rotation operators may be used [RIGID/EXP(ROT)]⁴. The recommended method is RIGID/QUATERN (the default).

User Input:

1. Toggle the option until the desired method is displayed, RIGID/QUATERN or RIGID/EXP(ROT).

⁴ H. Goldstein, *Classical Mechanics*, Addison-Wesley Publishing Co., Reading, Mass., 1980, p 156.

5.13 MODIFY PARAMS

The operations on the MODIFY PARAMS menu shown in Figure 5.30 allow the user to modify the force field parameters. PARAM INOUT may be used to read in parameters from an external file, input the parameters for a selected force field, and write out the current parameter list. MISSING PARAM allows the user to check for missing parameters. CHECK and MODIFY are used in conjunction with the remaining options to check and/or modify the current force field parameters. The current force field is also listed toward the bottom of the menu (default is Dreiding II). A more detailed discussion of parameter files and their formats is presented in Appendix I.

Modify params:
param inout
missing param
check
modify
Valence:
bond
angle
inversion
ang-ang invers
bond-x-angle
torsion
urey-bradley
Non valence:
new atom
van der Waals
offdiag nonbnd
charge
hydrogen bond
FF= DREIDII
return

Figure 5.30:
MODIFY
PARAMS menu

5.13.1 PARAM INOUT

The options on the PARAM INOUT menu shown in Figure 5.31 allow the user to read in parameters from an external file (READ PAR FILE) and write out the current parameter list.(WRITE PAR FILE). The input parameters may be modified if desired using MODIFY on the MODIFY PARAMS menu (Section 5.13.3). RESET ALL PARAM resets all the parameters to zero.

READ PAR FILE

The READ PAR FILE option allows the user to read a new set of parameters to be used with the force field and to add special parameters to the existing parameter set. After toggling READ PAR FILE the program asks for the name of the new parameter file. If the new parameters differ with any of the old parameters, the old parameters will be overwritten with the new ones. Any old parameters not contained in the new parameter set will be left unchanged.

User Input:

1. Pick READ PAR FILE.
2. Enter the name of the file containing new force field parameters.

WRITE PAR FILE

The WRITE PAR FILE option allows the user to write the current set of force field parameters to disk. After the user toggles WRITE PAR FILE the program asks whether the new parameter file should be in ASCII or binary format. Parameter files in ASCII format can be examined and modified by a user using a text editor. Files in binary format can not be read by users, but the computer reads them more quickly than files in ASCII format. Finally the program asks for the name of the new parameter file.

User Input:

1. Pick WRITE PARAMS.
2. Pick ASCII or BINARY.
3. Enter the name for the new force field parameter file.
4. Enter the remarks for the parameter file. Terminate by picking OK.

RESET ALL PARAM

Toggling RESET ALL PARAM sets the arrays containing force field parameters to zero.

Param Inout:
read par file
write par file
reset all param
return

Figure 5.31:
PARAM INOUT
menu

5.13.2 MISSING PARAM

Toggle MISSING PARAM to check for missing force field parameters which are needed by the current energy expression. If any force field parameters are missing, the program prompts the user to add them for particular bonds, angles, dihedral angles, *etc.*, in the structure. See Appendix G for an explanation of the force field parameters.

5.13.3 CHECK/MODIFY

CHECK is used to examine the current force field parameters, and MODIFY is used to modify the force field parameters for selected parts of the structure. The user picks CHECK or MODIFY, then picks the force field parameter to be checked or modified from those listed on the MODIFY PARAMS menu (six valence and five non valence parameters are listed, see Section 5.13.4 below).

5.14 ANIMATE

The operations on the ANIMATE menu shown in Figure 5.33 allow the user to step through or animate a sequence of conformations contained in a trajectory (or *trj*) file which has been generated from dynamics, constrained minimization, or conformational search calculations. Statistical analysis of the trajectory files may also be performed. An AVERAGE COORDS option is available which does coordinate averaging of all the snapshots read in; RMS COORDS generates RMS fluctuations for these coordinates and displays RMS plots on the screen. An interface has also been provided which allows the user to perform custom statistical analysis using their own routines (USER TRAJECT).

The name of the trajectory file must be entered first using the define trajectory option (DEFINE TRAJECT). The user may restrict the sequence of conformations to be included by specifying a starting and stopping sequence number or time value with the START and STOP options.

The conformers may be listed (LIST) in two modes: using their associated energy values to make comparisons or using the root mean square (RMS) values obtained from RMS matching. These modes appear on the LIST menu as BY ENERGY and BY RMS. Lists may be obtained for all conformers, the best one, those that are consecutively better, or for those that are within a specified energy or RMS difference value from the best. The user may also EXTRACT the information for a particular conformation (*e.g.*, the best) and create a file for it so that energy minimizations or other analyses may then be performed on it. A group will also be created for the extracted structure if the AUTO GROUP toggle is turned on.

The positions of the conformers must be read in (READ) before any animations or statistical analysis can be performed. As with the listings, the conformations may be read in BY ENERGY or BY RMS mode. When reading BY RMS, the conformers are matched to a reference conformation specified by the user. After moving the atoms to obtain the best possible fit, the RMS value for each of the conformers is calculated. The user may restrict the sequences read in to those that are consecutively better or to those that fall within a specified energy or RMS difference value.

The DISPLAY operations are used to step through or animate the conformations read in with the previous READ operation. The user may cycle through the sequences manually, or they may be scanned automatically. All the conformations may be displayed simultaneously (on top of each other) as well. This is helpful in illustrating the degree to which different parts of the molecule move or remain stationary. The conformations can be color-coded by sequence number, by energy, or by RMS value. The current sequence number and energy value is shown.

Statistical analysis may be performed on a number of trajectory snapshots. The AVERAGE COORDS option averages all the snapshots which have been read in; RMS COORDS generates

Animate:
define traject
start
- - -
stop
- - -
list
read
display
extract
auto group
average coords
RMS coords
user traject
return

Figure 5.33
ANIMATE menu

RMS fluctuations for these coordinates and displays RMS plots on the screen. These options operate on the sequences read in with the previous READ operation, allowing the user to restrict the snapshots included based on energy, RMS match, time and frequency.

Note: Conformers are animated as either single color or half-bonded vectors. The Titan allows animation of groups rendered in solid mode (CPK); the half-bond colors must be used, however (animation of single color solid renderings is not allowed).

Warning: Once the ANIMATE menu is exited, all the set up data is lost. The user must define the trajectory file and read in the conformations again before any animations or statistical analysis can be done.

5.14.1 DEFINE TRAJECT

The DEFINE TRAJECT option is used to define the trajectory file which is to be animated. If more than one file has been read into the system, the user will be prompted to pick the file for which the trajectory was written from the Files menu. The name of the trajectory (*trj*) file is then entered in the dialog box. When the trajectory file has been read in, the sequence number or time value for the first and last conformer in the file is displayed under the START and STOP options on the menu. None of the other options on the menu are operative until the trajectory file has been defined.

User Input:

1. Pick DEFINE TRAJECT.
2. Pick the file for which the trajectory was written from the Files menu. (This step is skipped if only one file is in the system.)
3. Enter the name of the trajectory (*trj*) file to be animated.

5.14.2 START, STOP

The START and STOP options are used to restrict the sequence of conformations to be included in the animation; they apply to the LIST, READ, EXTRACT, and DISPLAY operations described below. A starting and stopping sequence number or time value must be specified. If the trajectory file was created using dynamics calculations, then the starting and stopping conformations are specified by giving their time values (in picoseconds). If conformational search or constrained minimization calculations were used to generate the trajectory file, then the conformations' sequence numbers are used.

The values initially displayed under START and STOP on the menu are those for the first and last conformation in the file.

User Input:

1. Pick START.
2. Enter the sequence number or time value (in picoseconds) for the first conformation to be included in the animation.

3. Pick STOP.
4. Enter the sequence number or time value (in picoseconds) for the last conformation to be included in the animation.

5.14.3 LIST

The operations on the LIST menu shown in Figure 5.34 below allow the user to list the conformers in the trajectory file. The conformers may be listed in two modes: using their associated energy values to make comparisons (BY ENERGY) or using the root mean square (RMS) values obtained from RMS matching (BY RMS). In order to get listings BY RMS, the user must first read in the trajectory file in BY RMS mode (see READ in Section 5.14.4 below). These modes are described under BY ENERGY and BY RMS below.

The lists are shown in the text window and include the sequences defined by the START and STOP values. The conformers are listed with their associated energy values (if dynamics calculations were used to generate the trajectory, then the time associated with each of the conformers is also listed). When in BY RMS mode, the RMS values are also given. Lists may be obtained for all conformers (ALL), for those that are sequentially better than the best one so far (BETTER), or for the conformer with the lowest energy or RMS value (BEST). Comparisons are made using energy values when in BY ENERGY mode, while RMS values are used when in BY RMS mode. Those that are within a specified energy or RMS difference value from the best (WITHIN) may also be listed. These difference values are set using the ENERGY DIFF and RMS DIFFERENCE options.

List:
all
better
best
within
by energy
energy diff
10.000
by RMS
RMS difference
1.000
return

Figure 5.34: LIST menu

User Input:

1. Pick LIST from the ANIMATE menu.
2. Set the desired mode (BY ENERGY or BY RMS) on the LIST menu.
3. Pick one of the four listing options (ALL, BETTER, BEST, or WITHIN). If choosing WITHIN, first set ENERGY DIFF or RMS DIFFERENCE to the desired value.
4. Repeat steps 2 - 3 as often as desired.

ALL, BETTER, BEST, WITHIN

The ALL, BETTER, BEST and WITHIN options allow the user to list the conformers in the trajectory file in the text window. The conformers are listed with their associated time and energy values and are taken from the range of sequences defined by the START and STOP values. If in BY RMS mode, then the RMS difference values are also listed.

The ALL option lists all the conformers, BETTER lists those that are sequentially better than the best one so far, and BEST displays the conformer with the lowest energy or RMS value. Compar-

isons are made using energy values when in BY ENERGY mode, while RMS values are used when in BY RMS mode. The WITHIN option lists the conformers that are within the specified energy or RMS difference value from the best; when in BY ENERGY mode, the ENERGY DIFF value is used, while the RMS DIFFERENCE value is applied when in BY RMS mode.

User Input:

1. Pick the listing option desired (ALL, BETTER, BEST or WITHIN). If choosing WITHIN, first set ENERGY DIFF or RMS DIFFERENCE to the desired value.

BY ENERGY, BY RMS

BY ENERGY and BY RMS are mutually exclusive toggles; they set the mode to be used in the listings. When BY ENERGY is on, the energy values are used to compare the conformers and determine those listed with the BETTER, BEST and WITHIN options; when BY RMS is on, the root mean square values obtained from RMS matching are used. BY ENERGY is the default .

In RMS matching, the atoms in each of the conformers are moved to obtain the best possible fit to a reference conformation which is specified by the user. The RMS value for each of the conformers is then calculated by comparing the atom positions in each of the conformers to the same atoms in the structure which is being matched. This matching and calculation of RMS values is done when the trajectory file is read in. Therefore, in order to get BY RMS listings, the user must first do a READ using the BY RMS mode (see Section 5.14.4 below).

User Input:

1. Toggle BY ENERGY or BY RMS on.

ENERGY DIFF

ENERGY DIFF is used to set the energy difference value. Conformers with an energy value that is within this amount from the best conformer (the one with the lowest energy) will be included in the WITHIN listing. This is applicable only when in BY ENERGY mode. The default ENERGY DIFF value is 10.000 kcal/mol.

User Input:

1. Pick ENERGY DIFF.
2. Enter the energy difference value (in kcal/mol).

RMS DIFFERENCE

RMS DIFFERENCE is used to set the RMS difference value. Conformers with an RMS value that is within this amount from the best conformer (the one with the lowest RMS value) will be included in the WITHIN listing. This is applicable only when in BY RMS mode. The default RMS DIFFERENCE value is 1.0 Å.

User Input:

1. Pick RMS DIFFERENCE.
2. Enter the RMS difference value (in Å).

5.14.4 READ

The operations on the READ menu shown in Figure 5.35 are used to read the positions of the conformers in the trajectory file. A READ operation is required before any animations (DISPLAY operations) or statistical analysis can be performed. Similarly, READ using BY RMS mode is required for doing BY RMS listings.

As with the LIST operations, the trajectory file may be read in two modes: using the conformers' energy values to make comparisons (BY ENERGY) or using the RMS values obtained from RMS matching (BY RMS). These modes are described under BY ENERGY and BY RMS below.

The sequences to be read in are restricted to the range defined by the START and STOP values on the ANIMATE menu. The ALL option reads all the conformers specified by a read increment entered by the user. BETTER reads just those conformers which are better than the best one so far, and WITHIN reads those conformers that are within a specified energy or RMS difference value from the best. These difference values are set using the ENERGY DIFF and RMS DIFFERENCE options.

After picking READ, the user must specify a group to contain the trajectory file coordinates which will be read in. The READ menu will then appear and the user may select the READ options desired. When the coordinates have all been read in, the user will automatically be returned to the previous menu (the ANIMATE menu).

Note: The F12 key may be used to abort the READ operation. The conformers whose coordinates were read in prior to the abort are retained and may be displayed.

User Input:

1. Pick READ from the ANIMATE menu.
2. Pick the group which will contain the trajectory coordinates from the Groups menu. Pick RETURN to complete the selection.
3. Set the desired mode (BY ENERGY or BY RMS) on the READ menu.
4. If in BY RMS mode, pick the group containing the atoms to be matched (the reference conformation) from the Groups menu. Pick RETURN to complete the selection.
5. Pick one of the three read options (ALL, BETTER or WITHIN). If choosing WITHIN, first set ENERGY DIFF or RMS DIFFERENCE to the desired value.
6. If ALL was chosen, enter the read increment desired.

Read:
all
better
within
by energy
energy diff
10.000
by RMS
RMS diff
1.000
return

Figure 5.35:
READ menu

ALL, BETTER, WITHIN

The ALL, BETTER and WITHIN options read in the coordinates of the conformers in the trajectory file so that they may be animated. The conformers to be included are taken from the range of sequences defined by the START and STOP values. If in BY RMS mode, the conformers are first matched to the reference conformation and their RMS values are calculated.

The ALL option reads all the conformers specified by a read increment entered by the user (it will read every n th point). BETTER reads those that are sequentially better than the best one so far. Comparisons are made using energy values when in BY ENERGY mode, while RMS values are used when in BY RMS mode. The WITHIN option reads the conformers that are within the specified energy or RMS difference value from the best; when in BY ENERGY mode, the ENERGY DIFF value is used, while the RMS DIFFERENCE value is applied when in BY RMS mode.

User Input:

1. Pick the read option desired (ALL, BETTER or WITHIN). If choosing WITHIN, first set ENERGY DIFF or RMS DIFFERENCE to the desired value.
2. If ALL was chosen, enter the read increment desired.

BY ENERGY, BY RMS

BY ENERGY and BY RMS are mutually exclusive toggles; they set the mode to be used when reading the trajectory files. When BY ENERGY is on, the energy values are used to compare the conformers and determine those read in with the BETTER and WITHIN options. BY ENERGY is the default mode.

When BY RMS is on, RMS matching is done first. Each of the conformers is matched to a reference conformation which is specified by the user. The atoms in each of the conformers are moved to obtain the best possible fit. The RMS value for each of the conformers is then calculated by comparing the conformer's atom positions to the same atoms in the structure which is being matched. These RMS values are used to determine which conformers will be read in with the BETTER and WITHIN options.

User Input:

1. Toggle BY ENERGY or BY RMS on.
2. If in BY RMS mode, pick the group containing the atoms to be matched (the reference conformation) from the Groups menu. Pick RETURN to complete the selection.

ENERGY DIFF, RMS DIFFERENCE

These options function in the same way as they do in the LIST operations described above, except that they determine which conformers will be read in rather than listed when using the WITHIN option (see Section 5.14.3).

5.14.5 EXTRACT

The EXTRACT option is used to extract the information for a particular conformation and create a file for it. If desired, an auto group for the conformer may also be done by turning on the AUTO GROUP toggle (see Section 5.14.6). EXTRACT allows the user to use the ANIMATE operations to choose a particular conformation (*e.g.*, the best) and then save it so that he can then perform energy minimizations or other analyses on it.

If the trajectory file was created using dynamics calculations, the conformation to be extracted is specified by entering its time value. Otherwise, the conformation's sequence number is used. The time value or sequence number may be obtained using the LIST operations (see Section 5.14.3). A new file, named CONF n where n is the sequence number, will appear in the Files menu.

If the user chooses to do an auto group, he will be prompted to pick the group containing the original structure for which the trajectory was written (unless there is only one group in the system). The extracted conformation will then be displayed on the screen, and a new group, also named CONF n , will be created.

User Input:

1. Toggle AUTO GROUP on to create a group for the extracted conformer; toggle it off if this is not wanted.
2. Pick EXTRACT.
3. Enter the time value or conformation sequence number of the conformer to be extracted.
4. If AUTO GROUP is on, pick the group containing the original structure for which the trajectory was written from the Groups menu. Pick RETURN to complete the selection.

5.14.6 AUTO GROUP

The AUTO GROUP option is an on/off toggle which allows the user to automatically create a group for the conformation being saved in the EXTRACT operation described above (see Section 5.14.5). When the EXTRACT procedure is completed, the extracted conformation will be displayed on the screen and a new group named CONF n will be created. AUTO GROUP is on by default.

User Input:

1. Toggle AUTO GROUP on to create a group for the extracted conformer; toggle it off if this is not wanted.

5.14.7 DISPLAY

The operations on the DISPLAY menu shown in Figure 5.36 below are used to step through or animate the sequence of conformations read in with the previous READ operation. The display of the conformations may be controlled manually using ANIMATE, or they can be scanned automatically using AUTO SCAN. A SHOW ALL option is also provided which allows the user to display all of the sequences simultaneously (on top of each other). This is helpful in illustrating the degree to which different parts of the molecule move or remain stationary. CLEAR is used to remove them from the screen.

Conformers are animated as either single color or half-bond colored vectors. The Titan also allows animation of groups rendered in solid mode (CPK); the half-bond colors must be used, however (animation of single color solid renderings is not allowed).

Conformations displayed in a single color can be color-coded by sequence number (COLOR BY SEQNCE), by energy (COLOR BY ENERGY) or by RMS value (COLOR BY RMS). (To change a half-bonded group to a single color so that it can be color-coded, render the group with the half-bond option turned off (see RENDER GROUPS, Section 4.2). The COLOR INCREMENT option controls the size of the increments on the color wheel when color-coding by sequence number. The scale used in color-coding conformations by energy or by RMS can be changed using the ENERGY DIFF and RMS DIFFERENCE options. These options also set the scale used for the Y axis in the plot described below. A CAPTION option is provided at this level so that captions, labels *etc.* can be added to the displayed conformations. The PICK ORIGIN option is also provided so that a selected atom can be made the center of rotation.

The conformations being animated are renumbered from 1 to the total number read in. While cycling through the conformations, the current sequence number and its energy value are shown in the graphics window.

ANIMATE

The ANIMATE option allows the user to cycle through the conformations manually. This is done by rotating dial 8 (normally the depth cueing dial), or by depressing the middle mouse button and the control key and moving the mouse (move the mouse up or to the right to display later conformations, and down or to the left to display earlier conformations). The current sequence number and its energy value are displayed in the lower left corner of the graphics window.

User Input:

1. Pick ANIMATE.

Display:
animate
auto scan
show all
clear
color by seqnce
color increment
5.000
color by energy
energy diff
40.000
color by RMS
RMS diff
1.000
caption
pick origin
return

Figure 5.36:
DISPLAY menu

2. Pick anywhere to start the animation.
3. Rotate dial 8 (normally the depth cueing dial) to cycle through the conformations. If using the mouse, depress the middle mouse button and the control key and move the mouse up or to the right to display later conformations, and down or to the left to display earlier conformations).
4. Pick anywhere to stop the animation.

AUTO SCAN

AUTO SCAN does the same thing as ANIMATE described above, except that it scans through the conformations automatically. In this case, dial 8 is used to set the scan speed. Similarly, the middle mouse button plus the control key are used to set the scan speed for those using the mouse.

User Input:

1. Pick AUTO SCAN.
2. Pick anywhere to start scanning.
3. Rotate dial 8 (normally the depth cueing dial) to change the scan speed (rotate it clockwise to scan faster). If using the mouse, depress the middle mouse button and the control key and move the mouse up or to the right to scan faster, down or to the left to scan slower.
4. Pick anywhere to stop scanning.

SHOW ALL, CLEAR

The SHOW ALL option allows the user to display all of the conformations simultaneously (on top of each other). This is helpful in illustrating the degree to which different parts of the molecule move or remain stationary. The user may restrict the conformations displayed to a subset by specifying a first and last conformation. In specifying this subset, the user should refer to the conformation numbers which are shown in the upper right corner of the screen when doing ANIMATE or AUTO SCAN. The user is also prompted for a display increment which limits the conformations shown to every n th one.

CLEAR is used to remove the displayed structures from the screen.

User Input:

1. Pick SHOW ALL.
2. Enter the number of the first conformation to be displayed.
3. Enter the number of the last conformation to be displayed.
4. Enter the increment for the conformations to be displayed.
5. Pick CLEAR to remove the displayed structures from the screen.

COLOR BY SEQNCE, COLOR BY ENERGY, COLOR BY RMS

These options only apply to monochrome groups. They determine how the conformations are color-coded when animated: by sequence number (COLOR BY SEQNCE), by energy (COLOR BY ENERGY) or by RMS value (COLOR BY RMS). COLOR BY SEQNCE is the default.

With **COLOR BY SEQNCE**, red is the color of the first conformation; the color of each successive conformation is then changed by **COLOR INCREMENT** on the color wheel. Red is 120 and blue is 360 on the color wheel. For more information on the color wheel, see Section 4.3.11.

With **COLOR BY ENERGY** and **COLOR BY RMS**, red corresponds to the best conformation (lowest energy or RMS value), and blue corresponds to the conformations that are greater than or equal to **ENERGY DIFF** or **RMS DIFFERENCE** from the best. The conformations in between are interpolated between red and blue (**COLOR INCREMENT** is not used). **COLOR BY ENERGY** is only available if the trajectory file has been read in **BY ENERGY** mode. Similarly, **COLOR BY RMS** is only available if the file has been read in **BY RMS** mode.

When one of these options is picked, it is momentarily highlighted (until the program is finished setting up the colors). This option will then remain in effect until one of the other two options is chosen or until a **READ** operation is performed and the default (**COLOR BY SEQNCE**) again becomes effective.

User Input:

1. Set the associated parameter to the desired value (**COLOR INCREMENT** for **COLOR BY SEQNCE**, **ENERGY DIFF** for **COLOR BY ENERGY**, and **RMS DIFFERENCE** for **COLOR BY RMS**).
2. Pick **COLOR BY SEQNCE**, **COLOR BY ENERGY** or **COLOR BY RMS**.

COLOR INCREMENT

The **COLOR INCREMENT** option allows the user to set the increment on the color wheel which is used when color-coding conformations by sequence number (**COLOR BY SEQNCE**). For more information on the color wheel, see Section 4.3.11. The default value is 5.0.

User Input:

1. Pick **COLOR INCREMENT**.
2. Enter the increment on the color wheel to be used.

ENERGY DIFF

The **ENERGY DIFF** option allows the user to change the scale for color-coding conformations by energy (**COLOR BY ENERGY** option). The lowest energy conformation is assigned to red and all conformations with energies **ENERGY DIFF** or higher are assigned to blue. Conformations in between are interpolated between red and blue. The default **ENERGY DIFF** value is 10.000 kcal/mol.

User Input:

1. Pick **ENERGY DIFF**.
2. Enter the energy difference value desired (in kcal/mol).

RMS DIFFERENCE

The RMS DIFFERENCE option allows the user to change the scale for color-coding conformations by RMS value (COLOR BY RMS option). The conformation with the lowest RMS value is assigned to red and all conformations with RMS values RMS DIFFERENCE or higher are assigned to blue. Conformations in between are interpolated between red and blue. The default RMS DIFFERENCE value is 1.0 Å.

User Input:

1. Pick RMS DIFFERENCE.
2. Enter the RMS difference value desired (in Å).

5.14.8 AVERAGE COORDS

The AVERAGE COORDS option averages the coordinates for the sequences which have been read in with the READ operation (see Section 5.14.4). READ allows the user to restrict the sequences included based on energy difference, RMS match, time, and frequency. The averaged coordinates are held in a new file (AVE_XYZ). If desired, this file can be saved as a BioDesign General Format or Brookhaven file so that other operations may be performed on it (*i.e.*, it may be minimized, analyzed, displayed, *etc.*) (see WRITE, Section 2.2). In order to view the structure associated with the averaged coordinates, the user must first create a group using AVE_XYZ. This can be done using the MAKE GROUPS operations described in Section 4.1.

User Input:

1. Use the READ option to read in the sequences whose coordinates are to be averaged (see Section 5.14.4).
2. Pick AVERAGE COORDS.
3. To save the file containing the averaged coordinates, use the WRITE operations (see Section 2.2). To view the structure, use MAKE GROUPS and create a group from the AVE_XYZ file (see Section 4.1).

5.14.9 RMS COORDS

The RMS COORDS option calculates the root mean square of the coordinates for the trajectory sequences which have been read in with the READ operation. READ allows the user to restrict the sequences included based on energy difference, RMS match, time, and frequency (see Section 5.14.4). The RMS fluctuations generated provide a statistical indication of the motion of atoms during the trajectory. The x, y, and z components of the RMS fluctuations are held in a new file (RMS_XYZ). This file can be saved as a *bgf* file if needed for future processing (see WRITE, Section 2.2). After calculating the RMS values, a new menu appears containing four options which allow the user to display plots of the RMS fluctuation information on the screen: RMS/ATOM #, RMS/RES #, RMS MAIN/RES #, and RMS SIDE/RES # (see Figure 5.37). Toggling RMS/ATOM # displays a plot of RMS versus atom number. The remaining three options plot RMS averages versus residue number and apply only to proteins and peptides. RMS/RES # plots the RMS fluctuations averaged over all the atoms in each residue versus residue number. RMS MAIN/RES # plots RMS averages for the main chain atoms in each residue versus residue number and RMS SIDE/RES # does the same thing for the side chain atoms.

RMS plot type:
RMS/atom #
RMS/res #
RMS main/res #
RMS side/res #
return

Figure 5.37: RMS Plot type menu

User Input:

1. Use the READ option to read in the sequences to be analyzed (see Section 5.14.4).
2. Pick RMS COORDS.
3. To display a plot of the RMS fluctuation information, pick one of the options from the RMS Plot type menu which appears (RMS/ATOM #, RMS/RES #, RMS MAIN/RES #, or RMS SIDE/RES #).
4. To save the RMS coordinate file generated, use the WRITE operations (see Section 2.2).

5.14.10 USER TRAJECT

The USER TRAJECT option provides an interface which allows the user to perform custom statistical analysis on trajectory file data using their own routines. Toggling USER TRAJECT calls the user-written routine to accumulate statistics and/or data over the course of the trajectory. See the *Molecular Simulations Programmer's Manual* for a description of how to write and link a user-written routine with the program.

Chapter 6

Analyze Operations

The operations on the ANALYZE menu shown in Figure 6.1 allow the user to obtain atom identification information (IDENTIFY), to evaluate distances and angles between atoms and identify linkages (GEOMETRY), and to analyze the energies and forces on an atom, bond or angle (ENERGY-FORCE). For these three types of analysis, the user may do one operation at a time (for a single atom, bond or angle), or he may have the information for all the atoms in a group displayed at once. CONTACTS allows the user to check for "bad" van der Waals contacts. Both the energy and force analysis and the contacts checking operations require that an energy calculation has already been done for the structure. The ONE ENERGY option is therefore provided here for this purpose. It automatically sets up an energy expression with all the atoms being allowed to move and calculates the energy for the current structure.

The MATCH option uses least squares fits to match structures. TRAJECTORY allows the user to analyze and plot data in trajectory files and PLOT is used to display plots, change their appearance, and merge plot file data. PROTEIN check proteins for irregular structural features including D-amino acids, *cis*-peptides, and non-Ramachandran phi and psi angles. The dihedral angle values for disulfide linkages can also be determined. A VOLUMES option is also provided for measuring volumes and surface areas.

With all these types of analysis, the user may save the displayed information if desired.

Analyze:
identify
geometry
one energy
energy-force
contacts
match
trajectory
plot
protein
volumes
return

Figure 6.1:
ANALYZE menu

6.1 IDENTIFY

The operations on the IDENTIFY menu shown in Figure 6.2 allow the user to display atom identification information. Options are available to show eight types of information: atom type, atom charge, atom mass, atom label, atom radius, atom number, residue name, and chain. The user may choose to show the information for one atom at a time (SINGLE ATOM mode), or he may have the values for all the atoms in a group displayed at once (GROUP mode). He may have the displayed information cleared each time a new operation is performed (CLEAR) or he may have it saved using the NEW GROUP option. NEW GROUP allows the user to create a new group which can be used later for reviewing or for making hard copy.

The atom identification information is displayed next to the atom it refers to and is color-coded according to the option selected as shown:

Option	Color
ATOM TYPE	red
ATOM CHARGE	yellow
ATOM MASS	aqua
ATOM LABEL	green
ATOM RADIUS	magenta
ATOM NUMBER	orange
RESIDUE NAME	light blue
CHAIN	pink

Identify:
single atom
group
atom type
atom charge
atom mass
atom label
atom radius
atom number
residue name
chain
clear
new group
return

Figure 6.2:
IDENTIFY menu

The colors of the labels displayed with the analyze operations are set and cannot be changed (the CHANGE COLORS operations do not apply to groups created here using the NEW GROUP function).

6.1.1 SINGLE ATOM

SINGLE ATOM and GROUP are mutually exclusive toggles. The one which is highlighted indicates which of the two possible modes of operation of the identify function is to be used. Picking SINGLE ATOM mode allows the user to select individual atoms directly on the display. After choosing one of the eight atom identification options (ATOM TYPE, ATOM CHARGE, ATOM MASS, ATOM LABEL, ATOM RADIUS, ATOM NUMBER, RESIDUE NAME, CHAIN), the user picks an atom on the display. The information for that atom is then shown next to it in the appropriate color. This process can be repeated as often as desired.

If the user wishes to have the displayed information saved, he should first set up a new group using the NEW GROUP option as described in Section 6.1.12.

User Input:

1. Pick SINGLE ATOM.
2. Pick one of the eight atom identification options (ATOM TYPE, ATOM CHARGE, ATOM MASS, ATOM LABEL, ATOM RADIUS, ATOM NUMBER, RESIDUE NAME, CHAIN).
3. Pick the atom(s) desired on the display.
4. Pick a non-atom to terminate the operation.
5. Repeat steps 2-5 as often as desired.

6.1.2 GROUP

When in GROUP mode, atom identification information is displayed for all the atoms in the selected group(s). The user selects the groups to be analyzed from the Groups menu. He then chooses one of the eight atom identification options (ATOM TYPE, ATOM CHARGE, ATOM MASS, ATOM LABEL, ATOM RADIUS, ATOM NUMBER, RESIDUE NAME, CHAIN). The information is then displayed in the appropriate color next to each atom in the group.

If the user wishes to have the displayed information saved, he should first set up a new group using the NEW GROUP option as described in Section 6.1.12.

User Input:

1. Pick GROUP.
2. Pick the group(s) to be analyzed from the Groups menu. Groups created using the NEW GROUP operation are not allowed. Pick RETURN to terminate the list.
3. Pick one of the eight atom identification options (ATOM TYPE, ATOM CHARGE, ATOM MASS, ATOM LABEL, ATOM RADIUS, ATOM NUMBER, RESIDUE NAME, CHAIN).
4. Pick a non-atom to terminate the operation.
5. Repeat steps 2-4 as often as desired.

6.1.3 ATOM TYPE

The ATOM TYPE option allows the user to display the force field atom type(s) for a selected atom or group of atoms. They are shown in red, using a five character format as described in Appendix I.3.1.

6.1.4 ATOM CHARGE

The ATOM CHARGE option allows the user to display the magnitude of the point charge on a selected atom or on each of the atoms in a selected group. The atomic charge values are displayed in yellow.

6.1.5 ATOM MASS

The ATOM MASS option allows the user to display the atomic mass for a selected atom. This can also be shown for each of the atoms in a selected group. The values given are in atomic mass units (amu) and are displayed in aqua.

6.1.6 ATOM LABEL

The ATOM LABEL option allows the user to display the atom label(s) for a selected atom or group of atoms. The atom labels are displayed in green, using the same format used in the input file. (See Appendix D.3.)

6.1.7 ATOM RADIUS

The ATOM RADIUS option allows the user to display the van der Waals radius for a selected atom. This can also be shown for each of the atoms in a selected group. The values are in Å and are displayed in magenta.

6.1.8 ATOM NUMBER

Each atom known to the program is given a unique atom number. The ATOM NUMBER option allows the user to display this number for a selected atom or for each of the atoms in a selected group. Atoms input by reading a coordinate file are assigned atom numbers sequentially in the order in which they appear in the input file; when multiple input coordinate files are read, atoms read in the second and subsequent coordinate files are assigned higher atom numbers. The atom numbers are displayed in orange.

6.1.9 RESIDUE NAME

The RESIDUE NAME option allows the user to display the residue name(s) and number(s) for a selected atom or group of atoms. The residue names and numbers are displayed in light blue. They are separated by an underscore () and are the same as those designated in the input file. The residue name is displayed first; it is usually alphabetic and can be up to three characters long. For example, the names of the standard residues found in the Brookhaven Protein Data Bank files are listed in Appendix A.1.

6.1.10 CHAIN

The CHAIN option allows the user to display the chain designator(s) for a selected atom or group of atoms. This refers to the chains of peptides or nucleotides which are often contained in biological macromolecules. The chain designators are displayed in pink and are the same as those given in the input coordinate file. Chains are denoted by a one-letter (upper case) symbol in Brookhaven files. An underscore () is used if the chain identifier is blank.

6.1.11 CLEAR

CLEAR and NEW GROUP are mutually exclusive toggles; one or the other is always on. When CLEAR is on, the displayed atom identification information is not saved from one operation to the next, but is cleared from the display each time an operation is terminated.

6.1.12 NEW GROUP

The NEW GROUP operation allows the user to save the atom identification information to be displayed. This is done by creating a new group which can be reviewed later or can be used to make hard copy.

User Input:

1. Pick NEW GROUP.
2. Enter a name for the new group. (The name will appear in the Groups menu.)
3. Perform the identifying operation(s) desired. All labels generated by these operations are added to this group.
4. Pick CLEAR to terminate the items to be included in the group.

Once the new group is created, the atom identification information it contains can be shown or hidden at any time by toggling the group on or off. The colors of the labels saved in the new group are set, however, and cannot be changed (the CHANGE COLORS operations do not apply to groups created using the NEW GROUP function).

6.2 GEOMETRY

The operations on the GEOMETRY menu shown in Figure 6.3 allow the user to calculate distances and angles between atoms and to have these values displayed on the screen. Bond angles, dihedral and inversion angles, as well as the angle between two selected planes may be shown. The geometry information displayed is color-coded according to the type of measurement selected: DISTANCE (blue), ANGLE (pink), DIHEDRAL (chartreuse), INVERSION (orange), and PLANE (chartreuse).

A LINKAGE function is also provided which helps identify portions of the molecule which may affect its tertiary structure. These include disulfide bonds, nearby phenylalanine rings, and salt bridges. Angle or distance measurements are also displayed with these linkage options.

As in the IDENTIFY analysis, the user may choose to do one measurement at a time (SINGLE mode), or he may have the values for all the atoms in a group displayed at once (GROUP mode). (The PLANE operation is available in SINGLE mode only).

Any of these measurements may be done using a DIFFERENCE function. This allows the user to compare two similar structures and have the differences between the two displayed. If the user wishes to save any of the displayed geometry information, he may first choose the NEW GROUP option. This creates a new group which can then be used later for reviewing or for making hard copy. Similarly, if the user wishes to keep a listing of the measurement values along with their associated atom numbers, this data can be saved in an output file (WRITE FILE option). The GROUP ORIGIN and PICK ORIGIN options allow the user to change the center of rotation without having to leave the GEOMETRY menu.

6.2.1 SINGLE

SINGLE and GROUP are mutually exclusive toggles. The one which is highlighted indicates which of the two possible modes of operation of the geometry function is to be used. Picking SINGLE mode allows the user to select individual atoms directly on the display. The user chooses one of five measurement types: DISTANCE, ANGLE, DIHEDRAL, INVERSION, or PLANE (the LINKAGE operations are only available in GROUP mode). Picking an atom on the display will place a red dot on that atom. Atom identification information (*i.e.*, atom number, atom name, atom type, chain, residue name, and residue number) will also appear in the text window. When the proper number of atoms have been picked for the type of measurement selected, the calculated value is displayed and lines are drawn illustrating the distance or angle between the selected atoms. The order in which the atoms

Geometry:
single
group
distance
angle
dihedral
inversion
plane
linkage
difference
write file
clear
new group
group origin
pick origin
return

Figure 6.3:
GEOMETRY
menu

are picked is important for all of the angle measurements and is described in Sections 6.2.4 - 6.2.6 below. This process can be repeated as often as desired.

If the user wishes to compare two similar structures and display the differences between them, he first chooses the DIFFERENCE option and selects the two structures to be compared from the Groups menu. Then when one of the measurement types is picked, the value displayed will be the difference between the two structures. The DIFFERENCE option remains in effect until it is toggled off.

If the user wishes to have the displayed information saved, he should first set up a new group using the NEW GROUP option as described in Section 6.1.12. Similarly, if the user wishes to keep a listing of the measurement values along with their associated atom numbers, he should save this information in an output file by first choosing the WRITE FILE option. This is described in Section 6.2.10.

When in SINGLE mode, geometric information may be determined for any of the atoms on the screen; the atoms do not have to be connected by bonds.

User Input:

1. Pick SINGLE.
2. To display the differences between two structures:
 - a) Toggle DIFFERENCE on.
 - b) Pick the first (reference) structure for analysis from the Groups menu (pick RETURN to complete your selection).
 - c) Pick the second structure for analysis from the Groups menu (pick RETURN to complete your selection).Otherwise, make sure DIFFERENCE is toggled off.
3. Pick one of five measurement types: DISTANCE, ANGLE, DIHEDRAL, INVERSION or PLANE (the LINKAGE operations are available in GROUP mode only).
4. For INVERSION only: pick the method to be used to calculate the inversion angle (see Section 6.2.6).
5. Pick the desired atoms on the screen (2 for DISTANCE, 3 for ANGLE, 4 for DIHEDRAL and INVERSION, 6 for PLANE) in the proper order.
6. To continue making the same type of measurement, repeat step 5.
7. To make a different type of measurement, repeat steps 3 - 5.

6.2.2 GROUP

When in GROUP mode, geometric information is displayed for all the atoms in the selected group(s). The user selects the groups to be analyzed from the Groups menu. He then chooses one of five measurement types: DISTANCE, ANGLE, DIHEDRAL, INVERSION, or LINKAGE (the PLANE option is not available in GROUP mode). Values for all the connected atoms in the groups are then calculated and displayed on the screen. The numbers shown are color-coded according to measurement type and are positioned so that they correspond to the appropriate atoms on the screen. For the DIHEDRAL operation, a line is also drawn to indicate which atoms are in-

volved in the torsion angle value that is displayed. Similarly, illustrative lines are also drawn for the PHENYL RINGS and SALT BRIDGES options under LINKAGE.

If the user wishes to compare two similar structures and display the differences between them, he chooses the DIFFERENCE option and selects the two structures to be compared from the Groups menu. Then when one of the measurement types is picked, the value displayed will be the difference between the two structures. The DIFFERENCE option remains in effect until it is toggled off.

If the user wishes to have the displayed information saved, he should first set up a new group using the NEW GROUP option as described in Section 6.1.12. Similarly, if the user wishes to keep a listing of the measurement values along with their associated atom numbers, he should save this information in an output file by first choosing the WRITE FILE option. This is described in Section 6.2.10.

When in GROUP mode, only values for appropriately connected atoms are displayed.

User Input:

1. Pick GROUP.
2. Pick the group(s) to be analyzed from the Groups menu. Groups created using the NEW GROUP operation are not allowed. Pick RETURN to terminate the list.
3. To display the differences between two structures:
 - a) Toggle DIFFERENCE on.
 - b) Pick the first (reference) structure for analysis from the Groups menu. Pick RETURN to complete the selection.
 - c) Pick the second structure for analysis from the Groups menu. Pick RETURN to complete the selection.

Otherwise, make sure DIFFERENCE is toggled off.

4. Pick one of the five measurement types: DISTANCE, ANGLE, DIHEDRAL, INVERSION, or LINKAGE (the PLANE option is not available in GROUP mode). See Section 6.2.8 for a description of the LINKAGE operations.
5. For INVERSION only: pick the method to be used to calculate the inversion angle from the menu (see Section 6.2.6).
6. Pick anywhere on the screen to terminate the operation.
7. To make another type of measurement, repeat steps 4 - 6.

6.2.3 DISTANCE

The DISTANCE operation calculates the distance between two atoms. The value is given in Å and is displayed in blue at a point midway between the two atoms.

6.2.4 ANGLE

The ANGLE operation calculates the angle between the line segments connecting two atoms to a central atom. The value is displayed in pink near the central atom. When in SINGLE mode, the second atom selected is the central atom and the lines describing the angle are drawn in pink.

6.2.5 DIHEDRAL

Given four atoms (I, J, K, and L), the DIHEDRAL operation calculates the angle between plane IJK and plane JKL (see Appendix G.3). The bond between atoms J and K represents the common bond for the torsion. When in SINGLE mode, atoms I, J, K, and L are selected in that order, with the second and third ones selected representing the common bond atoms J and K. A positive angle is clockwise when looking from J toward K. This is shown in Appendix G, Figure G.5.

To indicate which atoms are involved in the torsion angle, a chartreuse line is drawn from the midpoint of line IJ to the midpoint of line KL and displays the torsion angle value θ halfway down this line as shown in Figure 6.4. This becomes particularly important when using GROUP mode since several torsion angle values are displayed at once.

6.2.6 INVERSION

The INVERSION operation allows three ways to calculate the inversion angle. These correspond to the common descriptions used in various force fields and are the Umbrella, AMBER, and CHARMM or Dreiding methods (see Appendix G.4). After picking INVERSION, the user chooses one of these methods from the INVERSION menu. With each of these methods, the inversion angle is defined by four atoms I, J, K and L. I is the central atom connected to the other three, and L is the special outer atom. When in SINGLE mode, the central atom I is selected first, followed by the special outer atom L. The third and fourth atoms may be selected in any sequence.

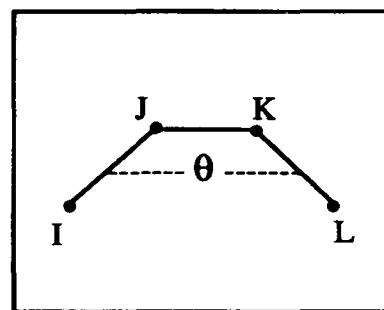


Figure 6.4: Display of Torsion Angle

The CHARMM and Dreiding methods define inversion in the same way: as if it were an improper torsion IJKL. The inversion angle is the angle between the IJK and JKL planes. The AMBER method also defines it as an improper torsion, but for JILK. In this case, the inversion angle is the angle between the IJL and IKL planes. The Umbrella method defines the inversion angle as the angle between the IL axis and the IJK plane. These inversion angles are illustrated in Appendix G.4, Figures G.6 - G.8.

The inversion angle value is displayed in orange at a position close to the central atom (I), 1/4 way between it and the special outer atom (L). If in SINGLE mode, orange lines are also drawn connecting the central atom to each of the other three atoms.

6.2.7 PLANE

The PLANE operation is used to display the angle between two selected planes. The user defines each plane by picking three atoms from the screen (a red cross marks the selected atoms). A triangle connecting the three atoms is drawn in chartreuse to represent each plane; the two planes are then joined by a line and the angle between them is displayed next to this line. The displayed val-

ues are between 0° and 90° . This is illustrated in Figure 6.5 below. The PLANE option is only available in SINGLE mode.

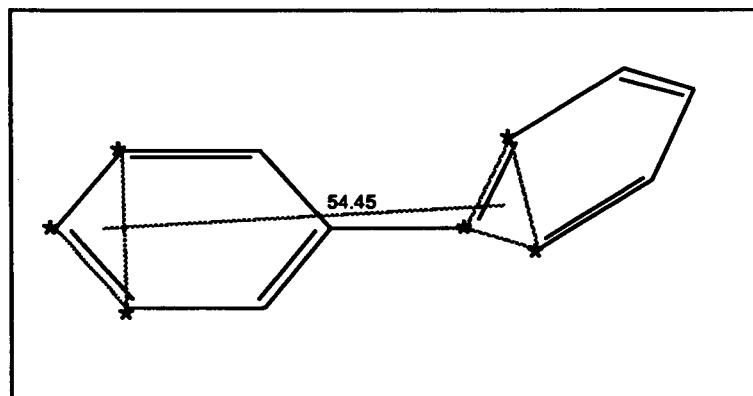


Figure 6.5: Display of Angle Between Two Selected Planes

6.2.8 LINKAGE

The operations on the LINKAGE menu shown in Figure 6.6 allow the user to identify portions of the molecule which may affect the molecule's tertiary structure. The linkages shown apply primarily to peptides. Options are available for identifying disulfide bonds and displaying their dihedral angles (DISULFIDE), for displaying the angle between the planes of two phenylalanine rings (PHENYL RINGS), and for identifying atoms that may form salt bridges (SALT BRIDGES). Default cutoff distances between the rings on phenylalanine residues and between the positively and negatively charged atoms on the salt bridges are set such that only those that are close enough to have an effect on the tertiary structure will be displayed. The user may reset these distances if desired, however, using the PHENYL CUTOFF and SALT BR CUTOFF options.

The user must be in GROUP mode before any linkage operations become available. The groups that are to be included in the analysis are selected from the Groups menu. LINKAGE may then be picked to bring up the LINKAGE menu. If the IDENTIFY option is toggled on, then when any of the three linkage operations is selected, additional information will be displayed in the graphics window or in the text window.

User Input:

1. Toggle GROUP mode on (if not already on).
2. Pick LINKAGE (must be in GROUP mode first).
3. Leave IDENTIFY toggled on to display additional information in the graphics window or in the text window; otherwise, toggle it off.
4. Perform the linkage operation desired (see the DISULFIDE, PHENYL RINGS and SALT BRIDGES options described below).

Linkage:
disulfide
phenyl rings
phenyl cutoff
6.0
salt bridges
salt br cutoff
3.5
identify
return

Figure 6.6:
LINKAGE menu

DISULFIDE

The DISULFIDE option displays the dihedral angle values for all the disulfide bonds in the groups selected. The atoms and bonds making up the dihedral angles are highlighted in yellow and the angle value is displayed nearby. If the IDENTIFY option was toggled on when DISULFIDE was selected, then atom identification information is also shown in the text window.

User Input:

1. Pick DISULFIDE.

PHENYL RINGS

The PHENYL RINGS option identifies phenylalanine residues whose rings are close enough to interact and possibly affect the tertiary structure of the molecule. A line connecting the nearby rings is drawn in green and the angle between their planes is displayed. Phenylalanine rings that are separated by more than a specified cutoff distance are excluded. The default cutoff value is 6.0 Å, but the user may reset this distance if desired using the PHENYL CUTOFF option described below. If the IDENTIFY option was toggled on when PHENYL RINGS was selected, then the residue names and numbers are also displayed.

If using the DIFFERENCE option to display the differences between two structures, then the PHENYL CUTOFF distance is applied to the reference structure (the first one selected) to determine which phenylalanine rings are close enough to be included in the display.

User Input:

1. Set the PHENYL CUTOFF distance to the desired value (default is 6.0 Å).
2. Pick PHENYL RINGS.

PHENYL CUTOFF

The PHENYL CUTOFF option is used to set the cutoff distance between nearby phenylalanine rings. When the PHENYL RINGS option is chosen, phenylalanine rings that are separated by more than this value will not be included in the display. The default cutoff distance is 6.0 Å.

User Input:

1. Pick PHENYL CUTOFF.
2. Enter the cutoff distance in Å.

SALT BRIDGES

The SALT BRIDGES option identifies positively and negatively charged atoms that are close enough together that they may form salt bridges and thus affect the tertiary structure of the molecule. A line connecting the two oppositely charged atoms is drawn in magenta and the dis-

tance between them is displayed. Charged atoms that are separated by more than a specified cutoff distance are excluded. The default cutoff value is 3.5 Å, but the user may reset this distance if desired using the SALT BR CUTOFF option described below. If the IDENTIFY option was toggled on when SALT BRIDGES was selected, then atom identification information is also shown in the text window.

If using the DIFFERENCE option to display the differences between two structures, then the SALT BR CUTOFF distance is applied to the reference structure (the first one selected) when identifying the charged atoms to be displayed.

User Input:

1. Set the SALT BR CUTOFF distance to the desired value (default is 3.5 Å).
2. Pick SALT BRIDGES.

SALT BR CUTOFF

The SALT BR CUTOFF option is used to set the cutoff distance between nearby oppositely charged atoms. When the SALT BRIDGES option is chosen, oppositely charged atoms that are separated by more than this value will not be included in the display. The default cutoff distance is 3.5 Å.

User Input:

1. Pick SALT BR CUTOFF.
2. Enter the cutoff distance in Å.

IDENTIFY

If the IDENTIFY option is toggled on, then when any of the three linkage operations is selected, additional information will be displayed in the graphics window or in the text window. The residue name and number are shown for the PHENYL RINGS option, and atom identification information is displayed on the text display for the DISULFIDE and SALT BRIDGES options. The default is on.

User Input:

1. Toggle IDENTIFY on prior to selecting one of the three linkage options (DISULFIDE, PHENYL RINGS, or SALT BRIDGES) in order to display additional information.

6.2.9 DIFFERENCE

The DIFFERENCE operation allows the user to compare two similar structures and have the differences between the two displayed. Any of the geometry measurements may be done using the DIFFERENCE function. Thus, the differences between the two structures in bond lengths, angles, torsions, or inversions may be shown. Differences in the angle between two selected planes or in the angle or distance of any linkages may also be displayed. The two structures must have

the same number of atoms and identical connectivities, but the individual atoms may be different or they may have different conformations. For example, one may want to look at the differences between two conformers of the same molecule, or to observe the effect of replacing one atom with another atom in two otherwise identical structures.

Differences may be displayed for a single measurement (SINGLE mode) or for all the atoms in a selected group (GROUP mode). Once in the proper mode, the user chooses the DIFFERENCE option and selects the two structures to be compared from the Groups menu. Then when one of the measurement types is picked (DISTANCE, ANGLE, DIHEDRAL, INVERSION, PLANE, or LINKAGE), the values displayed will be the differences between the two structures. (The LINKAGE option provides three measurement types; these are described in Section 6.2.8 above.) The numbers shown are color-coded according to measurement type and are positioned so that they correspond to the appropriate atoms on the screen. Illustrative lines are also drawn in some cases. The DIFFERENCE option remains in effect until it is toggled off.

The use of the DIFFERENCE option is described under SINGLE and GROUP in Sections 6.2.1 and 6.2.2 above.

6.2.10 WRITE FILE

The WRITE FILE option allows the user to save the geometry data: a listing of the measurement values along with their associated atom numbers is kept in an output file. Using the appropriate system commands, this file may then be viewed or printed.

The geometry measurements are listed in the same order in which they were executed. For each measurement made, the atom numbers are shown on the same line with the measurement value. The atom numbers given are those assigned by the program and are relative to the beginning of the first internal file in the system (File A). These atom numbers may be correlated with other atom identification information (atom label, residue name, charge etc.), by printing out a listing of the coordinate file data for the structure (the *bgf* file). Difference measurements are indicated as such; in this case, the atom numbers given are those of the reference structure. An example output file listing is shown in Figure 6.7.

User Input:

1. Toggle WRITE FILE on.
2. Enter a name for the output file. Include appropriate disk and directory specifications if needed (see Section 1.6).
3. Perform the geometry operations desired.
4. Use the appropriate system commands to display the saved data or to print it.

6.2.11 CLEAR, NEW GROUP

The CLEAR and NEW GROUP options function in the same way as they do in the atom identification operations described above (Sections 6.1.11- 6.1.12), except that in this case, the operations being performed are geometric measurements, and the displayed information that is being saved or cleared is geometry information.

BOND DISTANCE:	I	J	DISTANCE		
	28	42	2.770		
	45	49	2.384		
BOND ANGLE: (difference)	I	J	K	ANGLE	
	43	45	48	3.876	
BOND TORSION:	I	J	K	L	TORSION
	42	39	40	43	-50.575
BOND INVERSION:	I	J	K	L	INVERSION
	43	44	48	45	85.873
FIRST PLANE			SECOND PLANE		
59	57	58	36	33	32
					85.096
DISULFIDE BRIDGE:	C1	S1	S2	C2	TORSION
	44	45	121	122	72.184
SALT BRIDGE:	+ve atom	-ve atom	DISTANCE		
	62	84	1.805		
PHENYLALANINE PLANE ANGLES:	RESIDUE 1		RESIDUE 2		ANGLE
	2		12		179.905

Figure 6.7: Example of an output file listing obtained with WRITE FILE option

6.2.12 GROUP ORIGIN

The GROUP ORIGIN operation sets the center of rotation to the geometric center of a set of groups selected from the Groups menu. Groups created under ANALYZE (using the NEW GROUP operation from the IDENTIFY, GEOMETRY, ENERGY-FORCE or CONTACTS menus) are not allowed.

User Input:

1. Pick GROUP ORIGIN.

2. Pick the group(s) to be used for centering rotations from the Groups menu.
3. Pick RETURN to terminate the list.

6.2.13 PICK ORIGIN

The PICK ORIGIN operation sets the center of rotation to an atom picked from the display.

User Input:

1. Pick PICK ORIGIN.
2. Pick the atom desired for the rotation center.

6.3 ONE ENERGY

The ONE ENERGY option calculates the energy of the current conformation and displays the results in the Simulation Results menu. It automatically sets up the energy expression so that all the atoms are allowed to move (none are fixed). The energy results are displayed in the menu and when the energy calculations are finished, a message appears in the message box. ONE ENERGY is provided at this location as a shortcut since both the energy and force analysis and the contact checking operations require that an energy calculation has already been done for the structure.

User Input:

1. Pick ONE ENERGY.
2. When finished, continue by picking another option on the ANALYZE menu.

6.4 ENERGY-FORCE

The operations on the ENERGY-FORCE (Energy and Force) menu shown in Figure 6.8 allow the user to analyze the energies and forces on individual atoms, bonds and angles. Energy calculations must have been done on the structure before any of these analyses can be performed. As in the IDENTIFY and GEOMETRY analyses, the user may choose to do one measurement at a time, selecting the atoms directly from the display, or he may have the values for all the atoms in a group displayed at once.

Three options do a single measurement at a time (SINGLE ENERGY, SNGL ENERGY/ATOM and SNGL FORCE/ATOM); there are also three options which perform the same operations but on groups of atoms (GROUP ENERGY, GRP ENERGY/ATOM and GRP FORCE/ATOM). When selecting atoms directly from the screen, atom identification information is provided on the text window.

The energies and forces that may be displayed are divided into three types: internal, nonbond, and constraints. These are displayed in the Energy terms menu (see Figure 6.9). The internal or bonded interactions consist of bond stretching, bond angle bending, dihedral angle torsion and inversion terms and are represented by the first four items in the menu (BOND, ANGLE, DIHEDRAL, and INVERSION). The nonbonded interactions are listed next and consist of van der Waals, electrostatic and hydrogen bond terms (VDW, ELECTROSTATIC, and H BOND). Another nonbond option (NONBOND) is also provided which sums the van der Waals and electrostatic energy terms. If any constraints had been imposed when the energy calculation for the structure was done, the energies and forces associated with these may also be evaluated using the corresponding constraint terms (DIST CONSTRNT, ANGLE CONSTRNT, DIHED CONSTRNT, INVER CONSTRNT).

In order to display the energy values for any single energy term, the user selects SINGLE ENERGY or GROUP ENERGY. He then specifies which energy term he wants by choosing from the Energy Terms menu. The user may also select any number (or ALL) of the energy terms and have their combined values displayed. This is done using the SNGL ENERGY/ATOM or GRP ENERGY/ATOM options. The SNGL FORCE/ATOM and GRP FORCE/ATOM options allow

Energy-force:	
single energy	
group energy	
sngl energy/atom	
grp energy/atom	
sngl force/atom	
grp force/atom	
defaults	
group orign	
pick orign	
clear	
new group	
return	

Figure 6.8:
ENERGY-FORCE
menu

Energy terms:	
bond	
angle	
dihedral	
inversion	
nonbond	
vdw	
electrostatic	
H bond	
dist constrnt	
angle constrnt	
dihed constrnt	
inver constrnt	
all	
return	

Figure 6.9: Energy
terms menu

the user to display the net force on individual atoms. In addition to a force value, a force vector showing the magnitude and direction of the force is also displayed.

After selecting the atoms or groups to be analyzed, color-coded values will then be displayed on the screen. These may include energy, force, and distance or angle measurements, depending on the energy term selected. Illustrative lines are also drawn in some cases. The distances given are in Å and the energy values are in kcal/mol. The units for force depend on the energy term and are as follows: the force vectors and bond stretches are in kcal/(mol·Å), and the bond angle bending, dihedral angle torsion and inversion terms are in kcal/(mol·rad·Å).

The DEFAULTS option brings up another menu which allows the user to modify the way in which the energy and force analysis is done. The simulation parameters used in energy calculations may be changed. Options are provided which allow the user to change the distance and energy cutoff values and thus affect which atoms are included in nonbond interactions. Toggles are also available for turning the display of geometry, energy, and force values on or off and for setting the length of the force vectors.

The GROUP ORIGIN and PICK ORIGIN options allow the user to change the center of rotation without having to leave the ENERGY-FORCE menu. As in the other three types of analysis, CLEAR and NEW GROUP are mutually exclusive toggles. When NEW GROUP is on, the energy and force values displayed will be saved; when CLEAR is on, these values will be cleared from the display at the end of each operation.

6.4.1 SINGLE ENERGY

The SINGLE ENERGY operation allows the user to evaluate the energies and forces on individual atoms, bonds, and angles selected directly from the display. Any of twelve different energy terms may be analyzed (one at a time). These are selected from the Energy Terms menu.

Internal Energy Terms: When choosing one of the internal energy terms (BOND, ANGLE, DIHEDRAL, or INVERSION), the bond or angle is specified by picking atoms on the screen. The order in which the atoms are selected is important for all the angle terms (see Sections 6.2.4 - 6.2.6). After the appropriate number of atoms have been selected, three color-coded numbers will be displayed, representing the distance or angle, the energy, and the force. Lines are also drawn in light blue to highlight the bond or angle being evaluated.

Nonbond Interactions Terms: When evaluating van der Waals and/or electrostatic energies with the VDW, ELECTROSTATIC, or NONBOND terms, an atom is picked directly from the screen. Lines are then drawn from the selected atom to show all the nonbond interactions that are within a specified cutoff distance and that have van der Waals or electrostatic energies greater than or equal to a specified value. The default cutoff distance is 10 Å and the van der Waals and electrostatic energy cutoffs are -9999.999 kcal/mol. One of these nonbond interactions is then selected for detailed examination by picking the second atom. The energy and distance values for the interaction are then displayed. For the ELECTROSTATIC option, the electrostatic charge on each of the atoms is also shown. The user may specify the distance and energy cutoffs and whether the non-

bond interactions displayed include 1-4 nonbond interactions by using the DEFAULTS option described in Section 6.4.7.

For the HBOND option, the atoms selected must be a hydrogen bond donor/acceptor pair. The donor and acceptor may be selected in any order. In this case, the hydrogen bond force is also shown in addition to the distance and energy values.

Constraint Terms: If any constraints had been imposed when the energy calculation for the structure was done, the energies and forces associated with these may be evaluated using the corresponding constraint term for distance (DIST CONSTRNT), angle (ANGLE CONSTRNT), torsion (DIHED CONSTRNT), and inversion (INVER CONSTRNT). After selecting one of these constraint terms from the menu, the user picks the set of atoms that was initially used to define the constraint from the screen. The order in which the atoms are selected is important for the angle, torsion and inversion terms (see Sections 6.2.4 - 6.2.6). A line is then drawn and three color-coded values are displayed indicating the constrained distance or angle, the energy, and the force.

User Input:

1. Pick SINGLE ENERGY.
2. Pick one of the twelve energy terms from the Energy Terms menu.
3. Internal Energy Terms: Pick the appropriate number of atoms on the screen (2 for BOND, 3 for ANGLE, 4 for DIHEDRAL and INVERSION) in the proper order.
 - Nonbond Energy Terms: Pick the atom to be analyzed on the display.
Pick the second atom from the nonbonded atoms shown to get the energy and distance values for the interaction.
 - Constraint Terms: Pick the set of atoms that was initially used to define the constraint from the display (2 for distance, 3 for angle, 4 for torsion or inversion).
4. To continue making the same type of measurement, repeat step 3 as often as desired.
5. Pick a non-atom to terminate the operation.
6. Repeat steps 2-5 as often as desired.

6.4.2 GROUP ENERGY

The GROUP ENERGY option provides the same function as SINGLE ENERGY described above, except that instead of doing one evaluation at a time, values for a whole group of atoms are displayed at once. Instead of selecting individual atoms from the display, the user selects groups of atoms from the Groups menu. The only other differences are as follows: SINGLE ENERGY displays the distances between the selected atoms for the van der Waals and electrostatic energy terms, while GROUP ENERGY does not; and SINGLE ENERGY shows the charges on the atoms for the electrostatic energy term, while GROUP ENERGY does not. Also the van der Waals and electrostatic energy values shown next to each atom with GROUP ENERGY represent the net contribution of that atom to all the nonbond interactions, while with SINGLE ENERGY the value shown represents the energy due to its interaction with just one other atom.

The values are displayed next to each atom where appropriate. For the options where two atoms are involved, the values are positioned midway between. For the options that involve angles, however, the position of the displayed value relative to the atoms to which it refers is not as obvious. This is done in the same way as in the GEOMETRY analysis, and is described in Sections 6.2.4 - 6.2.6.

User Input:

1. Pick GROUP ENERGY from the ENERGY-FORCE menu.
2. Pick the appropriate energy term desired from the menu.
3. Pick the group(s) to be analyzed from the Groups menu.
4. Pick RETURN to terminate the list.

6.4.3 SNGL ENERGY/ATOM

The SNGL ENERGY/ATOM operation allows the user to show summed energies for a single atom selected from the display. The user may choose any number (or ALL) of the energy terms from the menu and have their combined value displayed next to the atom being analyzed. Switches are also provided using the ENERGY VAR option under DEFAULTS which turn off individual energy terms, so that when ALL is selected, these terms are not included in the summation (see Section 6.4.7).

User Input:

1. Pick SNGL ENERGY/ATOM.
2. Pick the energy terms desired from the Energy Terms menu.
3. Pick RETURN to terminate the list (not needed if ALL was chosen).
4. Pick the atom to be analyzed on the display.
5. Repeat step 4 as often as desired.
6. Pick a non-atom to terminate the operation.
7. Repeat steps 2-6 as often as desired.

6.4.4 GRP ENERGY/ATOM

The GRP ENERGY/ATOM option allows the user to display summed energies for a selected group of atoms. It thus provides the same function as SNGL ENERGY/ATOM described above, except that instead of doing one evaluation at a time, values for a whole group of atoms are displayed at once. Instead of selecting individual atoms from the display, the user selects groups of atoms from the Groups menu.

The values are displayed next to each atom where appropriate. For the options where two atoms are involved, the values are positioned midway between. For the options that involve angles, however, the position of the displayed value relative to the atoms to which it refers is not as obvious. This is done in the same way as in the GEOMETRY analysis, and is described in Sections 6.2.4 - 6.2.6.

User Input:

1. Pick GRP ENERGY/ATOM from the ENERGY-FORCE menu.
2. Pick any combination of term(s) or ALL.
3. Pick the group(s) to be analyzed from the Groups menu.
4. Pick RETURN to terminate the list.

6.4.5 SNGL FORCE/ATOM

The SNGL FORCE/ATOM operation illustrates the net force on an individual atom selected from the display. The force value is displayed next to the selected atom along with a force vector showing the direction and magnitude of the force. These are shown in light blue for movable atoms and in red for fixed atoms. These force displays can be turned off for fixed atoms if desired using the FIX FORCE toggle under the DEFAULTS option described in Section 6.4.7 below.

The length of the vector shown is scaled relative to the maximum force value for all the atoms (movable atoms only if FIX FORCE is off). The atom with the largest net force is given a vector length equal to the scale factor (the default is 1.000 Å), and the rest of the force vectors are scaled appropriately. The scale factor can be changed with the SCALE FACTOR option, and the force values can be turned off so that just the vectors are displayed by using the DISPLAY FORCE option. Both of these options are provided under the DEFAULTS option described in Section 6.4.7 below.

User Input:

1. Pick SNGL FORCE/ATOM.
2. Pick the atom to be analyzed on the screen.
3. Repeat step 2 as often as desired.
4. Pick a non-atom terminate the operation.

6.4.6 GRP FORCE/ATOM

The GRP FORCE/ATOM option provides the same function as SNGL FORCE/ATOM described above, except that instead of doing one evaluation at a time, values for a whole group of atoms are displayed at once. Instead of selecting individual atoms from the display, the user selects groups of atoms from the Groups menu.

The values are displayed next to each atom where appropriate. For the options where two atoms are involved, the values are positioned midway between. For the options that involve angles, however, the position of the displayed value relative to the atoms to which it refers is not as obvious. This is done in the same way as in the GEOMETRY analysis, and is described in Sections 6.2.4 - 6.2.6.

User Input:

1. Pick GRP FORCE/ATOM from the ENERGY-FORCE menu.
2. Pick any combination of term(s) or ALL.

3. Pick the group(s) to be analyzed from the Groups menu.
4. Pick RETURN to terminate the list.

6.4.7 DEFAULTS

Picking DEFAULTS brings up the menu shown in Figure 6.10. The ENERGY VAR option allows the user to modify the simulation parameters used in energy calculations. Other options (NB CUTOFF DIST, VDW ENERGY CUT, ELEC ENERGY CUT) are provided which allow the user to change the distance and energy cutoff values and thus affect which atoms are included in displaying the nonbond interactions. Toggles are also available for setting the length of the force vectors (SCALE FACTOR) and for turning the display of the geometry, energy and force values on or off (DISPLAY GEOM, DISPLAY ENERGY, DISPLAY FORCE). The 1-4 NONBOND option determines whether 1-4 interactions are included in the analysis, and the FIX FORCE toggle determines whether fixed atoms are included in the forces that are displayed.

All of these options except for ENERGY VAR affect only the analysis display and not the energy calculations. These options are described in more detail below.

Defaults:
energy var
nb cutoff dist
10.000
vdw energy cut
-9999.999
elec energy cut
-9999.999
scale factor
1.000
display geom
display energy
display force
1-4 nbond
fix force
return

Figure 6.10:
DEFAULTS menu

ENERGY VAR

The ENERGY VAR (energy variables) option is used to turn the basic components of the energy expression on and off, allowing energy calculations to include or omit specific terms. It is exactly like the ENERGY VAR option accessed under SIMULATE/DEFAULTS, and brings up the menu shown in Figure 6.11. Please refer to Section 5.12.1 for more details.

Changes made using the options on the ENERGY VAR menu affect the analysis display immediately. It is not necessary to do another energy calculation first.

NB CUTOFF DIST, VDW ENERGY CUT, ELEC ENERGY CUT

The NB CUTOFF DIST option is used to set the cutoff distance for nonbond interactions. Atoms which are farther than this distance from the selected atom are excluded when analyzing nonbond interactions with the SINGLE ENERGY or GROUP ENERGY options. The default cutoff distance is 10.000 Å. Similarly, VDW ENERGY CUT sets the van der Waals energy cutoff value and ELEC ENERGY CUT sets the electrostatic energy cutoff. The default for both of these is -9999.999 kcal/mol.

User Input:

1. Pick NB CUTOFF DIST, VDW ENERGY CUT, or ELEC ENERGY CUT.
2. Enter the cutoff value (in Å for distance, in kcal/mol for van der Waals or electrostatic energies).

SCALE FACTOR

The SCALE FACTOR option is used to set the scale factor which determines the length of the force vectors displayed when evaluating net forces with the SNGL FORCE/ATOM and GRP FORCE/ATOM options. The atom with the largest net force is given a vector whose length is equal to the scale factor (the default is 1.000 Å), and all the rest of the force vectors are scaled appropriately. If forces are being displayed for fixed atoms (the FIX FORCE option described below is on), then fixed atoms are included when determining the largest net force. Otherwise, only the movable atoms are considered. The SCALE FACTOR has no effect on the force values themselves, just on the length of the vectors.

User Input:

1. Pick SCALE FACTOR.
2. Enter the scale factor in Å.

Energy var:
bonds
angles
inversions
use all inv
ang-ang invers
bnd-x-ang inv
torsions
use all tor
ang-ang torsion
bnd-bnd torsion
constraints
nonbonds
coulomb
Q equil/update
H bonds
user energy
return

Figure 6.11:
ENERGY VAR
menu (Dreiding
force field)

DISPLAY GEOM, DISPLAY ENERGY, DISPLAY FORCE

The DISPLAY GEOM, DISPLAY ENERGY, and DISPLAY FORCE options are switches which turn the display of geometry, energy, and force values on or off. The DISPLAY GEOM option applies to the display of any distance or angle measurement. Values that one is not particularly interested in may thus be eliminated, making the display easier to read. For example, it might be useful to turn off the force values so that one can see the force vectors more clearly when using the GRP FORCE/ATOM option. The defaults for these options are all on.

User Input:

1. Toggle DISPLAY GEOM, DISPLAY ENERGY, or DISPLAY FORCE on or off.

1-4 NONBOND

When analyzing nonbond interactions with the SINGLE ENERGY or GROUP ENERGY options, 1-4 interactions may or may not be included in the analysis. The 1-4 NONBOND option allows the user to determine this. When 1-4 NONBOND is toggled on, these interactions are included. Otherwise, they are not. The default is on.

User Input:

1. Toggle 1-4 NONBOND on or off.

FIX FORCE

When analyzing forces using the SNGL FORCE/ATOM and GRP FORCE/ATOM options, fixed atoms may or may not be included in the forces that are displayed. To have the force values and force vectors for fixed atoms displayed, toggle FIX FORCE on. Otherwise, toggle FIX FORCE off. The default is off.

User Input:

1. Toggle FIX FORCE on or off.

6.4.8 GROUP ORIGIN, PICK ORIGIN

The GROUP ORIGIN and PICK ORIGIN options are described in Sections 6.2.12 and 6.2.13.

6.4.9 CLEAR, NEW GROUP

The CLEAR and NEW GROUP options function in the same way as they do in the IDENTIFY and GEOMETRY analyses (see Sections 6.1.11- 6.1.12), except that in this case, the operations being performed are energy and force analyses, and the displayed information that is being saved or cleared is energy and force information.

6.5 CONTACTS

The operations on the CONTACTS menu shown in Figure 6.12 allow the user to check for "bad" van der Waals contacts. When EXECUTE is picked, lines are drawn to show all the "bad" contacts; that is, all the nonbond interactions with a van der Waals energy above a specified cutoff value. The VDW ENERGY CUT option is used to specify this cutoff value. A 1-4 NONBOND toggle is also provided to allow the user to include or exclude 1-4 nonbond interactions in the analysis. Typically, only interactions with a positive van der Waals energy would be considered "bad", and 1-4 nonbond interactions would not be included. The default value used for the van der Waals energy cutoff is therefore zero, and the 1-4 NONBOND default is off.

Options are provided which allow the user to display the interaction distances if desired (DISPLAY DISTANCE) or to show the van der Waals energy values for each of the "bad" contacts (DISPLAY ENERGY).

As in the other three types of analysis, CLEAR and NEW GROUP are mutually exclusive toggles. When NEW GROUP is on, the contact information displayed will be saved, otherwise it will be cleared from the screen.

An energy calculation must have been done on the structure before the contact checking analysis can be performed.

User Input:

1. Pick CONTACTS.
2. Set the 1-4 NONBOND, DISPLAY DISTANCE, DISPLAY ENERGY, and VDW ENERGY CUT options on the CONTACTS menu as desired.
3. Pick EXECUTE.
4. Repeat steps 2-3 as often as desired.

6.5.1 1-4 NONBOND

When checking for "bad" van der Waals contacts, 1-4 interactions may or may not be included in the analysis. The 1-4 NONBOND option allows the user to determine this. When 1-4 NONBOND is toggled on, these interactions are included. Otherwise, they are not. The default is off.

User Input:

1. Toggle 1-4 NONBOND on or off.

Contacts:
1-4 nbond
display distance
display energy
vdw ener cut
0.000
execute
clear
new group
return

Figure 6.12:
CONTACTS
menu

6.5.2 DISPLAY DISTANCE

The DISPLAY DISTANCE option is a switch which turns the display of the distances for all the "bad" contacts on or off. Distance values are given in Å and are displayed in cyan midway between the two atoms. The default is off.

User Input:

1. Toggle DISPLAY DISTANCE on or off.

6.5.3 DISPLAY ENERGY

The DISPLAY ENERGY option is a switch which turns the display of the van der Waals energy values for all the "bad" contacts on or off. Energy values are given in kcal/mol and are displayed in red midway between the two atoms. The default is off.

User Input:

1. Toggle DISPLAY ENERGY on or off.

6.5.4 VDW ENERGY CUT

The VDW ENERGY CUT option is used to specify the van der Waals energy cutoff value which will be used in checking for "bad" contacts. All nonbond interactions with a van der Waals energy greater than this cutoff value will be displayed on the screen. Typically, only interactions with a positive van der Waals energy would be considered "bad", so the default value used is zero. The value specified cannot be negative.

User Input:

1. Pick VDW ENERGY CUT.
2. Enter the van der Waals energy cutoff value in kcal/mol (negative values are not allowed).

6.5.5 EXECUTE

Picking EXECUTE initiates the contact checking analysis. The current settings for the 1-4 NON-BOND and VDW ENERGY CUT options are used. Lines are drawn to show all the nonbond interactions with a van der Waals energy above the specified cutoff value. If the DISPLAY DISTANCE or DISPLAY ENERGY toggles are on, the interaction distances and van der Waals energy values will also be displayed.

6.5.6 CLEAR, NEW GROUP

The CLEAR, and NEW GROUP options function in the same way as they do in the other three types of analysis (IDENTIFY, GEOMETRY and ENERGY-FORCE), except that in this case, the operation being performed is checking for "bad" contacts, and the displayed information that is

being saved or cleared is contact information. These options are described in Sections 6.1.11-6.1.12.

6.6 MATCH

The operations on the MATCH menu shown in Figure 6.13 allow the user to match structures in three ways:

1. by least-squares fitting the list of atomic coordinates (GROUPS),
2. by least-squares fitting the coordinates of three or more pairs of atoms in the two structures (PAIRS, EXECUTE),
3. by least-squares fitting the coordinates of the geometric centers of mass of three or more groups of atoms in the two structures (CENTROIDS, EXECUTE).

The matched structures may be saved so that they can be read in together later, preserving the matched coordinates (WRITE).

Match:
groups
pairs
centroids
execute
write
return

Figure 6.13:
MATCH menu

6.6.1 GROUPS

The GROUPS option allows the user to perform a least-squares fit on the lists of atomic coordinates for two structures and thereby effect a match. The program prompts the user to pick the groups that contain the atoms of the structure to be matched (and moved). The groups containing template atoms must then be selected. Finally, the user picks the file containing the coordinates to be matched (and modified). The match is performed, the coordinates of the atoms of the matching structure are moved accordingly, and the calculated RMS results are reported in the text window.

Note: The two groups should contain the same number of atoms, since the two lists of atomic coordinates are compared directly in the least-squares fitting.

User Input:

1. Pick GROUPS.
2. Pick the groups containing the atoms to be matched (and moved) from the Groups menu.
3. Pick the groups containing the template atoms from the Groups menu.
4. Pick the file to be matched (and modified) from the Files menu.

6.6.2 PAIRS

The PAIRS option allows the user to select pairs of atoms from the two structures to be least-squares fit. *Note:* The first atom in each pair must be picked from the structure to be *moved* in effecting the match. At least three pairs of atoms must be selected.

User Input:

1. Pick PAIRS.
2. Pick the pairs of atoms from the display to be least-squares fit or to be constrained harmonically. (For each pair: pick the atom to be matched first, then pick the corresponding atom from the template structure).

3. Pick a non-atom to terminate the list.
4. Use the EXECUTE function to effect the match (see Section 6.6.4 below).

6.6.3 CENTROIDS

The CENTROIDS option allows the user to do a least squares fit on the geometric centers (centroids) of selected groups of atoms from two structures. At least three pairs of centroids must be selected. The range of atoms to be included in each group is defined by picking a base atom and one or more stopping atoms. The range selection for the group is completed by picking EXECUTE from the adjacent menu (ABORT aborts the process). (A description of range picking is given in Section 1.5.1.) A colored square appears at the geometric center (centroid) of the selected group. The user defines the first group by picking atoms from the structure to be *moved*. The corresponding group whose centroid is to be matched is then selected from the template structure. This is repeated until all the centroid pairs have been defined (at least three pairs of centroids must be selected). The user then indicates that the whole centroid selection process is completed by picking EXECUTE again from the adjacent menu. ABORT may be picked at any time to abort the whole process. The EXECUTE option on the MATCH menu is then picked to effect the match.

User Input:

1. Pick CENTROIDS.
2. Pick the groups of atoms from the display whose centroids are to be least squares fit. (For each pair of centroids: pick the group to be matched (and moved) first, complete by picking EXECUTE from the adjacent menu; then pick the corresponding group from the template structure, complete by picking EXECUTE from the adjacent menu) At least three pairs of centroids must be selected. See Section 1.5.1 on how to define a range of atoms.
3. Pick EXECUTE from the adjacent menu to complete the centroids selection process.
4. Use the EXECUTE function on the MATCH menu to effect the match (see Section 6.6.4).

6.6.4 EXECUTE

Toggling EXECUTE performs a least-squares fit of the coordinates of the atom pairs selected (PAIRS), or a least-squares fit of the coordinates of the centers of mass of the groups selected (CENTROIDS). The program prompts the user to select the file that contains the structure to be moved (and modified) to effect the match. As the match is performed, the coordinates of the atoms of the matching structure are moved accordingly. Upon completion, the program reports the calculated RMS results in the text window.

User Input:

1. Pick EXECUTE.
2. Pick the coordinate file to be modified from the Files menu.

6.6.5 WRITE

The WRITE option allows the user to save matched structures in BioDesign General Format so that they can be read in together later, preserving the matched coordinates.

User Input:

1. Pick WRITE.
2. Pick the files to be output from the Files menu. Terminate by picking RETURN. (This step is skipped if there is only one file in the system).
3. Enter a name for the output file including disk and directory information if needed (*bgr* is used for the extension).
4. Enter remarks. Terminate by pressing <Enter> or clicking the OK button.

6.7 TRAJECTORY

The options on the TRAJECTORY menu shown in Figure 6.14 allow the user to make two-dimensional (2D) or three-dimensional (3D) plots from data in a trajectory file. Color coding of an additional variable may also be used in conjunction with the 2D or 3D plots. The user first specifies the name of the trajectory file which will be used to make the plot (DEFINE). With MAKE PLOT, the user may choose from a list of standard 2D plots where one of the energy terms is plotted against time (or conformation number), or he may create his own custom plot where he specifies the plot type and variables for each of the axes. The variables available for the custom plots include all the energy terms available for standard plots, plus some additional values. Instantaneous energy values or average energy values may be used. Geometric measurements such as distances, angles, or dihedrals may also be plotted. Defaults are available which give the user control over the appearance of the plot (PLOT DEFAULTS) and the range of data to be included (TRAJ DEFAULTS). The user may specify a scatter plot, bar plot, or continuous line plot. He has the option of using labels, tick marks, window lines, and origin lines. He can change the color of the plot, axes, or labels, and may rescale the x, y, or z axes or reposition the origin lines if desired. After the plot has been calculated and displayed, the user may change any of the defaults and have it replotted using the new default values (REPLOT). He may also select any point on the plot and have the values for that point displayed (IDENTIFY). Similarly, the structure associated with a selected point may be displayed (EXTRACT) and saved (WRITE BGF). The plot itself may also be saved (WRITE PLOT) and read in again later (READ PLOT). Options are also available to change the center of rotation making viewing easier (FILE ORIGIN, PICK ORIGIN, GROUP ORIGIN, ZERO ORIGIN). If captions, arrows, or borders are desired, the CAPTION option on the IN-OUT menu may be used (see Section 2.4).

6.7.1 DEFINE

The DEFINE option allows the user to define the trajectory file which will be used to make a plot. If more than one file has been read into the system, the user is prompted to pick the file for which the trajectory was written from the Files menu. The name of the trajectory (*trj*) file is then entered.

User Input:

1. Pick DEFINE.
2. Pick the the file for which the trajectory was written from the Files menu. (This step is skipped if there is only one file in the system).
3. Enter the name of the trajectory file to be plotted. (See Section 1.6 on entering file names.)

Trajectory:
define
traj defaults
plot defaults
read plot
make plot
replot
identify
extract
write plot
write bgf
file origin
pick origin
group origin
zero origin
return

Figure 6.14:
TRAJECTORY
menu

6.7.2 TRAJ DEFAULTS

The options on the TRAJ DEFAULTS menu shown in Figure 6.15 allow the user to restrict the sequence of conformations to be included in the plot. The START and STOP options specify the range of conformations to be included by defining the starting and stopping sequence number or time value. The read increment (READ INC) specifies that every n th snapshot will be included. Initially, all the data in the trajectory file is included. Thus, the values initially displayed under START and STOP on the menu are those for the first and last conformation in the file. Similarly, the default READ INCREMENT is 1. Changes in the trajectory default values are not implemented until MAKE PLOT or REPLOT are used (see Sections 6.7.5 and 6.7.6).

Traj defaults:
start
0.100
stop
20.001
read increment
1
return

START, STOP

The START and STOP options are used to restrict the sequence of conformations to be included in the plot. A starting and stopping sequence number or time value must be specified. If the trajectory file was created using dynamics calculations, then the starting and stopping conformations are specified by giving their time values (in picoseconds). If conformational search or constrained minimization calculations were used to generate the trajectory file, then the conformations' sequence numbers are used. Initially, all the data in the trajectory file is included. Thus, the values initially displayed under START and STOP on the menu are those for the first and last conformation in the file.

Figure 6.15:
TRAJ DEFAULTS
menu

User Input:

1. Pick START.
2. Enter the sequence number or time value (in picoseconds) for the first conformation to be included in the plot.
3. Pick STOP.
4. Enter the sequence number or time value (in picoseconds) for the last conformation to be included in the plot.

READ INCREMENT

The READ INCREMENT option specifies that every n th conformation will be included in the plot. The default READ INC is 1.

User Input:

1. Pick READ INCREMENT.
2. Enter the new read increment.

6.7.3 PLOT DEFAULTS

The options on the PLOT DEFAULTS menu allow the user to change the appearance of plots created from trajectory file data. The PLOT DEFAULTS options are described under PLOT, Section 6.8.3. Changes made are implemented using the MAKE PLOT and REPLOT functions described below.

6.7.4 READ PLOT

The READ PLOT option allows the user to read in a plot file and have the plot displayed on the screen. This is intended to be used for plot files associated with the trajectory file read in.

User Input:

1. Pick READ PLOT.
2. Enter the name of the plot file to be read in. See Section 1.6 on entering file names (*plot* is used for the extension).

6.7.5 MAKE PLOT

The MAKE PLOT option allows the user to make two-dimensional (2D) or three-dimensional (3D) plots from the data in the trajectory file specified. The user selects a group to hold the plot from the Groups menu. He then specifies whether he wants to create a standard 2D plot where one of the energy terms is plotted against time (or conformation number), or to create his own custom plot. This is done by choosing STANDARD or CUSTOM from the Plot type menu (see Figure 6.16).

Selections:
standard
custom
return

Figure 6.16: Plot type menu

Standard Plots:

If doing a standard plot, the list shown in Figure 6.17 appears in the menu. Twelve different energy terms may be plotted against time (or conformation number): the total potential energy (POTENTIAL ENER), the energy due to bonds (BONDS), angles (ANGLES), torsions (TORSIONS), inversions (INVERSIONS), van der Waals interactions (VDW), electrostatic energy (ELECTROSTATIC), hydrogen bonds (H BONDS), constraints (CONSTRAINTS), user energy terms (USER), total kinetic energy (KINETIC ENERGY), and the sum of the total potential energy and the total kinetic energy (POTENT+KINETIC). After choosing the energy variable desired, the plot is calculated and appears on the display. The plot is positioned at $z = 0$; if the plot appears dim, the z translate or depth cueing functions may be used to brighten it. Time or conformation number is always plotted on the x axis. The range of data included and the appearance of the plot are determined by the current values specified for the trajectory defaults and the plot defaults (see TRAJ DEFAULTS and PLOT DEFAULTS, Sections 6.7.2 and 6.7.3). If desired, these values may be changed and the data may be replotted using REPLOT (see Section 6.7.6).

The names of the two variables plotted also appear in the Identify points menu. After toggling IDENTIFY, the user may select a point on the plot and have the x and y values for that point displayed in the boxes beneath the variable names as shown in Figure 6.20 (see Section 6.7.7).

User Input:

1. Pick STANDARD from the Plot type menu.
2. Pick the energy term to be plotted against time (or conformation number) from the Energy terms menu.

Custom Plots:

If choosing to do a custom plot, the user must first select the plot type desired from the Plot types menu shown in Figure 6.18a: 2D, 2D COLOR, 3D, or 3D COLOR. He then selects the quantity (variable) to be plotted for each of the axes. If a 2D COLOR or 3D COLOR plot is being done, the quantity to be color-coded is also selected. The four types of quantities available are: time (TIME) or conformation number (CONF #), the average energies over each dynamics cycle (AVERAGE ENERGY), the instantaneous energies (INSTANT ENERGY), and geometric measurements (GEOMETRY) (see Figure 6.18b).

Selections:
potential ener
bonds
angles
torsions
inversions
vdw
electrostatic
H bonds
constraints
user
kinetic energy
potent+kinetic
return

Figure 6.17:
Energy terms
menu (Standard
plots)

Selections:
2D
2D color
3D
3D color
return

Figure 6.18 a:
Plot types menu
(Custom plots)

Selections:
time
average energy
instant energy
geometry
return

Figure 6.18 b:
Quantity types
menu (Custom
plots)

An additional menu is displayed for all quantity types except TIME and CONF #. If AVERAGE ENERGY or INSTANT ENERGY is selected, the list of energy terms shown in Figure 6.19a is displayed. All the energy terms available for standard plots are included as well as some additional terms: internal energy (INTERNAL) which is the sum of the energies due to bonds, angles, torsions and inversions; nonbond energy (NONBOND) which is the sum of the van der Waals, electrostatic, and hydrogen bond energies; and the dynamics run temperature (ATOM TEMP). The geometric measurements available are DISTANCE, ANGLE, and DIHEDRAL, and RADIUS GYRATION (see Figure 6.19b). The user specifies the particular distance, angle, or dihedral to be plotted by picking the appropriate atoms from the structure on the display. The user may specify a label name for the axis if a geometric measurement is being plotted. (All other variables use the name shown on the menu; "_A" is appended to the labels for the AVERAGE ENERGY terms to indicate they are averages). Custom plots are calculated and displayed as described above for standard plots.

Selections:
potential ener
bonds
angles
torsions
inversions
vdw
electrostatic
H bonds
constraints
user
internal
nonbond
kinetic energy
potent+kinetic
atom temp
return

a) Energy terms menu

Selections:
distance
angle
dihedral
radius gyration
return

b) Geometric measurements menu

User Input:

1. Pick CUSTOM from the MAKE PLOT menu.
2. Pick the plot type desired from the Plot types menu (2D, 2D COLOR, 3D, or 3D COLOR).
3. Pick the quantity type for the x axis from the Quantity types menu (TIME or CONF #, AVERAGE ENERGY, INSTANT ENERGY, or GEOMETRY).
4. If TIME or CONF # was chosen, skip to step 6. Pick the energy term or geometric measurement to be plotted from the menu.
5. If a geometric measurement is being plotted:
 - 1) Pick the atoms which define it from the display.
 - 2) Enter a label for the x axis.
6. Repeat steps 3 - 5 for the y and z axes (if applicable), and for the color coded quantity if doing a 2D COLOR or 3D COLOR plot.

6.7.6 REPLOT

The REPLOT option allows the user to change the range of data included or the appearance of the plot. The options on the TRAJ DEFAULTS and PLT DEF menus are first used to change the trajectory defaults and plot defaults desired (see Section 6.7.2 and 6.7.3). The user then toggles

REPLOT and the data is replotted using the new values. REPLOT applies to the currently active plot only.

User Input:

1. Use the options on the PLOT DEFAULTS or TRAJ DEFAULTS menus to change the plot defaults and trajectory defaults desired (see Sections 6.7.2 and 6.7.3).
2. Toggle REPLOT.

6.7.7 IDENTIFY

The IDENTIFY option allows the user to select a point on the currently active plot and have the values for that point displayed in the Identify points menu. The values appear directly beneath the label names for each of the data columns. An example is shown in Figure 6.20.

User Input:

1. Pick IDENTIFY.
2. Select the point on the plot whose values are to be displayed.
3. Repeat step 2 as many times as required.
4. Pick any non-point on the plot to terminate the function.

Identify points:
time
1.500
energy
225.000

Figure 6.20:
Identify points
menu example

6.7.8 EXTRACT

The EXTRACT option allows the user to select a point on the currently active plot and have the structure associated with that point displayed on the screen. The plot must have been created from the trajectory file specified under DEFINE.

User Input:

1. Pick EXTRACT.
2. Select the point on the plot whose structure is to be displayed.

6.7.9 WRITE PLOT

The WRITE PLOT option allows the user to save the currently active plot information in a plot file. The plot can then be read in later using the READ PLOT option (see Section 6.7.4).

User Input:

1. Pick WRITE PLOT.
2. Enter a name for the plot file including directory information if needed (*plot* is used for the extension).

6.7.10 WRITE BGF

The WRITE BGF option allows the user to save the coordinates of the current structure in a BioDesign General Format (*bgf*) file.

User Input:

1. Pick WRITE BGF.
2. Enter a name for the output file including disk and directory information if needed (*bgf* is used for the extension).

6.7.11 FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN

These options are used to reset the center of rotation. FILE ORIGIN sets it to the geometric center of a selected input data file, GROUP ORIGIN sets it to the geometric center of a set of groups selected from the Groups menu, and PICK ORIGIN sets it to an atom picked from the display. ZERO ORIGIN sets the center of rotation to $x, y, z = 0, 0, 0$. These options are described in Sections 3.9.8 - 3.9.11.

6.8 PLOT

The PLOT option allows the user to display a plot, change its appearance, or merge its data with the data from another plot file. The PLOT menu is shown in Figure 6.21. DEFINE is used to specify the plot file which is then read in and appears on the display. Any plot file which meets the format requirements can be used; these include those generated using conformational search, trajectory analysis, and polymer analysis. A group is automatically created to hold the plot. The user may select any point on the plot and have the values for that point displayed (IDENTIFY). He may change the defaults which determine the plot type and variables used, or alter those which control the appearance of the plot (PLOT DEFAULTS) and then REPLOT it using the new values. The current plot may be retained by creating a NEW GROUP to hold the replotted data. There is also a MERGE function which allows the user to merge data from multiple plot files into a single plot file. MERGE also allows simple arithmetic operations to be performed on the data in columns with the same name. The plot may be saved using the WRITE PLOT option. Options are also available to change the center of rotation making viewing easier (FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN). If captions, arrows, or borders are desired, the CAPTION option on the IN-OUT menu may be used (see Section 2.4).

Plot:
define
new group
plot defaults
replot
identify
merge
write plot
file origin
group origin
pick origin
zero origin
return

6.8.1 DEFINE

The DEFINE option is used to specify the plot file which is to be used. Any plot file which meets the format requirements outlined in Appendix D can be used. Plot files generated within the program are all appropriate. These include those generated using conformational search, trajectory analysis, and polymer analysis (see GRID SEARCH and M-CARLO SEARCH, Sections 5.10 and 5.11, and TRAJECTORY, Section 6.7). After entering the name of the plot file, the user specifies a name for the group which will be created to display the plot. The plot then appears on the display, and the names for each of the variables are shown in the DEFINE menu.

Figure 6.21:
PLOT menu

User Input:

1. Pick DEFINE.
2. Enter the name of the plot file. Include disk and directory information if needed (*plot* is used for the extension).
3. Pick the group to be created from the Groups menu.
4. Enter a name for the group.

6.8.2 NEW GROUP

The NEW GROUP option is used to create a new group which will hold the display information for the next plot. This option is used in conjunction with REPLOT (see Section 6.8.4). For ex-

ample, the user may wish to change the plot defaults and replot the data. In order to retain the current plot, however, he must create a new group to hold the updated plot. (If NEW GROUP is not used, REPLOT overwrites the plot data stored in the last group.)

User Input:

1. Pick NEW GROUP.
2. Pick the group to hold the next plot from the Groups menu.
3. Enter a name for the new group.
4. Pick REPLOT to replot the data and store it in the new group.

6.8.3 PLOT DEFAULTS

The options on the PLOT DEFAULTS menu shown in Figure 6.22 allow the user to specify the plot defaults which will be applied when the REPLOT function is used (see Section 6.8.4).

PLOT TYPE defines the plot type. Two-dimensional (2D), three-dimensional (3D), and contour plots are available. Color-coding of an additional variable may also be used in conjunction with the 2D and 3D plots. COLUMNS specifies which variables will be used and assigns each to the axis desired. CONTOUR specifies the number of contour lines which will be drawn, and sets the base value and increment used for contour plots. Currently, contour plots can only be created by the program using conformational search (see Sections 5.10 and 5.11). The WINDOW option may be used to stretch out or contract the data along any of the axes, or to crop it so that only a portion of the data is included. This is done by specifying new minimum or maximum values for the axes (initially default values are chosen so that the plot displays all available data). Similarly, default values are used for the x, y, and z origins, but these may also be changed using the ORIGIN option. The SYMBOL option is used to specify a scatter plot, a bar plot, or a continuous line plot. COLORS is used to change the color of the axes, labels, or data; it also allows the user to specify the high and low-end colors which define the color range used when color coding. DISPLAY allows the user to determine whether display features such as labels, tickmarks, the window around the plot, and origin lines are included or not. The LINE WIDTH option specifies the width of the lines used to draw the plot, and RESET resets all the plot defaults to their original default values.

Plot defaults:
plot type
columns
window
origin
contour
symbol
colors
display
line width
1
reset
return

Figure 6.22:
PLOT
DEFAULTS menu

User Input:

1. Change the plot default settings as desired using the options on the PLOT DEFAULTS menu.
2. Use the REPLOT function to implement the changes (see Section 6.8.4).

PLOT TYPE

The PLOT TYPE option is used to define the plot type. After toggling PLOT TYPE, the user selects from the list of options shown in the Plot types menu (see Figure 6.23). Two-dimensional (2D), three-dimensional (3D), and 2D contour (2D CONTOUR) plots are available. Color-coding of an additional variable may also be used in conjunction with the 2D and 3D plots (2D COLOR, 3D COLOR).

User Input:

1. Pick PLOT TYPE.
2. Select the plot type desired from the Plot types menu (2D, 2D COLOR, 2D CONTOUR, 3D, or 3D COLOR).

Selections:
2D
2D color
2D contour
3D
3D color
return

Figure 6.23: Plot types menu

COLUMNS

The COLUMNS option allows the user to specify which variables will be used in the plot and to assign them to the axis desired. After choosing COLUMNS, a list of the column labels contained in the plot file is shown in the Column labels menu. They are listed in the same order as they appear in the file. An example is shown in Figure 6.24. The user is then successively prompted to select the variable to be plotted against the x, y, and (if applicable) z axes; if a color plot is being done, he is also asked to select the variable which is to be color coded. Changes to COLUMNS are implemented via the REPLOT function.

User Input:

1. Pick COLUMNS.
2. Select the variable for the indicated axis from the Column labels menu.
3. Repeat step 2 as required.

Selections:
conformer #
dihedral 1
dihedral 2
energy
return

Figure 6.24: Column labels menu (example)

WINDOW

Initially, the minimum and maximum values for each of the axes are chosen so that the plot displays all available data. The WINDOW option allows the user to change these values so that the data is stretched out or contracted along any of the axes, or is cropped so that only a portion of the data is included. Similarly, for color-coded plots, initially the LOW DATA color is assigned to the

lowest data point for the variable being color coded, and the HIGH DATA color is assigned to the highest data point; the points in between are assigned colors based on a linear interpolation. The minimum and maximum data values for the color coded variable can be changed. For example, by specifying a lower maximum data value outlying data can be excluded thus broadening the range of colors used for the rest of the data.

After toggling WINDOW, the user picks the item whose value he wants to change from the Window items menu (see Figure 6.25). The first six items set the minimum and maximum values for the x, y, and z axes; the last two (COLOR QUAN MIN and COLOR QUAN MAX) determine the minimum and maximum values for the color coded quantity. Initially, dashed lines are shown beneath each of the menu items, indicating that default values are being used. The user then enters the new value; this number then appears beneath the selected item on the menu.

User Input:

1. Pick WINDOW.
2. Pick the item whose value is to be changed from the Window items menu.
3. Enter the new minimum or maximum value (or enter 999.99 to select the default setting).
4. Repeat steps 2 and 3 to change another value or pick RETURN to complete the operation.

Selections:
x min
- - -
x max
- - -
y min
- - -
y max
- - -
z min
- - -
z max
- - -
color quan min
- - -
color quan max
- - -
return

Figure 6.25:
Window items
menu

ORIGIN

The ORIGIN option allows the user to specify the values to be used for the x, y, and (if applicable) z origins. Initially, default values are used, so that the origin lines are coincident with the window lines. However, the origin lines may be placed elsewhere if desired. For example, if the y variable includes both positive and negative values, it might be useful to have the y origin at $y = 0$. The origin to be changed is selected from the Origins menu (see Figure 6.26). Initially, dashed lines are shown beneath the menu items, indicating that default values are being used. When the new origin value is entered, it is displayed beneath the appropriate menu item.

User Input:

1. Pick ORIGIN.
2. Select the origin which is to be changed from the Origins menu (X ORIGIN, Y ORIGIN, or Z ORIGIN).
3. Enter the new origin value (or enter 999.99 to select the default setting).
4. Repeat steps 2 and 3 to change additional origin values or pick RETURN to complete the operation.

Selections:
x origin
- - -
y origin
- - -
z origin
- - -
return

Figure 6.26:
Origins menu

CONTOUR

The CONTOUR option allows the user to specify the number of contour lines which will be drawn, and to set the base value and increment used for contour plots. Currently, contour plots can only be created by the program using data generated under conformational search (see Sections 5.10 and 5.11). After toggling CONTOUR, the user selects the item to be changed from the Contour items menu (BASE, INCREMENT, or NUMBER) (see Figure 6.27) and then enters the new value. NUMBER defines the number of contour lines to be drawn. INCREMENT specifies the value which is initially added to the BASE to define the first contour line and is subsequently added to the previous contour line to define each additional contour line. The newly entered value will appear beneath the appropriate item on the menu. Initially, dashed lines appear beneath BASE and INCREMENT indicating that default values are being used. By default, the NUMBER of contour lines is one, the BASE is the minimum value, and the $INCREMENT = (maximum - minimum) / (NUMBER + 1)$. If only a single contour line is specified, then the line is color coded using HIGH DATA (see COLORS, below). If more than one contour line is specified, then the lines are automatically color coded using the range of colors specified by LOW DATA and HIGH DATA. The first contour line is assigned the LOW DATA color, the last contour line is assigned the HIGH DATA color, and the lines in between are assigned colors based upon a linear interpolation.

Selections:
base
- - -
increment
- - -
number
1
return

Figure 6.27:
Contour items
menu

User Input:

1. Pick CONTOUR.
2. Pick the item to be changed from the Contour items menu (BASE, INCREMENT, or NUMBER).
3. Enter the new value (or enter 999.99 to select the default setting).
4. Repeat steps 2 and 3 to make additional changes or pick RETURN to complete the operation.

SYMBOL

The SYMBOL option is used to specify whether a scatter plot, a bar plot, or a connected line plot will be used. After toggling SYMBOL, the user selects the type desired from the Symbol types menu (see Figure 6.28). The first three items specify a scatter plot; three symbol types are allowed: BOX, PLUS, or X. BAR is used for a bar plot and CONNECT indicates that a connected line plot is to be used. CONNECT is the default.

User Input:

1. Pick SYMBOL.
2. Pick the type of plot desired from the Symbol types menu (BOX, PLUS, X, BAR or CONNECT).
3. Pick RETURN to complete the operation.

Selections:
box
plus
X
bar
connect
return

Figure 6.28:
Symbol types
menu

COLORS

The COLORS option is used to change the color of the window, origin, tick marks, labels, or data; it also allows the user to specify the high- and low-end colors which define the color range used when color coding. After toggling COLORS, the user picks the item whose color he wants to change from the Plot items menu (WINDOW, ORIGIN, TICK MARKS, LABELS, LOW DATA, or HIGH DATA) (see Figure 6.29). The current color for each of these items is displayed directly beneath the item on the menu. WINDOW refers to the box drawn around the plot, ORIGIN refers to the origin lines, and TICK MARKS refers to the tick marks. LABELS is used to change the color of the text, and HIGH DATA specifies the color of the symbols or lines used to plot the data. HIGH DATA is also used together with LOW DATA to define the range of colors used when doing 2D COLOR or 3D COLOR plots. After picking the item, the user selects the new color from the Colors menu which appears (see Figure 6.30). He is then returned to the Plot items menu where he can make additional changes if desired.

Selections:
window
green
origin
green
tick marks
green
labels
red
low data
blue
high data
red
return

Figure 6.29: Plot items menu

Selections:
purple
magenta
pink
red
orange
yellow
chartreuse
green
aqua
cyan
light blue
blue
white
return

Figure 6.30: Colors menu

The default colors are: green for the WINDOW, ORIGIN, and TICK MARKS, red for the LABELS, blue for LOW DATA, and red for HIGH DATA.

User Input:

1. Pick COLORS.
2. Select the item whose color is to be changed from the Plot items menu (WINDOW, ORIGIN, TICK MARKS, LABELS, LOW DATA, or HIGH DATA).
3. Pick the color desired from the Colors menu. The new color will be displayed beneath the item on the menu.
4. Repeat steps 2 and 3 to change the color of another item or pick RETURN to complete the operation.

DISPLAY

The **DISPLAY** option allows the user to determine whether display features such as labels, tickmarks, the window around the plot, and origin lines are included or excluded from the plot. After toggling **DISPLAY**, the Display items menu shown in Figure 6.31 appears. Items which are "on" (highlighted) are included in the plot, while those which are "off" (not highlighted) are not. Toggling the item switches it "on" or "off". **WINDOW** refers to the box drawn around the plot, **ORIGIN** refers to the origin lines, and **TICK MARKS** refers to the tick marks; three tick marks are used for each axes: one at each of the end points, and one in the middle. **LABELS** refers to the text which appears directly beneath each of the axes. The labels include the name of the variable being plotted (placed in the middle) and the minimum and maximum data values (placed at the low and high ends of the axes respectively). The default is to include all the display features in the plot.

Selections:
window
origin
tick marks
labels
return

Figure 6.31:
Display items
menu

User Input:

1. Toggle the display item in the menu "on" to include it in the plot, or toggle it "off" to exclude it. Pick **RETURN** to complete the selection.

LINE WIDTH

The **LINE WIDTH** option allows the user to change the width of the lines used to draw the plot symbols or lines, the window, the origin lines, and the tick marks. The default **LINE WIDTH** is 1 pixel.

User Input:

1. Toggle **LINE WIDTH**.
2. Enter the new line width value in pixels.

RESET

Toggling **RESET** resets all the plot defaults to their original default values.

6.8.4 REPLOT

The **REPLOT** option implements changes made in the plot type, variables, or appearance of the plot, or in the plot file data. The options on the **PLOT DEFAULTS** menu are first used to change the plot defaults desired (see Section 6.8.3) or the **MERGE** option is used to merge data from multiple plot files (see Section 6.8.6). The user then toggles **REPLOT** and the data for the currently active plot is replotted using the new values. If a new group was created, then the replotted data will be stored there; otherwise it will overwrite the plot data stored in the last created group (see **NEW GROUP**, Section 6.8.2)

User Input:

1. Use the options on the PLOT DEFAULTS menu to change the plot defaults desired (see Section 6.8.3) or use MERGE to merge data from multiple plot files (see Section 6.8.6).
2. Use NEW GROUP to create a new group if you do not want to overwrite the last group (see Section 6.8.2).
3. Toggle REPLOT.

6.8.5 IDENTIFY

The IDENTIFY option allows the user to select a point on the currently active plot and have the values for that point displayed in the menu. The values appear directly beneath the label names for each of the data columns. An example is given in Figure 6.32.

User Input:

1. Pick IDENTIFY.
2. Select the point on the plot whose values are to be displayed.
3. Repeat step 2 as many times as required.
4. Pick any non-point on the plot to terminate the function.

Identify:
time
1.500
energy
225.000

Figure 6.32:
Identify menu
(example)

6.8.6 MERGE

The MERGE option allows the user to merge data from multiple plot files into a single plot file. Merging is done in pairs (only one file may be added to the current one at a time). If there are more columns of data in the file being merged, the new file will expand to include them. Simple arithmetic operations may be performed on the data in columns with the same name. However, the two files must have the same number of data records. (See Appendix D for a description of the plot file format.)

The user reads in and displays the first plot file (see DEFINE, Section 6.8.1). A new menu appears which lists the columns of data contained in the plot file. He then toggles MERGE and specifies the name of the plot file to be merged with the current one. If the two files have columns of data with the same name, the program prompts the user to select the arithmetic operations to be performed on the data in common columns. The options appear in the MERGE menu (see Figure 6.33). The user may keep the data from the current file (KEEP OLD), keep the data from the new file (KEEP NEW), add the two together (OLD + NEW), subtract them (OLD - NEW or NEW - OLD), multiply them (OLD * NEW), or divide them (OLD/NEW or NEW/OLD). The merge is implemented and a new plot is displayed by using the REPLOT function (see Section 6.8.4).

Selections:
keep old
keep new
old + new
old - new
new - old
old * new
old / new
new / old
return

Figure 6.33:
Arithmetic
operations menu
(Merge function)

User Input:

1. Use DEFINE to read in the first plot file (see Section 6.8.1).
2. Pick MERGE.
3. Enter the name of the plot file to be merged. See Section 1.6 on entering file names (*plot* is used for the extension).
4. If there are commonly named columns, select the arithmetic operation to be performed on the specified column from the menu. Repeat for each common column.
5. Use the REPLOT function to implement the merge (see Section 6.8.4).

6.8.7 WRITE PLOT

The WRITE PLOT option allows the user to save the currently active plot information in a plot file. The plot can then be read in later using DEFINE (Section 6.8.1) or with the READ PLOT option under trajectory analysis (see Section 6.7.4).

User Input:

1. Pick WRITE PLOT.
2. Enter a name for the plot file including directory information if needed (*plot* is used for the extension).

6.8.8 FILE ORIGIN, GROUP ORIGIN, PICK ORIGIN, ZERO ORIGIN

These options are used to reset the center of rotation. FILE ORIGIN sets it to the geometric center of a selected input data file, GROUP ORIGIN sets it to the geometric center of a set of groups selected from the Groups menu, and PICK ORIGIN sets it to an atom picked from the display. ZERO ORIGIN sets the center of rotation to $x, y, z = 0, 0, 0$. These options are described in Sections 3.9.8 - 3.9.11.

6.9 PROTEIN

The PROTEIN operations allow the user to check proteins for irregular structural features including D-amino acids, *cis*-peptides, and non-Ramachandran phi and psi angles. The dihedral angle values for disulfide linkages can also be determined. The PROTEIN menu is shown in Figure 6.34. After choosing one of the options on the menu, the program checks the structure for the feature and lists the results in the text window. If any irregularities are found, the associated residue number(s) and values are listed.

User Input:

1. Pick the file(s) to be analyzed from the Files menu. (This is skipped if there is only one file in the system.)

Protein:
stereochemistry
peptide linkage
phi-psi angles
disulfide bonds
return

Figure 6.34:
PROTEIN menu

6.9.1 STEREOCHEMISTRY

The STEREOCHEMISTRY option allows the user to check proteins for correct stereochemistry. If any D-amino acids are found, or if any threonines or isoleucines do not have the correct stereochemistry, the associated residue number(s) are listed in the text window.

User Input:

1. Pick STEREOCHEMISTRY. The results are listed in the text window.

6.9.2 PEPTIDE LINKAGE

The PEPTIDE LINKAGE option allows the user to check proteins for *cis*-peptides. If any are found, the associated residue number(s) are listed on the text display. *Cis*-peptides preceding prolines are also indicated.

User Input:

1. Pick PEPTIDE LINKAGE. The results are listed in the text window.

6.9.3 PHI-PSI ANGLES

The PHI-PSI ANGLES option allows the user to check proteins for non-Ramachandran phi and psi angles. Phi angles of -44° to $+180^\circ$ and psi angles of -160° to -80° and -20° to $+20^\circ$ are considered unusual. The phi and psi angles for all the residues are listed in the text window; residues with unusual phi or psi angle values are also listed separately.

User Input:

1. Pick PHI-PSI ANGLES. The results are listed in the text window.

6.9.4 DISULFIDE BONDS

The DISULFIDE BONDS option checks proteins for disulfide linkages and if any are found lists the dihedral angle values and associated residue numbers in the text window.

User Input:

1. Pick DISULFIDE BONDS. The results are listed in the text window.

6.10 VOLUMES

The VOLUMES operations allow the user to calculate and display volumes or surface areas for selected groups of atoms. The VOLUMES menu is shown in Figure 6.35.

The user toggles CALCULATE GRID and then selects the groups for volume analysis from the Groups menu. The uniform finite element method of T. R. Stouch and Peter C. Jurs¹ is used to calculate the volumes and surface areas. Basically, the volume of a structure is determined based on the number of cubic grid elements it occupies. If the center of a cubic grid element lies within the radius of any atom, then the element is considered occupied. The atom radii are calculated using the van der Waals radius multiplied by a user adjustable scale factor (VDW SCALE) plus a user adjustable probe radius (PROBE RADIUS) for solvent accessible volumes:

$$\text{atom radius} = (\text{VDW radius} \times \text{VDW SCALE}) + \text{PROBE RADIUS}$$

The VDW SCALE and PROBE RADIUS values are set under DEFAULTS. A DENSITY option is also available which defines the density of dots used to display the volumes. CALCULATE GRID calculates the volume element grids for the groups selected. Grids with the same names are also created and will appear later in a Grid menu. CUBIC VOLUME may then be used to compute the volume of selected grids; similarly, their surface area can be computed using the SURFACE AREA option. Grid combination operations can be performed on selected grids to give their UNION, intersection (INTERSECTION), the difference between the union and intersection (RESIDUAL), and the free volume not occupied by any grid (INVERSE). A SURFACE ONLY option is available which creates a volume grid of the surfaces of a selected grid. SURFACE ONLY is used in combination with DISPLAY to display volume surfaces as an array of dots. DISPLAY will also display volumes (as a 3 dimensional array of dots), but it is not intended to be used for this purpose. (Since the maximum number of dots that can be displayed is 20,000, only small structures can be displayed in this way.) CHANGE DISPLAY brings up the Groups menu so that selected groups may be turned on or off. The READ option allows the user to read and display dotted volume or surface files, and WRITE saves the coordinates of a grid as a dot file so that it can be read in and displayed later.

Volumes:
defaults
calculate grid
union
intersection
residual
inverse
surface only
display
cubic volume
surface area
read
write
return

Figure 6.35:
VOLUMES menu

User Input:

1. Use the DEFAULTS options to set the probe radius, van der Waals scale factor, and dot density to the desired values.
2. Use the CALCULATE GRID option to calculate volume element grids for the groups desired.

¹ T. R. Stouch and P. C. Jurs, *J. Chem. Inf. Comput. Sci.*, **26** (1986) 4-12.

3. If desired, use the grid combination operations (UNION, INTERSECTION, RESIDUAL, INVERSE) to form new grids or use SURFACE ONLY to create a surface area grid.
4. The following operations may be performed in any order on the grids created previously.
 - a) Use the CUBIC VOLUME or SURFACE AREA options to compute the volume or surface area of a selected grid.
 - b) Use the DISPLAY option to create a group and display a selected grid as an array of dots.
 - c) Use the WRITE option to save the coordinates of a grid as a dot file, or use READ to read in and display a dotted volume or surface file.

6.10.1 CALCULATE GRID

The CALCULATE GRID option calculates volume element grids for selected groups. Grids of the same name are also created which appear later in a Grid menu. The calculations are based on the uniform finite element method of T. R. Stouch and Peter C. Jurs¹. Basically, the volume of a structure is determined based on the number of cubic grid elements it occupies. If the center of a cubic grid element lies within the radius of any atom, then the element is considered occupied. The atom radii are calculated using the van der Waals radius multiplied by a user adjustable scale factor (VDW SCALE) plus a user adjustable probe radius (PROBE RADIUS) for solvent accessible volumes:

$$\text{atom radius} = (\text{VDW radius} \times \text{VDW SCALE}) + \text{PROBE RADIUS}$$

The VDW SCALE and PROBE RADIUS values are set under DEFAULTS (see Section 6.10.2). The default settings are 0.890 for VDW SCALE and 0 for PROBE RADIUS. After calculating the volume element grids, CUBIC VOLUME may be used to compute their volumes or SURFACE AREA may be used to calculate their surface areas.

Note: Using CALCULATE GRID erases all previous grids.

User Input:

1. Pick CALCULATE GRID.
2. Pick the groups for volume analysis from the Groups menu. Pick RETURN to complete the group selection. This step is skipped if there is only one group.

6.10.2 DEFAULTS

The options on the DEFAULTS menu shown in Figure 6.36 allow the user to change the defaults used in calculating the volume grids. VDW SCALE sets the van der Waals scale factor used in determining the atom radii and PROBE RADIUS sets the radius of the surface probe. The default values (0.890 for VDW SCALE and zero for PROBE RADIUS) yield the standard CPK radii. The DENSITY option defines the density of dots used for the grids in dots per linear Å.

DENSITY

The DENSITY option sets the density of dots used for the grids. The default value of 4.000 dots/Å gives about 1% accuracy in computing volumes. The range of grid densities is limited only by the amount of virtual memory and CPU time available. The DISPLAY function, however, is limited to a total of 20,000 dots (see Section 6.10.5). Grids containing more than 20,000 dots will only be partially displayed.

User Input:

1. Pick DENSITY.
2. Enter the new value (in dots/Å).

PROBE RADIUS

The PROBE RADIUS option sets the radius of the surface probe used in determining the atomic radii for the volume grid calculations:

$$\text{atom radius} = (\text{VDW radius} \times \text{VDW SCALE}) + \text{PROBE RADIUS}$$

The PROBE RADIUS value allows solvent accessible volumes to be calculated. The default value of zero produces van der Waals surfaces. A value of 1.4 Å yields a water solvent accessible surface.

User Input:

1. Pick PROBE RADIUS.
2. Enter the new value (in Å).

VDW SCALE

The VDW SCALE option sets the van der Waals scale factor used in determining the atomic radii for the volume grid calculations:

$$\text{atom radius} = (\text{VDW radius} \times \text{VDW SCALE}) + \text{PROBE RADIUS}$$

Defaults:
density
4.000
probe radius
0.000
vdw scale
0.890
return

Figure 6.36:
DEFAULTS menu

The default value of 0.890 yields the standard CPK radii (providing the probe radius is set to zero).

User Input:

1. Pick VDW SCALE.
2. Enter the new value.

6.10.3 UNION, INTERSECTION, RESIDUAL, INVERSE

Grid combination operations may be performed on selected grids which allow the user to determine their union (UNION), intersection (INTERSECTION), the difference between the union and intersection (RESIDUAL), and the free volume not occupied by any grid (INVERSE). These correspond to the binary operations: logical OR (UNION), logical AND (INTERSECTION), exclusive OR of union and intersection (RESIDUAL), and logical NOR (INVERSE). Each of these operations takes one or more input grids and produces a new output grid. After picking the option desired, the user enters a name for the grid to be created. He then picks the input grid(s) for the operation and initiates the operation by choosing EXECUTE from the bottom of the Grid menu (picking RETURN will abort the operation). When the operation is completed, the Grid menu is updated with the name of the result.

User Input:

1. Pick the grid combination operation desired (UNION, INTERSECTION, RESIDUAL, INVERSE).
2. Enter a name for the grid to be defined.
3. Pick the input grid(s) required for the operation from the Grid menu.
4. Pick EXECUTE from the Grid menu to initiate the operation or choose RETURN to abort it.

6.10.4 SURFACE ONLY

The SURFACE ONLY operation allows the user to create a volume grid of the surfaces of a selected grid (only the volume elements which have faces adjoining unoccupied elements are included). This can then be used in conjunction with the DISPLAY option to display volume surfaces (see Section 6.10.5). The maximum number of dots that can be displayed is limited to 20,000. Grids requiring more than 20,000 dots are only partially displayed. It is therefore often advisable to display the volume surface of a structure rather than the volume itself so that the entire structure may be visualized. As in the grid combination operations, the user enters a name for the grid to be created. He then picks the input grid for the operation and initiates the operation by choosing EXECUTE from the bottom of the Grid menu (picking RETURN will abort the operation). When the operation is completed, the Grid menu is updated with the name of the result.

User Input:

1. Pick SURFACE ONLY.
2. Enter a name for the grid to be defined.
3. Pick the input grid(s) required for the operation from the Grid menu.
4. Pick EXECUTE from the Grid menu to initiate the operation or choose RETURN to abort it.

6.10.5 DISPLAY

The DISPLAY option allows the user to display a grid as a 3 dimensional array of dots. The density of dots is set with the DENSITY option (Section 6.10.2). After picking DISPLAY, the user selects a group to display the grid data from the Groups menu, then enters a name for it. He then picks the grid to be displayed from the Grid menu.

Note: The maximum number of dots that can be displayed is limited to 20,000. Grids requiring more than 20,000 dots are only partially displayed; a message to that effect appears in the Modules/Messages window. It is therefore often advisable to display the volume surface of a structure rather than the volume itself so that the entire structure may be visualized. The volume grid of the surfaces of a structure can be created using the SURFACE ONLY option (see Section 6.10.4). Alternatively, one might recalculate the grid after having set the dot density to a lower value (see DENSITY, Section 6.10.2 and CALCULATE GRID, Section 6.10.1).

User Input:

1. Pick DISPLAY.
2. Enter a name for the group to be created.
3. Pick the grid to be displayed from the Grid menu.

6.10.6 CUBIC VOLUME

The CUBIC VOLUME option computes the volume of the occupied elements in a selected grid. Volume is computed by counting cubes in the grid (see CALCULATE GRID for a description of how cubic grid elements are calculated, Section 6.10.1). The user selects the grid for the volume calculation from the Grid menu. The volume calculated is shown in the Modules/Messages window in $\text{\AA}^3$.

User Input:

1. Pick CUBIC VOLUME.
2. Select the grid for the volume calculation from the Grid menu.

6.10.7 SURFACE AREA

The SURFACE AREA option computes the surface area of the occupied elements in a selected grid. Surface area is computed by counting the surface faces of the cubes in the grid and multiplying the result by $2/3$. This is the mathematically correct factor. Representing a sphere by an infinite number of tiny cubes overestimates the surface area by exactly 1.5. (See CALCULATE GRID for a description of how cubic grid elements are calculated, Section 6.10.1.) The user selects the grid for the surface area calculation from the Grid menu. The calculated surface area value is shown in the Modules/Messages box in $\text{\AA}^2$.

User Input:

1. Pick SURFACE AREA.
2. Select the grid for the surface area calculation from the Grid menu.

6.10.8 READ

The READ option allows the user to read in and display dotted volume or surface files (dot files). The user picks the group to display the dot file from the Groups menu, then enters a name for it. He then enters the name of the dot file to be read in. The data contained in the dot file is displayed in the selected group. (Although the data in the dot file can be displayed, it cannot be reconverted to a grid for volume calculations).

User Input:

1. Pick READ.
2. Pick the group to display the dot file from the Groups menu. If the group has been previously defined, it will be redefined.
3. Enter a name for the group to be created.
4. Enter the name of the *dot* file to be read in and displayed. (See Section 1.6 on how to enter file names.)

6.10.9 WRITE

The WRITE option creates a dot file containing the coordinates for a selected grid. The user selects the grid from the Grid menu and enters a name for the dot file to be created. The resultant dot file can be read in and displayed later, but it cannot be reconverted to a grid for volume calculations.

User Input:

1. Pick WRITE.
2. Pick the grid to be saved from the Grid menu.
3. Enter a name for the dot file to be created. Include disk and directory specifications if needed (*dot* is used for the extension).

Chapter 7

Utilities

The operations on the UTILITIES menu shown in Figure 7.1 allow the user to execute user-written routines (USER), reset the program (PROGRAM RESET), change the color of the nonpickable menu items (*e.g.*, values and subtitles) (LABEL COLOR), and browse through the files in the system (BROWSER).

Utilities:
user
program reset
label color
browser
return

Figure 7.1:
UTILITIES menu

7.1 USER

The USER menu shown in Figure 7.2 allows up to 10 user-written routines. Contact Molecular Simulations Customer Support for more information.

User:
user 1
user 2
⋮
user 10
return

Figure 7.2: USER
menu

7.2 PROGRAM RESET

Toggle PROGRAM RESET to reset all the atom, group and file arrays, *i.e.*, to start over without restarting the program. Resetting the program is faster than exiting and restarting it since the force field parameters and defaults files do not have to be read in.

User Input:

1. Pick PROGRAM RESET to reset all the atom, group and file arrays.

7.3 LABEL COLOR

The LABEL COLOR option allows the user to change the color of the nonpickable menu items such as values and subtitles. When LABEL COLOR is picked, a dialog box with several slider bars in it appears (see Figure 4.14 in Section 4.5). The user moves the sliders with the mouse to specify values for the red (R), green (G), and blue components (B). The hue (H), saturation (S), and value (V) may also be adjusted. The box at the bottom indicates the current color (default is red). When satisfied with the color, the user picks OK.

User Input:

1. Pick LABEL COLOR.
2. Use the sliders to specify values for the red (R), green (G), and blue (B) components; the hue (H), saturation (S), and value (V) may also be adjusted. When the box at the bottom shows the desired color, pick OK. See Section 4.5 for more details.

7.4 BROWSER

The BROWSER option brings up the browser box and allows the user to browse through all the files in the system. The current working directory may also be changed if desired. The working directory comes up whenever the browser box is invoked (*e.g.*, when reading in files). See Section 1.6 for a more complete description.

User Input:

1. Pick BROWSER.
2. Use the browser box to search through the files. File names all end with a 3- or 4-letter extension (*e.g.*, *test.bgf*). Directory names have no extension and are listed in red. To move to a subdirectory, select the directory name from the listings. To move to the next higher-level directory, pick the name at the top of the listings. If there is no higher-level directory, the top listing will be *"/*". The directory may also be changed by picking the New Dir button and entering the name of the new directory in the dialog box. See Section 1.6 for more details.

Chapter 8

Exit Operations

The EXIT operations allow the user to exit the program. After picking EXIT, an alert box appears and the user is asked to verify that he wants to exit the program. The user clicks the OK button to exit or clicks Cancel if he has changed his mind and wants to continue.

User Input:

1. Pick EXIT.
2. Click the OK button to exit the program or click Cancel to continue.

Appendix A

Standard Group Elements

A.1 Standard Residues

Standard residues used in the RESIDUES option of the MAKE GROUPS menu (see Section 4.1.5):

AMINO ACIDS

ALA	Alanine
ARG	Arginine
ASN	Asparagine
ASP	Aspartic Acid
ASX	Asparagine or Aspartic Acid
CYS	Cysteine
CYX	Cysteine or Cystine
GLN	Glutamine
GLU	Glutamic Acid
GLX	Glutamic Acid or Glutamine
GLY	Glycine
HIS	Histidine
ILE	Isoleucine
LEU	Leucine
LYS	Lysine
MET	Methionine
PHE	Phenylalanine
PRO	Proline
SER	Serine
THR	Threonine
TRP	Tryptophan
TYR	Tyrosine
VAL	Valine

NUCLEOSIDES

A	Adenosine
C	Cytidine
G	Guanosine
T	Thymidine
U	Uridine
5MU	5-Methyluridine (Thymidine)

OTHER

ACE	Acetyl
FOR	Formyl
HOH	Water
HEM	Heme
NAD	Nicotinamide Adenine Dinucleotide
NDP	Nicotinamide Adenine Dinucleotide Monophosphate

A.2 Standard Structural Elements (Primitives)

Standard structural elements used in the PRIMITIVES option on the LIMITS menu under MAKE GROUPS (see Section 4.1.6):

STANDARD SET

STD SET	Standard set of elements (all listed except TRACE and LABELS CA)

PROTEINS

TRACE	α -carbons (C_α)
MAIN CHAIN	Peptide backbone or main chain without hydrogens (-N- C_α -C-)
HMAIN	Peptide backbone hydrogens
SIDE CHAIN	Peptide residue side chains without hydrogens
HSIDE	Side chain hydrogens
CARBONYL	Peptide C=O group
LABELS CA	Labels (residue name and sequence number) for α -carbons

NUCLEIC ACIDS

PHOSPHATE	Nucleotide phosphate without hydrogens
HPHOSPHATE	Phosphate hydrogens
SUGAR	Nucleotide sugar without hydrogens
HSUGAR	Sugar hydrogens
BASE	Nucleotide base without hydrogens
HBASE	Base hydrogens

OTHER

METAL	Metals (appear as 6-point stars)
WATER	Waters without hydrogens (appear as tetrahedra)
HWATER	Water with hydrogens
NONSTD	Nonstandard residues (<i>e.g.</i> , heme)
HNONSTD	Nonstandard residue hydrogens

A.3 Standard Structural Descriptors (Selectors)

Standard structural descriptors used in the SELECTORS option on the LIMITS submenu under MAKE GROUPS (see Section 4.1.6):

STRUCTURAL FEATURES (must be defined in the Brookhaven file to be selected)

HELIX	α -helix
SHEET	β -sheet
ACTIVE SITE	Active site
TURN	β -turn

SIDE CHAIN CHARACTERISTICS

POSITIVE	Positively charged residues (ARG, LYS)
NEGATIVE	Negatively charged residues (ASP, GLU)
AROMATIC	Aromatic residues (HIS, PHE, TRP, TYR)
FOBIC	Hydrophobic residues (ILE, LEU, MET, PHE, TRP, TYR, VAL)
FILIC	Hydrophilic residues (ARG, ASP, GLU, HIS, LYS)
AMBIVALENT	Ambivalent residues (ALA, ASN, CYS, GLN, GLY, PRO, SER, THR)

Appendix B

Fragment Libraries

The BUILD operations (Chapter 3) contain builders which are designed for constructing various classes of molecules based on fragments. Each builder is associated with an appropriate fragment library or libraries. The fragment libraries and their corresponding fragment library directories are listed below. The directory names are given in the UNIX¹ operating system format. Lists of the BioDesign General Format (*bgf*) fragment files contained within each of the directories are given in the following sections. *Note:* In the directory specification, *vn* = version number $\times$ 100, *e.g.*, biogv300 for BIOGRAF version number 3.0.

Fragment Library	Logical Name	Directory Specification (UNIX)
Organic fragment	BG_ORG_FRG	/biograf/biogvvn/organic/fragment/
Organic functional	BG_ORG_FUN	/biograf/biogvvn/organic/function/
Solvent	BG_SOLVENT	/biograf/biogvvn/solvent/
Peptide	BG_PEP_L	/biograf/biogvvn/peptide_l/
Peptide	BG_PEP_D	/biograf/biogvvn/peptide_d/
Carbohydrate	BG_CARBO	/biograf/biogvvn/carbohydrate/
Lipid backbone	BG_LIPID_BCK	/biograf/biogvvn/lipid/backbone/
Lipid fatty acids	BG_LIPID_FA	/biograf/biogvvn/lipid/fatty_acid/
Lipid head groups	BG_LIPID_HG	/biograf/biogvvn/lipid/head_group/
DNA nucleotides	BG_DNA_SINGLE	/biograf/biogvvn/dna/single/
DNA nucleotide pairs	BG_DNA_DOUBLE	/biograf/biogvvn/dna/double/
Nucleotide triplets	BG_DNA_TRIPLE	/biograf/biogvvn/dna/triple/

¹UNIX is a trademark of AT&T.

B.1 BG_ORG_FRG

The *bgf* files contained in the BG_ORG_FRG fragment library are listed below. All file names have a *bgf* extension.

Name	Description	Name	Description
acetylene	Acetylene	formic_acid	Formic Acid
ammonia	Ammonia	furan	Furan
anthracene	Anthracene	h3po4	Phosphoric Acid
benzene	Benzene	hcl	Hydrochloric Acid
cyclobutane	Cyclobutane	hcn	Hydrogen Cyanide
cyclohexane	Cyclohexane	hono2	Nitric Acid
cyclopentane	Cyclopentane	indene	Indene
cyclopropane	Cyclopropane	methane	Methane
decalin_cis	<i>cis</i> -Decalin	naphthalene	Naphthalene
decalin_trans	<i>trans</i> -Decalin	phenalene	Phenalene
ethane	Ethane	phenanthrene	Phenanthrene
ethylene	Ethylene	steroid	Steroid
fluorene	Fluorene	tetralin	Tetralin
formaldehyde	Formaldehyde	water	Water

Several of the compounds in the organic fragment library are standard fused polycyclic hydrocarbons. For more information see Robert C. Weast, *CRC Handbook of Chemistry and Physics*, 62nd ed. (Boca Raton, Florida: CRC Press, 1981), p.C-12 - C-17.

B.2 BG_ORG_FUN

The *bgf* files contained in the BG_ORG_FUN functional group library are listed below. All file names have a *bgf* extension.

Name	Description	Name	Description
acetyl	Acetyl	nitro	Nitro
amino	Amino	phenyl	Phenyl
carboxy	Carboxy	phosphino	Phosphino
chloro	Chloro	sulfo	Sulfo
cyano	Cyano	t_butyl	<i>tert</i> -Butyl
ethyl	Ethyl	thiol	Thiol
formyl	Formyl	tosyl	Tosyl
hydroxy	Hydroxy	vinyl	Vinyl
methyl	Methyl		

B.3 BG_SOLVENT

The *bgf* files contained in the BG_SOLVENT fragment library are listed below. All file names have a *bgf* extension.

Name	Description	Name	Description
acetone	Acetone	hexane	Hexane
benzene	Benzene	hexfisopropanol	Hexfluoroisopropanol
cyclohexane	Cyclohexane	methanol	Methanol
cyclopentane	Cyclopentane	nmp	N-methyl-P
dimethylether	Dimethyl Ether	pentane	Pentane
dmf	Dimethyl Formamide	phenol	Phenol
ethanol	Ethanol	toluene	Toluene
ether	Ether	water	Water
formamide	Formamide	water_no_H	Water without hydrogens

B.4 BG_PEP_L, BG_PEP_D

The *bgf* files contained in the BG_PEP_L and BG_PEP_D fragment libraries are listed below. All file names have a *bgf* extension.

Name	Description	Name	Description
ala_l	<i>l</i> -Alanine	leu_l	<i>l</i> -Leucine
arg_l	<i>l</i> -Arginine	lys_l	<i>l</i> -Lysine
asn_l	<i>l</i> -Asparagine	met_l	<i>l</i> -Methionine
asp_l	<i>l</i> -Aspartic Acid	phe_l	<i>l</i> -Phenylalanine
cys_l	<i>l</i> -Cysteine	pro_l	<i>l</i> -Proline
gln_l	<i>l</i> -Glutamine	ser_l	<i>l</i> -Serine
glu_l	<i>l</i> -Glutamic Acid	thr_l	<i>l</i> -Threonine
gly_l	<i>l</i> -Glycine	trp_l	<i>l</i> -Tryptophan
his_l	<i>l</i> -Histidine	tyr_l	<i>l</i> -Tyrosine
ile_l	<i>l</i> -Isoleucine	val_l	<i>l</i> -Valine

Peptide fragment library, L-form (BG_PEP_L)

Name	Description	Name	Description
ala_d	<i>d</i> -Alanine	leu_d	<i>d</i> -Leucine
arg_d	<i>d</i> -Arginine	lys_d	<i>d</i> -Lysine
asn_d	<i>d</i> -Asparagine	met_d	<i>d</i> -Methionine
asp_d	<i>d</i> -Aspartic Acid	phe_d	<i>d</i> -Phenylalanine
cys_d	<i>d</i> -Cysteine	pro_d	<i>d</i> -Proline
gln_d	<i>d</i> -Glutamine	ser_d	<i>d</i> -Serine
glu_d	<i>d</i> -Glutamic Acid	thr_d	<i>d</i> -Threonine
gly_d	<i>d</i> -Glycine	trp_d	<i>d</i> -Tryptophan
his_d	<i>d</i> -Histidine	tyr_d	<i>d</i> -Tyrosine
ile_d	<i>d</i> -Isoleucine	val_d	<i>d</i> -Valine

Peptide fragment library, D-form (BG_PEP_D)

B.5 BG_LIPID

The *bgf* files contained in the BG_LIPID fragment libraries are listed below. All file names have a *bgf* extension.

Name	Description	Name	Description
glycerol	Glycerol	p_sphingosine	Phosphosphingosine
p_glycerol	Glycerol-3-Phosphate	sphingosine	Sphingosine

Lipid Backbone (BG_LIPID_BCK)

Name	Description	Name	Description
arachidate	Arachidate	linolenate	Linolenate
arachidonate	Arachidonate	myristate	Myristate
behenate	Behenate	oleate	Oleate
laurate	Laurate	palmitate	Palmitate
lignocerate	Lignocerate	palmitoleate	Palmitoleate
linoleate	Linoleate	stearate	Stearate

Lipid Fatty Acid (BG_LIPID_FA)

Name	Description	Name	Description
choline	Choline	ethanolamine	Ethanolamine
glycerol	Glycerol	inositol	Inositol

Lipid Head Group (BG_LIPID_HG)

B.6 BG_CARBO

The files contained in the BG_CARBO fragment library are listed below. All file names have a *bgf* extension.

Name	Description	Name	Description
ad_2dribose	α -2-deoxy-D-ribose	bd_2dribose	β -2-deoxy-D-ribose
ad_allose	α -D-allose	bd_allose	β -D-allose
ad_altrose	α -D-altrose	bd_altrose	β -D-altrose
ad_arabinose	α -D-arabinose	bd_arabinose	β -D-arabinose
ad_fructose	α -D-fructose	bd_fructose	β -D-fructose
ad_galactose	α -D-galactose	bd_galactose	β -D-galactose
ad_galactsamn	α -D-galactosamine	bd_galactsamn	β -D-galactosamine
ad_galnac	α -N-acetyl-D-galactosamine	bd_galnac	β -N-acetyl-D-galactosamine
ad_glcamn_nac	α -N-acetyl-D-glucosamine	bd_glcamn_nac	β -N-acetyl-D-glucosamine
ad_glucosamn	α -D-glucosamine	bd_glucosamn	β -D-glucosamine
ad_glucose	α -D-glucose	bd_glucose	β -D-glucose
ad_gulose	α -D-gulose	bd_gulose	β -D-gulose
ad_idose	α -D-idose	bd_idose	β -D-idose
ad_lyxose	α -D-lyxose	bd_lyxose	β -D-lyxose
ad_mannosamn	α -D-mannosamine	bd_mannosamn	β -D-mannosamine
ad_mannose	α -D-mannose	bd_mannose	β -D-mannose
ad_mansa_nac	α -N-acetyl-D-mannosamine	bd_mansa_nac	β -N-acetyl-D-mannosamine
ad_psicose	α -D-psicose	bd_psicose	β -D-psicose
ad_ribose	α -D-ribose	bd_ribose	β -D-ribose
ad_sorbose	α -D-sorbose	bd_sorbose	β -D-sorbose
ad_tagatose	α -D-tagatose	bd_tagatose	β -D-tagatose
ad_talose	α -D-talose	bd_talose	β -D-talose
ad_xylose	α -D-xylose	bd_xylose	β -D-xylose
al_fucose	α -L-fucose	bl_fucose	β -L-fucose
al_rhamnose	α -L-rhamnose	bl_rhamnose	β -L-rhamnose
al_sorbose	α -L-sorbose	bl_sorbose	β -L-sorbose

B.7 BG_DNA_SINGLE

The *bgf* files contained in the BG_DNA_SINGLE fragment library are listed below². These are the standard deoxyribonucleotides for single-stranded chains. All file names have a *bgf* extension.

Name	Description	Name	Description
A_dA	A form - deoxyadenylic acid	B_dA	B form - deoxyadenylic acid
A_dC	A form - deoxycytidylic acid	B_dC	B form - deoxycytidylic acid
A_dG	A form - deoxyguanylic acid	B_dG	B form - deoxyguanylic acid
A_dT	A form - deoxythymidylic acid	B_dT	B form - deoxythymidylic acid

B.8 BG_DNA_DOUBLE

The *bgf* files contained in the BG_DNA_DOUBLE fragment library are listed below³. These are the standard deoxyribonucleotide pairs for double-stranded chains. All file names have a *bgf* extension.

Name	Description
A_dA-dT	A form - deoxyadenylic acid - deoxythymidylic acid pair
A_dC-dG	A form - deoxycytidylic acid - deoxyguanylic acid pair
A_dG-dC	A form - deoxyguanylic acid - deoxycytidylic acid pair
A_dT-dA	A form - deoxythymidylic acid - deoxyadenylic acid pair
B_dA-dT	B form - deoxyadenylic acid - deoxythymidylic acid pair
B_dC-dG	B form - deoxycytidylic acid - deoxyguanylic acid pair
B_dG-dC	B form - deoxyguanylic acid - deoxycytidylic acid pair
B_dT-dA	B form - deoxythymidylic acid - deoxyadenylic acid pair

² Coordinates for the single and double stranded fragments are the same as those used for the old DNA builder and were taken from Arnott and Hukins "Optimized Parameters for A-DNA and B-DNA", *Biochem. and Biophys. Research Comm.*, **47**:6, (1972).

B.9 BG_DNA_TRIPLE

The *bgf* files contained in the BG_DNA_TRIPLE fragment library are listed below³. These are nucleotide triplets for triple-stranded chains. All file names have a *bgf* extension.

Name	Description
I-A-I ⁴	Inosylic acid - adenylic acid - inosylic acid triplet (helix symmetry: 12, turn angle: 30.0°)
U-A-U-11	Uridylic acid - adenylic acid - uridylic acid triplet (helix symmetry: 11)
U-A-U-12	Uridylic acid - adenylic acid - uridylic acid triplet (helix symmetry: 12)
U-dA-U	Uridylic acid - deoxyadenylic acid - uridylic acid triplet
dC-dI-dC-1 ⁵	Deoxycytidylic acid - deoxyinosylic acid - deoxycytidylic acid triplet (2' oxygens on sugars)
dC-dI-dC-2 ⁶	Deoxycytidylic acid - deoxyinosylic acid - deoxycytidylic acid triplet (no 2' oxygens on sugars)
dT-dA-dT	Deoxythymidylic acid - deoxyadenylic acid - deoxythymidylic acid triplet

³ Coordinates for the triple stranded fragments were taken from Arnott *et al.*, "Models of triple stranded polynucleotides with optimized stereochemistry", *Nucleic Acids Research*, 3:10, (1976), with the following caveats (see footnotes 5 and 6).

⁴ The helix symmetry in Arnott *et al.*⁴ for poly(I)•poly(A)•poly(I) is described as 12-fold, however, the "turn angle" given in the corresponding table is 32.7° which indicates 11-fold symmetry (i.e., $360/11 = 32.7$). Since the pitch and residue height for this structure are given as 39.5 Å and 32.9 Å, respectively (which implies 12 units/turn - $39.5/3.29 = 12$), the helix symmetry was taken as 12 and a turn angle of 30.0° was used.

⁵ In Arnott *et al.*⁴, the structure poly(dC)•poly(dI)•poly(dC) has a name that implies 2' deoxy sugars. However, the corresponding table reports coordinates for the 2' oxygens. Hence, BIOGRAF provides both types of triplets, one with and one without the 2' oxygens.

Appendix C

Supplementary Files

The supplementary files supplied with Molecular Simulations software include data files containing the appropriate coordinate, group, and dot files for representative structures, macro files containing example macros, and user subroutines. These files are useful in demonstrating various features of the program or in extending its capabilities.

C.1 Data Files

The supplementary data files listed below contain representative structures which are useful for demonstrating various features of the program. They are located in the *bg_data* directory.

1. MORPHINE

Files to group and display morphine, including the dot surface. Test case for rotations and translations, molecular geometry, dot surface display, and setting colors of groups.

File names: *intro.bgf*, *intro.dot*, *intro.grp*

2. CYTOCHROME C

Files to group and display Cytochrome C. Test case for side chain replacement.

File names: *test.bgf*, *test.grp*

3. CYCLOHEXANE

Files to display cyclohexane. Test case for batch calculations.

File names: *chex.bgf*, *chex.grp*, *chex.bkv*, *chex.dot*

4. CRAMBIN

Files to group and display crambin. Test case for protein modification.

File names: *crambin.bkv*, *crambin.grp*, *crambin.bgf*

5. DNA

Files to group and display DNA with antiviral drug netropsin. Test case for docking.

File names: *dna.bkv, dna.grp, dna.bgf, dna.bgf*

6. THERMOLYSIN

Files to group and display thermolysin with the CLT inhibitor bound to the active site, including a solvent accessible surface. Test case for docking. Also files to group and display a water-accessible dot surface for the active site. Test case for surfacing.

File names: *cltxray.bgf, cltxray.bkv, therm.bgf, therm.grp, thermcav.dot*

7. LEUCINE ENKEPHALIN

Files to group, display and animate a dynamics trajectory with leucine enkephalin. Test case for animation. Van der Waals and solvent accessible surfaces.

File names: *leuenk1.bgf, leuenk1.grp, leuenk1.trj, leuenk1.sol, leuenk1.vdw*

8. ZEOLITE

Files to group and display para-xylene in ZSM-5 zeolite. Test case for constrained minimization and trajectory options.

File names: *zsm54.bkv, zsm5.grp, paraxy.bgf, zsm5.bgf, paraxy.bkv*

9. COBALT PHENANTHRENE

File for cobalt phenanthrene complex. Example of Cambridge file format.

File names: *cobalt_phen.cmb*

10. ETHANE

Files for ethane. Example of Chemlab CHM and MOL file formats.

File names: *ethane.chm, ethane.mol*

11. LA2CUO4

File for La₂CuO₄ unit cell. Example of unit cell file format.

File names: *la2cuo4.cel*

12. SOCCERBALL

File for C₆₀ cage structure.

File names: *soccerball.bgf*

13. SHUTTLE

File for space shuttle. Example of general vector list.

File names: *shuttle.vec*

14. POLYETHYLENE

File for fifty unit polymer of polyethylene. Example of polymer strand.

File names: *peth.bgf*

15. NYLON-6

File containing unit cell and extended cell for crystalline nylon-6. Example of periodic polymer.

File names: *nylon6.bgf*

16. NACL

File containing unit cell and extended cell for NaCl crystal. Example of periodic ionic crystal.

File names: *nacl.bgf*

17. DIAMOND

File containing unit cell and extended cell for diamond crystal. Example of periodic covalent crystal.

File names: *diamond.bgf*

18. DEFAULTS

File containing all the standard settings for all the defaults.

File names: *sample.def*

C.2 Macro Files

The supplementary macro files listed below contain example macros which illustrate how to write macros, demonstrate various features of the program, or extend its capabilities. They are located in the *bg_macro* directory. A brief description of the function of each macro is given below. For a more complete description, see the comments within the appropriate macro file.

1. *Energy_out.macro* - Reads a *bgf* file and writes out its energy to a file on disk.
2. *Replace.macro* - Replaces residues in a peptide with those listed in an accompanying data file.
3. *Poly_build.macro* - Builds a polymer by adding one monomer unit at a time, minimizing and running dynamics after each addition.
4. *Dihedral.macro* - Counts the total number of *trans* and *gauche* dihedrals in a polymer chain.
5. *Run_extract.macro* - Shows a simple way in which one might call the macro, *extract.macro* (see below).

6. ***Extract.macro*** - A macro subroutine that extracts conformations from a trajectory file.
7. ***Run_match.macro*** - Matches a series of peptides by calling the subroutine macro, *match.macro* (see below).
8. ***Match.macro*** - A macro subroutine that matches peptide residues by atom pairs.
9. ***Del_methyl_h.macro*** - Deletes the hydrogens on only the methyl groups on a polypropylene molecule and changes the methyl carbon atom types to C_33.
10. ***Show_res_label.macro*** - Uses keyboard commands to page through and display the residue labels in a peptide.

C.3 User Files

The supplementary user files listed below contain sample subroutines which illustrate the features of the user programming interface or extend the program's capabilities. These subroutines must be compiled and relinked with the program. They are located in the */biograf/biogv220/source* directory. A brief description of the function of each of the subroutines is given below. For a more complete description, see the comments within the appropriate user file.

1. ***User1.f*** - Changes the background color of the molecule display area on the IRIS and RS6000.
2. ***Edrusr.f*, *User5.f*, *Usermen1.f*** - These three subroutines together provide an example of how to add your own energy terms to the energy calculations. *Edrusr.f* calculates a spherical harmonic repulsive shell around a system of atoms at an adjustable radius. *User5.f* allows one to change parameter values used by *edrusr.f*. *Usermen1.f* generates the user menu, USMN001, and its submenu, USMN002, for use with *edrusr.f*.
3. ***User4.f*, *Usermen2.f*, *Usrda.f*, *Usubmn.f*** - These four subroutines together provide an illustration of the code used to perform some of the operations that are accessible through the user programming interface. *User4.f* displays and runs a user menu that performs a number of menu functions used within the program. *Usermen2.f* defines the user menu and submenu that are used with *user4.f*. *Usrda.f* and *usubmn.f* are very short subroutines that are called by *user4.f*.

Appendix D

File Formats

Some of the file formats used in the program are described below. These include the format for Brookhaven (*bkv*) files, the format for BioDesign General Format (*bgrf*) coordinate files, as well as the formats for trajectory files, constraints files, group files, plot files, the defaults file, Brookhaven connect files, and the atom and residue synonyms file.

D.1 Brookhaven

The program can read coordinate files which are written in the format used by the Brookhaven Protein Data Bank. The user is referred to the *Atomic Coordinate and Bibliographic Entry Format Description* issued by the Protein Data Bank in January 1985 for detailed information on the Brookhaven format. The Brookhaven or *bkv* files written by the program use a small subset of the possible record formats. A *bkv* file consists of 80 character records, each of which have a left-justified keyword of up to 6 characters, which indicates the type of record format. Brookhaven files written by the program use 6 different record formats: HEADER, REMARK, ATOM, HETATM, CONECT, END.

1. HEADER

FORMAT (A6,1X,A72)

File header.

Col.

1-6 'HEADER'

8-79 Descriptive title for the file.

63-66 'BGRF' is encoded into the title to indicate that the file was written by Molecular Simulations software.

2. REMARK

FORMAT (A6,1X,A72)

File description.

Col.

1-6 'REMARK'

8-79 Descriptive information for the file.

3. ATOM

FORMAT (A6,I5,1X,A4,1X,A3,1X,A1,I4,A1,3X,3F8.3)

Atomic coordinate records for "standard" groups. (See Appendix A.)

Col.

1-4 'ATOM'

7-11 Atom number

13-16 Atom label. (See Section D.3 for a discussion of Brookhaven and Molecular Simulations formats for atom labels.)

18-20 Residue name. (See Section A.1.)

22 Chain designator. (*Warning:* Must be upper case or blank.)

23-26 Residue number.

27 Code for insertions of residues, e.g., 66A, 66B, etc.

31-38 X-coordinate in Å.

39-46 Y-coordinate in Å.

47-54 Z-coordinate in Å.

4. HETATM

FORMAT (A6,I5,1X,A4,1X,A3,1X,A1,I4,A1,3X,3F8.3)

Atomic coordinate records for "nonstandard" groups. (See Appendix A.)

Col.

1-6 'HETATM'

7-54 Same as for ATOM record.

5. CONECT

FORMAT (A6,5I5)

Connectivity records.

Col.

1-6 'CONECT'

7-11 Number of central atom.

12-31 Numbers of atoms bonded to central atom. (*Note:* Uses two records if there are more than four bonds.)

6. END

FORMAT (A3)

End of file marker.

Col.

1-3 'END'

D.2 BioDesign General Format

The format for BioDesign General Format or *bgf* coordinate files is described below.

1. FILTYP

FORMAT (A6,I5)

File header.

1.1 'BIOGRF' or 'XTLGRF' (XTLGRF is the header for periodic files).

1.2 Version number of software in which file format was last changed, currently 300 for version number 3.0.

2. DESCRP

FORMAT ('DESCRP',1X,A8)

A descriptor for this file. This name appears in the FILES menu.

2.1 'DESCRP'

2.2 A short descriptor for the file. This name may contain upper or lower case letters or blanks.

3. REMARK

FORMAT ('REMARK',1X,A)

File description.

3.1 'REMARK'

3.2 Descriptive information for file.

4. FORCEFIELD

FORMAT ('FORCEFIELD',1X,A8)

ASCII name of the force field.

4.1 'FORCEFIELD'

4.2 Name of the force field.

5. PERIOD

FORMAT ('PERIOD',1X,3I1)

Specifies whether periodic in x, y, and z directions.

5.1 'PERIOD'

5.2 Whether periodic in x.

5.3 Whether periodic in y.

5.4 Whether periodic in z.

6. AXES

FORMAT ('AXES',3X,A)

The order of the axes.

- 6.1 'AXES'
- 6.2 The order of the axes (*e.g.*, zyx).
- 7. **SGNAME**
FORMAT ('SGNAME',1X,A8,1X,A8,2I5)
Information regarding the space group.
 - 7.1 'SGNAME'
 - 7.2 Space group name.
 - 7.3 Schoenflies symbol for the space group.
 - 7.4 Number of the space group.
 - 7.5 Choice of unique axis, unit cell origin, or geometry of axes for some space groups.
Include a case index only for space groups that have two cases listed in Appendix B of the *POLYGRAF Reference Manual*.
- 8. **CRYSTX**
FORMAT ('CRYSTX',1X,6F11.5)
Unit cell parameters. Lengths of axes (a,b,c) in Å and angles (α , β , γ) in degrees.
 - 8.1 'CRYSTX'
 - 8.2 Length of a axis.
 - 8.3 Length of b axis.
 - 8.4 Length of c axis.
 - 8.5 Size of angle α .
 - 8.6 Size of angle β .
 - 8.7 Size of angle γ .
- 9. **CELLS**
FORMAT ('CELLS',1X,6I5)
Limits of extension of the cells along the a, b, and c axes.
 - 9.1 'CELLS'
 - 9.2 Min a: number of rows of cells to be added along the a axis (negative).
 - 9.3 Max a: number of rows of cells to be added along the a axis (positive).
 - 9.4 Min b: number of rows of cells to be added along the b axis (negative).
 - 9.5 Max b: number of rows of cells to be added along the b axis (positive).
 - 9.6 Min c: number of rows of cells to be added along the c axis (negative).
 - 9.7 Max c: number of rows of cells to be added along the c axis (positive).

The following five records are only present in periodic files.

10. XTLUNQ

FORMAT ('XTLUNQ,I5,1X,A5,1X,A3,1X,A1,1X,3F9.6,1X,A5,1X,2I3,F8.4,I2,F6.2,F7.1)

These records are only in periodic files containing structures generated with Crystal Builder. They contain all the fractional coordinates and designator information for the unique atoms that are in the unit cell.

- 10.1 'XTLUNQ'
- 10.2 Unique atom number
- 10.3 Unique atom label.
- 10.4 Unique residue name.
- 10.5 Unique chain designator. (*Warning:* Must be upper case or blank.)
- 10.6 Unique residue number.
- 10.7 Fractional X-coordinate.
- 10.8 Fractional Y-coordinate.
- 10.9 Fractional Z-coordinate
- 10.10 Unique atom type. (See Appendix I.)
- 10.11 Absolute value is maximum number of covalent bonds. (*Note:* A negative number indicates that the atom should always be represented with a 6-point star, even if it is covalently bonded.)
- 10.12 Unique number of lone pairs.
- 10.13 Unique atomic charge.
- 10.14 Unique atom symbol code. The symbols are as follows:
 - 1 dots
 - 2 tetrahedron
 - 3 octahedron
 - 4 six-point jack
 - 5 12-point jack
 - 6 18-point jack
- 10.15 Unique atom symbol size in Å.
- 10.16 Unique atom symbol color (hue value in range 0-360).

11. FORMAT ATOM

FORMAT ('FORMAT ATOM',3X,A)

If the format record is present it will govern the format of the following **ATOM** records (if not present the default format described in Section 12 below is assumed).

- 11.1 'FORMAT ATOM'
- 11.2 Format specification for ATOM records.

12. ATOM or HETATM

FORMAT (A6,1X,I5,1X,A5,1X,A3,1X,A1,1X,A5,3F10.5,1X,A5,I3,I2,1X,F8.5)

Atomic coordinate and force field records for "standard" groups. (See Appendix A.)

- 12.1 'ATOM'
- 12.2 Atom number
- 12.3 Atom label.
- 12.4 Residue name.
- 12.5 Chain designator. (*Warning:* Must be upper case or blank.)
- 12.6 Residue number.
- 12.7 X-coordinate in Å.
- 12.8 Y-coordinate in Å.
- 12.9 Z-coordinate in Å.
- 12.10 Atom type. (See Appendix I.)
- 12.11 Absolute value is maximum number of covalent bonds. (*Note:* A negative number indicates that the atom should always be represented with a 6-point star, even if it is covalently bonded.)
- 12.12 Number of lone pairs.
- 12.13 Atomic charge.

13. FORMAT CONECT

FORMAT ('FORMAT CONECT',1X,A)

If the format record is present it will govern the format of the following CONECT records (if not present the default format described in Section 14 below is assumed).

- 13.1 'FORMAT CONECT'
- 13.2 Format specification for CONECT records.

14. CONECT

FORMAT ('CONECT',12I6)

Connectivity records.

- 14.1 'CONECT'
- 14.2 Number of central atom.
- 14.3 Atom numbers of atoms bonded to central atom. *Note:* The program is currently limited to 6 bonds or less.

15. ORDER

FORMAT ('ORDER',1X,12I6)

Bond order records. These records are only present if one or more connections is a multiple bond.

- 15.1 'ORDER'
- 15.2 Number of central atom.
- 15.3 Bond orders for atoms bonded to central atoms. 0 or 1 - single, 2 - double, 3 - triple bond).

The following three record classes are only present in periodic files.

The atom numbers in the **CONNECT** record are the atom numbers of the equivalent unit cell atoms. The following three records specify whether the central atom is bonded to an atom in the same cell or in a neighboring cell.

16. DISPX

FORMAT ('DISPX',1X,12I6)

Cell displacement records. These records are only present if there are one or more connections to a neighboring cell.

16.1 'DISPX'

16.2 Number of central atom.

16.3 X difference in cells of the two connected atoms. Possible values are -1,0,1.

17. DISPY

FORMAT ('DISPY',1X,12I6)

Cell displacement records. These records are only present if there are one or more connections to a neighboring cell.

17.1 'DISPY'

17.2 Number of central atom.

17.3 Y difference in cells of the two connected atoms. Possible values are -1,0,1.

18. DISPZ

FORMAT ('DISPZ',1X,12I6)

Cell displacement records. These records are only present if there are one or more connections to a neighboring cell.

18.1 'DISPZ'

18.2 Number of central atom.

18.3 Z difference in cells of the two connected atoms. Possible values are -1,0,1.

19. END

FORMAT (A3)

End of file marker.

19.1 'END'

D.3 Atom Labels

Brookhaven-formatted files employ a 4-character atom label, while BioDesign General Format files employ a 5-character atom type which is also the default for atom labels. The formats for the two atom labels are presented here in Tables D.1 and D.2.

Char.	Meaning
1-2	Chemical Symbol - right justified
3	Remoteness indicator (alphabetic)
4	Branch designator (numeric)

Table D.1: Brookhaven atom label format

Char.	Meaning
1-2	Chemical Symbol - left justified
3	Hybridization (3-tetrahedral, 2-trigonal, 1-linear)
4	Number of bound hydrogens
5	Alternate parameters indicator (alphabetic)

Table D.2: Molecular Simulations atom label format

D.4 Standard File Name Extensions

The following file name extensions are used for the standard files that come with Molecular Simulations software.

bgf	-	BioDesign General Format coordinate file
bgs	-	BIG STRAIN coordinate file
bkv	-	Brookhaven Protein Data Bank coordinate file
cel	-	Molecular Simulations Crystal Builder or unit cell file
chm	-	CHEMLAB CHM coordinate file
cmb	-	Cambridge crystallographic data base "FDAT" file
con	-	Molecular Simulations constraints file
cnv	-	Molecular Simulations conversion file
csf	-	UNIX command procedure file
cssr	-	Cambridge Structure Search and Retrieval coordinate file
def	-	Molecular Simulations defaults file
dot	-	Molecular Simulations dotted-surface file
exe	-	Molecular Simulations executable image file
grp	-	Molecular Simulations group information file
macro	-	Molecular Simulations macro command file
mcm	-	Macromodel coordinate file
mle	-	Moledit coordinate file
mol	-	MACCS MOLFILE coordinate file
mpc	-	MOPAC coordinate file
par	-	Molecular Simulations force-field parameter file
plt	-	Molecular Simulations intermediate plot file
plot	-	Molecular Simulations plot file
rot	-	Molecular Simulations rotation/translation transformation file
trj	-	Molecular Simulations trajectory file
vec	-	General vector file
xtl	-	Molecular Simulations periodic coordinate file (identical to periodic bgf files)

D.5 Trajectory Files

DISCLAIMER: the format of trajectory files is not finalized and will change. This information is provided for users who wish to write their own applications which need to read the information contained in trajectory files. Please be aware that Molecular Simulations cannot warrant the proper functioning of its software with a trajectory file written by any software package which was not supplied by Molecular Simulations, Inc.. Some parameters in trajectory files are reserved for future use and may be supported in future releases of the program.

Molecular Simulations trajectory files are unformatted sequential files written using the standard FORTRAN library function WRITE supplied with the UNIX operating system. The variables appearing in each record of a trajectory file are described below.

All variables which are "averages" are evaluated since the previous snapshot was written.

Unless otherwise noted, all REALs, INTEGERs, and LOGICALs are declared as such. Therefore, they are the default size.

The file format documented here is for version 2.0 and later.

1. CONTROL INFORMATION

Variables: HDR, (ICNTRL(I), I=1,20)

Data types: character*4, integer(20)

Description:

HDR character*4

Has one of the following values, depending on the type of trajectory in the trajectory file:

'MDTR'	molecular dynamics
'SETR'	potential energy surface from a conformational search
'QUTR'	quenched <i>or</i> annealed molecular dynamics
'CMTR'	constrained minimization

TAKE NOTE OF THIS VALUE: some of the variables written to the trajectory file are meaningful only for certain types of trajectories.

ICNTRL(I) integer(20)

At present, only ICNTRL(1) is being used:

ICNTRL(1)	version number. This integer value is 100 times the version number of the software in which the format of the trajectory file was last modified. Currently this is 220. For prior trajectory file formats, please consult the appropriate reference manual.

2. DESCRIPTIVE TITLE FROM ENERGY EXPRESSION FILE

Variables: NEEXTI, (EEXTIC(I), I=1,NEEXTI)

Data types: integer, character*80(10)

Description:

NEEXTI integer

Number of title records (maximum 10)

EEXTIC(I) character*80(10)

Title records from the energy expression file

3. INFORMATION FOR PERIODIC TRAJECTORIES

Variables: PERIOD, MOLXTL, LCAON, DEFCEL, LRES1, LNOSE, LRES2,
LTMPDAMP

Data types: logical, logical, logical, logical, logical, logical, logical, logical

Description:

PERIOD logical

True if periodic boundary conditions were on when trajectory was written (toggle set to PBC ON).

MOLXTL logical

True if periodic structure is being treated as a molecular crystal.

LCAON logical

True if canonical dynamics had been performed when trajectory was written.

DEFCEL logical

True if unit cell was also allowed to move when energy calculations were performed.

LRES1 logical. Must be set to false.

Reserved for future use.

LNOSE logical

True if Nosé formulation of canonical dynamics is being used.

LRES2 logical

Reserved for future use. Must be set to false.

LTMPDAMP logical

True if temperature damping form of molecular dynamics is being used.

4. INFORMATION FOR ATOMS AND FILES

Variables: NFLUSD, (MVATPF(I), I=1,NFLUSD), (NATPFU(I), I=1,NFLUSD),
(DECUSD(I), I=1,NFLUSD)

Data types: integer, integer(*), integer(*), character*8(*)

Description:

NFLUSD integer
Number of the files containing movable atoms.
MVATPF(I) integer(*)
Number of movable atoms in each file.
NATPFU(I) integer(*)
Total number of atoms in each file used.
DECUSD(I) character*8(*)
Descriptor of each file used.

Variables: TOTMOV, (MVOFST(I), I=1,TOTMOV)

Data types: integer, integer(*)

Description:

TOTMOV integer
Total number of movable atoms (*i.e.*, the sum of MVATPF(I)).
MVOFST(I) integer(*)
Relative atom number of each movable atom with respect to the file it belongs to (*e.g.*,
atom number 1 means the first atom number in the file).

5. NAME OF ENERGY EXPRESSION FILE

Variables: LEEXTI, EEXTIT(1:LEEXTI)

Data types: integer, character*80

Description:

LEEXTI integer
Length of file name
EEXTIT character*80
Complete file name of the energy expression file used in generating this trajectory, including a full directory specification. Under VMS, this also includes a version number.

6. NAME OF PARAMETER FILEVariables: LPARTI, PARTIT(1:LPARTI)Data types: integer, character*80Description:

LPARTI integer

Length of file name

PARTIT character*80

Complete file name of the parameter file used in generating this trajectory, including a full directory specification.

BEGINNING OF LOOP over each snapshot in trajectory.**7. ENERGY AND DYNAMICS CONTROL INFORMATION**Variables: CURTIM, ITSTEP, SYSTEM, AVETEM, DSTEP, FIRSTT, FINALT, E, EB, ET, EP, EI, ENB, EEL, EHB, EC, EU, TINT, TNB, EA, EBA, ETA, EPA, EIA, ENBA, EELA, EHBA, ECA, EUA, TINTA, TNBA, TOTE, TOTKE, TOTEA, TKEA, ICONMP, IMSTEP, LVELWR, ICONFS, ICSTEPData types: real, integer, 5(real), 28(real*8), integer, integer, logical, integer, integerDescription:**NOTE:** *Be sure to take note of the value of HDR in record 1.* Some of the following variables are meaningful only for certain types of trajectories. For example, CURTIM has no meaning in a conformational search trajectory.

CURTIM real

Current time in dynamics run (in ps)

ITSTEP integer

Current total number of steps in dynamics run

SYSTEM, AVETEM real, real

System temperature, average temperature (in K)

DSTEP real

Integration time step for dynamics (in ps)

FIRSTT, FINALT real, real

Initial, final temperature (in K) for the dynamics run

E real*8

Total potential energy for this conformation (in kcal/mole). Total potential energy = total internal energy + total nonbond energy + total user energy + total constraint energy + total NOE energy (see below).

EB, ET, EP, EI, ENB, EEL, EHB, EC, EU 9(real*8)

Totals of individual components of potential energy for this conformation (in kcal/mole): bonds, angles ("thetas"), torsions ("phis"), inversions (often called "improper torsions"), van der Waals ("nonbond"), electrostatic, hydrogen bond, constraint, user energy

TINT, TNB real*8, real*8

Total internal energy, total nonbond energy (in kcal/mole). Total internal energy = bond + angles + torsions + inversions; total nonbond energy = van der Waals + electrostatic + Hbond

EA, EBA, ETA, EPA, EIA, ENBA, EELA, EHBA, ECA, EUA

Average total potential energy; average totals of individual components of potential energy (in kcal/mole).

TINTA, TNBA real*8, real*8

Average total internal energy, average total nonbond energy (in kcal/mole)

TOTE, TOTKE, TOTEA, TKEA 4(real*8)

Grand total energy, total kinetic energy; average grand total energy, average total kinetic energy (in kcal/mole). Grand total energy = total potential energy + total kinetic energy.

ICONMP integer

Current point in a constrained minimization

IMSTEP integer

Total number of steps in minimization

LVELWR logical

True if velocities are being written

ICONFS integer

Number of this conformation in a conformational search

ICSTEP integer

Number of minimization cycles performed at each step in a conformational search or the number of minimization cycles performed at each point of a quenched molecular dynamics trajectory

8. PRESSURE-VOLUME INFORMATION

Variables: PRESSURA, VOLA, PVTOTA, PVKINA, PVPOTA, RADGYRA

Data types: 6(real*8)

Description:

NOTE: These variables only have meaning for periodic calculations.

PRESSURA real*8

Average total pressure (Gpa)

VOLA real*8

Average total volume

PVTOTA real*8

Average PV - total

PVKINA real*8

Average PV - kinetic stress

PVPOTA real*8
 Average PV - potential stress
 RADGYRA real*8
 Average radius of gyration

9. CANONICAL DYNAMICS (included only if LCA_{NON} is true - see item #3)

A) LNOSE = TRUE (the following variables apply if the Nosé formulation is being used)

Variables: SNOSE, SNOSEH, SSDOT, QCANON

Data types: real*8, real*8, real, real

Description:

SNOSE real*8
 Value for the Nose heat bath variable
 SNOSEH real*8
 Half step value for the Nose heat bath variable
 SSDOT real
 Time derivative of the SNOSE variable at full step
 QCANON real
 Mass-like variable for canonical dynamics

B) LNOSE = FALSE (the following variables apply if the Nosé formulation is *NOT* being used)

Variables: SIGNOSE, ZFRICT, ZPRFRICT, QCANON

Data types: real*8, real*8, real, real

Description:

SIGNOSE real*8
 Value for the log of Nose heat bath variable
 ZFRICT real*8
 Value for the thermodynamics friction
 ZPRFRICT real
 Time derivative of the thermodynamics friction
 QCANON real
 Mass-like variable for canonical dynamics

10. DEFORMED CELL (included only if PERIOD is true - see item #3)

Variables: TCELAT, S2R, S2RDOT

Data types: real*8, real(6)

Description:

TCELAT real*8
 Cell temperature

S2R real(6)
Information for transforming scaled to real coordinates
S2RDOT real(6)
Velocities for cell coordinates

11. PERIOD (included only if PERIOD is true - see item #3)

Variables: NATMCEL, STRSA, EXTSTRSA
Data types: integer, real*8(6), real*8

Description:

NATMCEL integer
Number of atoms in the unitcell
STRSA real*8(6)
Average stress values on the cell
EXTSTRSA real*8
Average contribution to energy due to external stresses

12. TEMPERATURE DAMPING

Variables: GAMTMP
Data types: real

Description:

GAMTMP real
Value of the temperature damping constant, $1/(2\tau)$

13. X CARTESIAN COORDINATES

Variables: (X(II), I=1,TOTMOV)
Data types: real(*)

Description:

X(II) real(*)
X-coordinates of moveable atoms.

NOTE: The indices for storing the coordinates are dependent on the number of files currently being used. These indices are calculated from the information held in record 4 (see above). If there is only one file and all atoms are movable, $II = I$.

14. Y CARTESIAN COORDINATES

Variables: (Y(II), I=1,TOTMOV)

Data types: real(*)

Description:

Y(I) real(*)

Y-coordinates of moveable atoms.

15. Z CARTESIAN COORDINATES

Variables: (Z(II), I=1,TOTMOV)

Data types: real(*)

Description:

Z(I) real(*)

Z-coordinates of moveable atoms.

If (LVELWR) then continue; else go to end of loop

16. X CARTESIAN VELOCITIES

Variables: (VELX(II), I=1,TOTMOV)

Data types: real(*)

Description:

VELX(I) real(*)

X-component of the velocities of moveable atoms. Currently, this array is dimensioned no larger than 20000.

17. Y CARTESIAN VELOCITIES

Variables: (VELY(II), I=1,TOTMOV)

Data types: real(*)

Description:

VELY(I) real(*)

Y-component of the velocities of moveable atoms.

18. Z CARTESIAN VELOCITIES

Variables: (VELZ(II), I=1,TOTMOV)

Data types: real(*)

Description:

VELZ(I) real(*)

Z-component of the velocities of moveable atoms.

END OF LOOP over each snapshot in trajectory.

Loop back to Record 7 unless end-of-file encountered.

D.6 Constraints Files

Constraints files store information on the constraints to be imposed during energy calculations. The format for Molecular Simulations constraints files is given below. The first two records must be HEADER and CONTROL records (in that order) for the file to be a valid data file. The last record must be an END record. All other records can be mixed in any order. Up to 1000 DISTANCE constraint records are allowed and limits for all other constraints are 100.

1. HEADER

FORMAT (A10,A60)

File header.

Col.

1-6 'HEADER'

11-21 'CONSTRAINTS'

2. CONTROL

FORMAT (A10,I4I5)

Control information (*e.g.*, version number and force field) under which the constraint data file was written.

Col.

1-7 'CONTROL'

11-15 Version number.

3. REMARK

FORMAT (A10,A60)

Descriptive information. The remarks are echoed to the terminal when a CONSTRAINTS file is read.

Col.

1-6 'REMARK'

11-70 Descriptive information for the file.

4. DISTANCE

FORMAT (A10,2I5,10X,I5,2F10.3)

Information for distance constraints.

Col.

1-8 'DISTANCE'

11-15 Atom I in distance constraint.

16-20 Atom J in distance constraint.

31-35 Integer code for the type of distance constraint (reserved for future use).

36-45 Force constant in (kcal/mol)/Å².

46-55 Distance constraint value in Å.

5. ANGLE

FORMAT (A10,3I5,5X,I5,2F10.3)

Information for angle constraints.

Col.

1-5 'ANGLE'**11-15** Atom I in angle constraint.**16-20** Atom J in angle constraint.**21-25** Atom K in angle constraint.**31-35** Integer code for type of angle constraint (reserved for future use).**36-45** Force constant in (kcal/mol)/rad².**46-55** Angle constraint value in degrees.**6. TORSION**

FORMAT (A10,4I5,I5,2F10.3)

Information for angle constraints.

Col.

1-7 'TORSION'**11-15** Atom I in torsion constraint.**16-20** Atom J in torsion constraint.**21-25** Atom K in torsion constraint.**26-30** Atom L in torsion constraint.**31-35** Integer code for the type of torsion constraint (reserved for future use).**36-45** Inversion constraint value in degrees (planar = 0°).**46-55** Force constant in (kcal/mol)/rad².**7. INVERSION**

FORMAT (A10,4I5,I5,2F10.3)

Information for angle constraints.

Col.

1-9 'INVERSION'**11-15** Atom I in inversion constraint.**16-20** Atom J in inversion constraint.**21-25** Atom K in inversion constraint.**26-30** Atom L in inversion constraint.**31-35** Integer code for type of inversion constraint (reserved for future use).**36-45** Force constant in (kcal/mol)/rad².**46-55** Angle constraint value in degrees (-180° to 180°).**8. END**

FORMAT (A80)

Record indicating end of file.

Col.

1-3 'END'

D.7 Group Files

The format for Molecular Simulations group files is given below. Note that all of the following information is created by picking the options on the MAKE GROUPS menu (see Section 4.1). The first three items (GROUP FILE, VERSION and REMARK) are required for each group file. The rest of the data records (items 4 through 30) are repeated for each group. However, they are only present if this group actually exists (that is, has been setup). Items 25 through 30 are repeated for each range within a group. The final 14 items are only present in type 2 group files. These records contain the specific atom numbers used in constructing the displayed group. Type 2 group files are provided to speed up the process of creating the groups when the group file is read in. Additionally, this format allows the sophisticated user to provide an arbitrary list of atoms to be grouped (from an external program).

1. GROUP FILE

FORMAT (A10)

File header. Must have the following value: 'GROUP FILE'.

2. VERSION

FORMAT (A8,1X,I1)

Version number of software and file type assigned when file was saved (type 1 or type 2).

Col.

1-5 'BIOGV'

6-8 Version number of software in which file format was last changed, currently 300 for version number 3.0.

9-10 File type assigned when group file was saved (1 or 2).

3. REMARK

FORMAT (A6,1X,A72)

File description.

Col.

1-6 'REMARK'

8-79 Descriptive information for the file.

4 - 30: The following set of records (items 4 through 30) are repeated once for each group.

4. GROUP

FORMAT (A12,':',1X,I2,1X,A10)

Group number and name.

Col.

1-14 'GROUP : '

15-16 Group number.

18-27 Group name.

5. GROUP TYPE

FORMAT (A12,':',1X,A4)

Type of group.

Col.

1-14 'GROUP TYPE : '**15-18** Type of group (*e.g.*, PRIM for atoms).**6. RENDER TYPE**

FORMAT (A12,':',1X,I2)

Type of rendering to be performed

Col.

1-14 'RENDER TYPE : '**15-16** Type of rendering to be performed (0 means vectors).**7. BALL SCALE**

FORMAT (A12,':',1X,E15.7)

Size parameter for balls in BALL AND STICK model.

Col.

1-14 'BALL SCALE : '**15-29** Size parameter for balls in BALL AND STICK model.**8. STICK SCALE**

FORMAT (A12,':',1X,E15.7)

Size parameter for sticks in BALL AND STICK model.

Col.

1-14 'STICK SCALE : '**15-29** Size parameter for sticks in BALL AND STICK model.**9. SHADE TYPE**

FORMAT (A12,':',1X,I2)

Type of shading to be performed.

Col.

1-14 'SHADE TYPE : '**15-16** Type of shading to be performed.**10. TESSELATN**

FORMAT (A12,':',1X,E15.7)

Degree of tessellation of spheres (number of triangles).

Col.

1-14 'TESSELATN : '**15-29** Degree of tessellation of spheres (number of triangles).

11. SPEC INTEN

FORMAT (A12,':',1X,E15.7)

Specular intensity (intensity of the mirror like part of the light source reflection).

Col.

1-14 'SPEC INTEN : '**15-29** Specular intensity (intensity of the mirror like part of the light source reflection).**12. SPEC FACTR**

FORMAT (A12,':',1X,E15.7)

Specular factor (roughness of the surface).

Col.

1-14 'SPEC FACTR : '**15-29** Specular factor (roughness of the surface).**13. HALF BOND**

FORMAT (A12,':',1X,L2)

Logical toggle that determines whether group is half bonded.

Col.

1-14 'HALF BOND : '**15-16** Logical toggle that determines whether group is half bonded.**14. COLOR**

FORMAT (A12,':',1X,2(1X,F10.5))

Color used to display the group.

Col.

1-14 'COLOR : '**15-25** Hue (0-360).**26-36** Saturation (0-1).**15. TOGL STATUS**

FORMAT (A12,':',1X,L2)

Logical toggle recording whether group is turned on.

Col.

1-14 'TOGL STATUS : '**15-16** Logical toggle recording whether group is turned on.**16. FULL FILE**

FORMAT (A12,':',1X,L2)

Logical toggle that determines whether to use full file for group.

Col.

1-14 'FULL FILE : '

15-16 Logical toggle that determines whether to use full file for group.

17. UNIT CELL

FORMAT (A12,':',1X,L2)

Logical toggle determining whether to be unit cell.

Col.

1-14 'UNIT CELL : '

15-16 Logical toggle determining whether to be unit cell.

18. FILES

FORMAT (A12,':',1X,I2,10(1X,A1):/(16X,10(1X,A1)))

List of the files that comprise the group.

Col.

1-14 'FILES : '

15-16 Total number of files used.

17-36 List of the files that comprise the group. (*Note:* Uses multiple records if there are more than 10 files).

19. CHAINS

FORMAT (A12,':',1X,I2,10(1X,A1):/(16X,10(1X,A1)))

List of the chains that comprise the group.

Col.

1-14 'CHAINS : '

15-16 Total number of chains used.

17-36 List of the chains that comprise the group. (*Note:* Uses multiple records if there are more than 10 chains).

20. RESIDUES

FORMAT (A12,':',1X,I2,10(1X,A4):/(16X,10(1X,A4)))

List of the residues that comprise the group.

Col.

1-14 'RESIDUES : '

15-16 Total number of residues used.

17-66 List of the residues that comprise the group. (*Note:* Uses multiple records if there are more than 10 residues).

21. HB PRIM

FORMAT (A12,':',1X,I2,10I3:/(16X,10I3))

Col.

1-14 'HB PRIM : '

- 15-16** Total number of primary groups.
17-46 List of the primary groups. (*Note:* Uses multiple records if there are more than 10 primary groups).

22. HB SCND

FORMAT (A12,':',1X,I2,10I3:/(16X,10I3))

Col.

1-14 'HB SCND : '

15-16 Total number of secondary groups.

17-46 List of the secondary groups. (*Note:* Uses multiple records if there are more than 10 secondary groups).

23. HB DCUT

FORMAT (A12,':',1X,2(1X,E15.7))

Col.

1-14 'HB DCUT : '

15-30 Hydrogen bond distance cutoff.

31-46 Hydrogen bond angle cutoff.

24. # OF RANGES

FORMAT (A12,':',1X,I2)

Number of ranges of primitives/selectors.

Col.

1-14 '# OF RANGES : '

15-16 Number of ranges of primitives/selectors.

25 - 30: The following set of records (items 25 through 30) are repeated once for each range.

25. RANGE

FORMAT (A12,':',1X,I2)

Number of the current range.

Col.

1-14 'RANGE : '

15-16 Number of the current range.

26. START/STOP

FORMAT (A12,':',1X,2(1X,I5))

Starting and stopping residue numbers for current range.

Col.

1-14 'START/STOP : '

15-20 Starting residue number for current range.

21-26 Stopping residue number for current range.

27. DSTNCE GRPS

FORMAT (A12,':',1X,10L2:/14X,10L2)

Logical toggles controlling which groups to compare when grouping by distance.

Col.

1-14 'DSTNCE GRPS : '**15-34** Logical toggles controlling which groups to compare when grouping by distance.
(Note: Uses two records if there are more than 10 groups).**28. DSTNCE CUT**

FORMAT (A12,':',1X,E15.7)

Value of distance cutoff for groups by distance.

Col.

1-14 'DSTNCE CUT : '**15-29** Value of distance cutoff for groups by distance.**29. PRIMITIVES**

FORMAT (A12,':',1X,I2,10(1X,A3):/(16X,10(1X,A3)))

List of primitives used to create group.

Col.

1-14 'PRIMITIVES : '**15-16** Total number of primitives used.**17-56** List of primitives used to create group. (Note: Uses multiple records if there are more than 10 primitives).**30. SELECTORS**

FORMAT (A12,':',1X,I2,10(1X,A3):/(16X,10(1X,A3)))

List of selectors used to create group.

Col.

1-14 'SELECTORS : '**15-16** Total number of selectors used.**17-56** List of selectors used to create group. (Note: Uses multiple records if there are more than 10 selectors).

The following items (31-44) are only present in type 2 group files.

31. NPE

FORMAT (A12,':',1X,I5)

Total number of picture elements (lists of vectors to be displayed).

Col.

1-14 'NPE : '

15-19 Total number of picture elements (lists of vectors to be displayed).

32. PENAME

FORMAT (A12,':',1X,5(A12,1X):/(16X,5(A12,1X)))

Picture element name.

Col.

1-14 'PENAME : '

15-79 List of picture element names that indicates which vectors belong to each group. (Note: Uses multiple records if there are more than 10 picture element names).

33. IPEBI

FORMAT (A12,':',1X,10I6:/(16X,10I6))

Position of first vector in picture element list for each group.

Col.

1-14 'IPEBI : '

15-74 Position of first vector in picture element list for each group. (Note: Uses multiple records if there are more than 10 picture element names).

34. IPELST

FORMAT (A12,':',1X,10I6:/(16X,10I6))

Picture element list.

Col.

1-14 'IPELST : '

15-74 List of pairs of atoms between which vectors are drawn. (Note: Uses multiple records if there are more than 5 vectors).

35. NPS

FORMAT (A12,':',1X,I5)

Total number of picture symbols (lists of symbols to be displayed).

Col.

1-14 'NPS : '

15-19 Total number of picture symbols (lists of symbols to be displayed).

36. PSNAME

FORMAT (A12,':',1X,5(A12,1X):/(16X,5(A12,1X)))

Picture symbol name.

Col.

1-14 'PSNAME : '

15-79 List of picture symbol names that indicates which symbols belong to each group. (Note: Uses multiple records if there are more than 10 symbol names).

37. IPSLST

FORMAT (A12,':',1X,10I6:/(16X,10I6))

Picture symbol list.

Col.

1-14 'IPSLST : '**15-74** List of atoms where symbols are drawn. (*Note:* Uses multiple records if there are more than 10 symbols).**38. IPSSMB**

FORMAT (A12,':',1X,10I6:/(16X,10I6))

List of symbol codes.

Col.

1-14 'IPSSMB : '**15-74** List of symbol codes. See Section D.2.10.14 above for a list of the symbols. (*Note:* Uses multiple records if there are more than 10 symbols).**39. PSCOLR**

FORMAT (A12,':',1X,10I6:/(16X,10I6))

List of symbol colors. (*Note:* Uses multiple records if there are more than 10 symbols).

Col.

1-14 'PSCOLR : '**15-74** List of symbol colors. The numbers correspond to the following colors:

1	Red
2	Yellow
3	Green
4	Cyan
5	Blue
6	Purple
7	Pink
8	Orange
9	Chartreuse
10	Aqua
11	Light Blue
12	Magenta

40. PSSCAL

FORMAT (A12,':',1X,6F10.5:/(14X,6F10.5))

List of symbol sizes. (*Note:* Uses multiple records if there are more than 6 symbols).

Col.

1-14 'PSCOLR : '**15-74** List of symbol sizes in Å.

41. NPL

FORMAT (A12,':',1X,I5)

Total number of picture labels (lists of labels to be displayed).

Col.

1-14 'NPL : '**15-19** Total number of picture labels (lists of labels to be displayed).**42. PLNAME**

FORMAT (A12,':',1X,5(A12,1X):/(16X,5(A12,1X)))

Picture label names. (*Note:* Uses multiple records if there are more than 5 label names).

Col.

1-14 'PLNAME : '**15-79** List of picture label names that indicates which labels belong to each group.**43. IPLLST**

FORMAT (A12,':',1X,10I6:/(16X,10I6))

Picture label list. (*Note:* Uses multiple records if there are more than 10 labels).

Col.

1-14 'IPLLST : '**15-74** List of atoms where labels are drawn.**44. PLABLS**

FORMAT (A12,':',1X,5(A12,1X):/(16X,5(A12,1X)))

Picture labels. (*Note:* Uses multiple records if there are more than 5 labels).

Col.

1-14 'PLABLS : '**15-79** Picture labels.

D.8 Plot Files

The format for plot files will be described at a later time.

D.9 Defaults File

The defaults file contains all the defaults which are used by the program. All the variables which can currently be reset are listed below with their default values; when appropriate, the other options or acceptable range of values is included in the description. If you wish to override any of these defaults, you should create a new file and eliminate any variables which are not changed (to speed up execution time). This is a free format ASCII file. Comments can be interspersed through this file by preceding with exclamation points. A line ending with a minus sign (-) is continued on the next line. Strings containing blanks or ending in (-) may be surrounded with quotes (").

D.9.1 Graphics Parameters

A) For Silicon Graphics Workstations

L_SGI	NO	Don't run graphics calls for Silicon Graphics workstations (license dependent)
XORIGIN	321	Lower left X position (in pixels) of main window (0 - 1279)
YORIGIN	48	Lower left Y position (in pixels) of main window (0 - 1023)
XSIZE	952	X size of the main window (in pixels) excluding border (1 - 1280)
YSIZE	952	Y size of the main window (in pixels) excluding border (1 - 1024)
L_WINPUT	YES	Accept the keyboard input only when the mouse is in the main window
L_STEREO	NO	Don't put the program into stereo monitor display mode

B) For IBM RISC System/6000

L_RS6000	NO	Don't run graphics calls for IBM RISC System/6000 workstations (license dependent)
XORIGIN	321	Lower left X position (in pixels) of main window (0 - 1279)
YORIGIN	48	Lower left Y position (in pixels) of main window (0 - 1023)
XSIZE	952	X size of the main window (in pixels) excluding border (1 - 1280)
YSIZE	952	Y size of the main window (in pixels) excluding border (1 - 1024)
L_WINPUT	YES	Accept the keyboard input only when the mouse is in the main window
L_STEREO	NO	Don't put the program into stereo monitor display mode

C) For Stardent Workstations

L_DORE	NO	Don't run graphics calls for Stardent Titan workstations (license dependent)

L_WINPUT	YES	Accept the keyboard input only when the mouse is in the main window
L_STEREO	NO	Don't put the program into stereo monitor display mode
MP_DORE	0	Number of processors for graphics (0 = don't multiprocessing, 1 = not recommended, ≥ 2 = multiprocessing)

D.9.2 General Defaults

PROGRAM		Sets which program is run (BIOGRAF, POLYGRAF, NMRGRAF, Polaris). Default depends on authorization.
L_CMND	NO	Don't start the program in command mode
L_LOGFILE	YES	Log the session in <i>logfile.macro</i>
LBUF_SIZE	1	Number of lines in <i>logfile.macro</i> output buffer (1-5)
SEED	12345	Seed for all random number generation. Specify 0 to generate a pseudo-random seed. Specify any positive integer to generate the same random number sequence each time.
STEREO_MODE	PROXIMAL	Stereo display mode (PROXIMAL, DISTAL)
STRO_ANGL	2.850	Angle used to display images in side by side stereo viewing
STRO_SPRN	0.010	X separation distance between the two stereo images (given in terms of display area coordinates where the left and right edges of the molecule display area are -1.0 and +1.0 respectively)
STRO_OFST	0.000	X offset distance from outer edge of each stereo image to inner edge of the screen (given in terms of display area coordinates where the left and right edges of the molecule display area are -1.0 and +1.0 respectively)
L_AUTOEXT	YES	Add file name extensions for writes as required

D.9.3 Energy Parameters

NSTEP	50	Number of steps
MIN_TYPE	CG	Minimization type (CG for conjugate gradient, SD for steepest descents, FP for Fletcher-Powell)
INFO_LVL	1	Information level (0 - 3)
E_CUTOFF	0.001	Energy convergence criterion for minimizations
PURG_COR	YES	Purge coordinate files in minimization
TIME_STP	0.002	Time step (in picoseconds) for dynamics
INIT_TEM	300.0	Initial temperature (in °K)
FINL_TEM	300.0	Final temperature (in °K)
WRIT_VEL	NO	Do not write velocities to <i>.trj</i> file
NQSTEP	10	Number of quenching steps for molecular dynamics
E_CYCLE	20	Energy cycle in constrained minimization
NPOINTS	10	Number of points in a constrained minimization
DIR_X	1.0	X component of direction vector (0 - 1.0)
DIR_Y	0.0	Y component of direction vector (0 - 1.0)
DIR_Z	0.0	Z component of direction vector (0 - 1.0)

INCREMNT	0.5	Increment in a constrained minimization (Å)
CON_TYPE	PLANE	Constraint type (PLANE, LINE, POINT)
TRAN_TEM	1000.0	Translational temperature
WRIT_PAR	NO	Do not write out new parameter file
SCL_B14	1.0	Scaling factor for 1-4 nonbond interactions (0 - 1.0)
NRUNCT	10 (Titan) 2 (Iris)	Nonbond sequential run length cutoff for vector operations (if < cutoff, scalar operations are used)

D.9.4 Energy Switches

L_BONDS	YES	Include bonds in energy calculation
L_ANGLES	YES	Include angles in energy calculation
L_TORSIN	YES	Include torsions in energy calculation
L_INVRSN	YES	Include inversions in energy calculation
L_NBONDS	YES	Include nonbonds in energy calculation
L_HBONDS	YES	Include hydrogen bonds in energy calculation
L_USER_E	NO	Don't include user energy in energy calculation
L_CONSTR	NO	Don't include harmonic constraints in energy calculation
L_ALLTOR	YES	Calculate and normalize all torsion terms for each torsion; NO = do so only for the first nonzero torsion term
L_ALLINV	YES	Calculate and normalize all inversion terms for each inversion center; NO = do so only for the first nonzero inversion term for each inversion center

D.9.5 Dynamics Switches

L_ANNEAL	NO	Don't run annealed dynamics
L_MAXFOR	YES	Check for maximum force
WRT_TRJ	NO	Don't write the trajectory file

D.9.6 Nonbond Switches

L_NBEXCB	YES	Exclude 1-2 and 1-3 interactions from nonbond lists
L_NBEXCT	YES	Exclude 1-3 interactions from nonbond lists

D.9.7 Nonbond Variables

NB_ON	8.0	Energy completely on below this distance (in Å)
NB_OFF	8.5	Energy off above this distance (in Å)
NB_SEL	9.0	All nonbonds selected within this distance (in Å)
EPSILON	1.0	Dielectric constant
L_R_EPS	YES	Dielectric = R*EPS

D.9.8 Hydrogen Bond Variables

HB_ON	4.0	Energy completely on below this distance (in Å)

HB_OFF	4.5	Energy off above this distance (in Å)
HB_SEL	5.0	All hydrogen bonds selected within this distance (in Å)
HBA_ON	65.0	Energy completely on below this angle (0 - 360.0)
HBA_OFF	75.0	Energy off above this angle (0 - 360.0)
HBA_SEL	90.0	All hydrogen bonds selected within this angle (0 - 360.0)

D.9.9 Miscellaneous Energy Variables

A) For Constraints

K_DIST	10000.0	Distance force constant for next constraints
K_TOR	1000.0	Torsion force constant for next constraints
K_ANG	1000.0	Angle force constant for next constraints
K_INV	1000.0	Inversion force constant for next constraints

B) For Minimization

MAX_STEP	1.5	Maximum displacement size during the minimization (in Å)
GRD_CVRG	0.1	Criterion for convergence (kcal/mol)/Å

C) For Dynamics

ANEL_INC	10.0	Temperature increment for each step of annealing (in °K)
DRAW_FRQ	1	The frequency per step with which structure is redrawn or updated on PS300 display. Also used for minimization.
T_CYCLE	0.1	Interval between updating NBond and HBond lists (picoseconds)
HT_FRQ	0.1	Time interval between temperature equilibrations (picoseconds)
WTRJ_FRQ	0.1	Time interval between writes to trajectory file (picoseconds)
ANEL_FRQ	0.02	Interval between annealing temperature adjustments (picoseconds)
MAX_FORC	200.0	Maximum force level allowed in dynamics (kcal/mol)
TIM_DYN	0.50	Time duration (picoseconds)
ANEAL_CY	1	Number of cycles of annealed dynamics

D) For Temperature

HEAT_INC	20.0	Heating increment after each equilibration cycle in dynamics (in °K)
TEM_WIND	50.0	Difference allowed between desired temperature and actual temperature in dynamics (in °K)

E) For Animation

COL_INC	5.0	Increment on the color wheel used in color coding conformations by sequence number or energy
E_DIFF	10.0	Sets the scale for color coding conformations by energy

F) For Constrained Minimization

NX_STEP	100	Number of steps of minimization at each point

G) For Search

IC_STEP	0	Number of steps of minimization at each point

H) For Impulse Dynamics

NS_STEP	100	Number of steps in dynamics trajectory

I) For Surfaces

DOT_DENS	8.0	Density of dots for calculating surface
PROB_RAD	0.0	Probe radius during surface calculations (Å)
MAX_CHRG	0.5	Scale factor used for defining property of dots by charge (0 -1.0)
GRS_CHRG	0.2	Cutoff value used for defining property of dots by gross charge -,0,+ (electrons)
MAX_ESP	100.0	Maximum electrostatic potential (electrons/Å)
VDW_SCL	1.0	van der Waals scale radius
TOT_COLR	20	Total number of colors to be displayed (machine-dependent)

J) For Connect Functions

CN_RMIN	0.8	Minimum allowed bond distance (Å)
CN_RMAX	2.0	Maximum allowed bond distance (Å)
CN_SCAL	0.55	Scale factor for radius

D.9.11 Energy and Force Analysis Variables

DIST_NBCT	10.0	Nonbond cutoff distance used for display purposes (Å)
VDW_ENRCT	-9999.99	VDW energy cutoff used during analysis (kcal/mol)
ELC_ENRCT	-9999.99	Electrostatic energy cutoff used during analysis (kcal/mol)
SCAL_FAC	1.0	Scale factor for displaying force vectors
L_DSPGEOM	YES	Display bond distance or angle during analysis
L_DSPENR	YES	Display energy values during analysis
L_DSPFR	YES	Display force values during analysis

L_EAF14	YES	Include 1-4 nonbonds during analysis
L_FIXFRC	NO	Don't include fixed atoms during force analysis

D.9.12 Contact Check Analysis Variables and Switches

L_CON14	NO	Don't include 1-4 nonbonds during contact analysis
L_DIS	NO	Don't display distance values during analysis
L_ENR	NO	Don't display energy values during analysis
VDWCT	0.0	VDW energy cutoff used during analysis (kcal/mol) (≥ 0)

To Read from another File

READ (DIR PATH)MYDEFAULTS.DEF Add defaults from another file (you must include the directory path specifications with the file name; see Chapter 1 for information on entering file names)

D.10 Brookhaven Connect Files

Molecular Simulations software can generate connectivity information for residues in the Brookhaven Protein Data Bank. The Brookhaven residue file does not contain any information about hydrogens (deuteriums) present in common ribonucleotides (*e.g.*, A,G,C,U,T) and about non standard groups (*e.g.*, HEM, NAD, NADP). Consequently no implicit connectivity information will be generated about these atoms. This can be fixed by checking the format for representing these hydrogens in a Brookhaven Data Bank file (neutron diffraction or protein data) and putting the data in this connect file.

Hydrogen representation in Kollman's format has been added for each residue (to read Kollman's files properly). In addition, several non-standard residues (used frequently in Kollman's data files) such as QNX, NCY *etc.* have been added, as well as several histidines (HIS, HID, HIP, HIE, HSD, HSP). Different terminal groups are also included. The format for the connect file is described below; "*" in the first column is used for comment lines.

1. RESIDUE

FORMAT (A7,2X,A3)

Residue name.

Col.

1-7 'RESIDUE'

10-12 Residue name

2. Connectivity Data

FORMAT (A5,8(A5,I1,1X))

Connectivity data for a given residue.

The maximum number of connected atoms allowed is six. "#" is used as a wild card; it will match blank or any number from 1 to 9. "-" and "+" in atom names indicate atom of previous or forward residue in peptide chain to which that atom should be connected (if any).

Col.

1-5 Central atom

6-10 Connected atom

11 Bond order of connected atom (0 = single or no bond, 2 = double, 3 = triple)

13-17 Connected atom

18 Bond order of connected atom

20-24 Connected atom

25 Bond order of connected atom

27-31 Connected atom

32 Bond order of connected atom

33-37 Connected atom

38 Bond order of connected atom

40-44 Connected atom

45 Bond order of connected atom

D.11 Atom and Residue Synonyms File

The synonyms file contains the synonyms for residue names and atom labels.

1. SYNONYMS

FORMAT (A8)

File header. Must have the following value: 'SYNONYMS'.

2. CONTROL

FORMAT (A10,14I5)

Control information under which the constraint data file was written.

Col.

1-7 'CONTROL'

11-15 Version number of Molecular Simulations software in which the file format was last changed, currently 207 for version number 2.07.

16-80 Reserved for future use.

3. REMARK

FORMAT (A6,1X,A72)

Descriptive information.

Col.

1-6 'REMARK'

8-79 Descriptive information.

4. RESIDUE NAMES

FORMAT (A13)

Keyword designating the beginning of the synonyms for residue names. Must have the following value: 'RESIDUE NAMES'.

Record 5 may appear any number of times.

5. RESIDUE NAME SYNONYM

FORMAT (A3,'/',n(A3,'/'))

Standard residue name followed by synonyms. At least one synonym must be specified.

Col.

1-3 Standard residue name.

4 Field delimiter: '/'.

5-7 Synonym for standard residue name.

8 Field delimiter: '/'.

optional:

9-11 Synonym for standard residue name.

12 Field delimiter: '/'.

⋮

6. ATOM LABELS

FORMAT (A11)

Keyword designating the beginning of the synonyms for atom labels. Must have the following value: 'ATOM LABELS'.

Record 7 may appear any number of times.

7. FOR RESIDUE

FORMAT (13X,A3)

Keyword designating start of atom label records for each residue. Must have the following value: 'FOR RESIDUE: '. The three letter name of the residue appears in columns 14-16.

Col.**1-13** 'FOR RESIDUE: '**14-16** Standard residue name.

Record 8 may appear any number of times.

8. ATOM LABEL SYNONYM

FORMAT (n(A5,'/'),) (n has a maximum value of 10)

Standard atom label followed by synonyms.

Col.**1-5** Standard atom label.**6** Field delimiter: '/'.**7-11** Synonym for standard atom label.**12** Field delimiter: '/'.*optional:***13-17** Synonym for standard atom label.**18** Field delimiter: '/'.

⋮

Appendix E

Units

The standard units of measure used by the program are as follows:

- energy - kcal/mol
- length - Angstroms (Å)
- angles - degrees (except in force constants where radians are used)
- mass - atomic mass units (amu; 12.0000 for ^{12}C)
- charge - electrons (+1.0 for a proton)

Convenient conversions are:

$$\text{K}[(\text{kcal/mol})/\text{\AA}^2] = 143.93 \text{ K}(\text{mdyn}/\text{\AA})$$

and for an observed vibrational mode with ν in units of 1000 cm^{-1}

$$\text{K}[(\text{kcal/mol})/\text{\AA}^2] = 84.80 \mu \nu^2,$$

where μ is the effective mass in amu ($^{12}\text{C} = 12.0000$). For a diatomic molecule with masses M_1 and M_2 ,

$$\mu = \frac{M_1 M_2}{M_1 + M_2}$$

Appendix F

Rotations

This appendix describes how the program stores and uses the data generated when one object (A) is docked with a second object (B). The (x, y, z) spatial coordinates of the atoms are stored on disk. The spatial coordinates are converted into graphics display coordinates with a four-by-four rotation-translation (RT) transformation matrix T_D . Rotating the dials (or using the middle or right buttons on the mouse) changes the transformation matrix and thereby translates, rotates and zooms the object on the display. (*Note:* Manipulating the image on the display does not affect the spatial coordinates stored on disk.) Thus, by using the dials or mouse to dock object A with B, the user constructs two transformation matrices T_A and T_B , which transform the spatial coordinates (stored on disk) of A and B into display coordinates for A and B, respectively. During the docking operation, the dials or mouse may be connected to both A and B (ROTATE ALL on) or to only A (ROTATE ALL off). When ROTATE ALL or DOCK is toggled and the dials or mouse are reconnected, the RT-transformation matrix, T_D , generated by the current settings of the dials is used to update the existing transformations, T_A^0 and T_B^0 , for A and B. The T_D matrix is then reset to the unit matrix.

For example, if the dials are currently connected to A but not to B, then the net transformation on A is $T_A^0 T_D$ and the net transformation on B is T_B^0 . If ROTATE ALL is toggled on, a new transformation matrix

$$\bar{T}_A^0 = T_A^0 T_D$$

is constructed and stored as the new T_A^0 . The dials are reset, *i.e.*, T_D is set to the unit matrix. The dials are now connected to both A and B. Rotating the dials generates a new transformation

$$\bar{T}_D$$

so that the net transformation on A is $\bar{T}_A^0 \bar{T}_D$, while the net transformation on B is $T_B^0 \bar{T}_D$.

The spatial coordinates of A and B are not modified until the user picks the MODIFY COORDS option (Section 3.10.2). The MODIFY COORDS operation modifies the coordinates of only object A by applying the differential transformation

$$T_{AB} = T_A T_B T_A^{-1}$$

constructed from the RT-transformation matrices.

The RT-transformations are initialized so that the object defined by the first coordinate file is moved to the center of the display. Toggling RESET DIALS sets the dials or mouse to zero and T_D to the unit matrix. Toggling RESET ALL sets the RT-transformations to unit matrices, the dials or mouse to zero, and T_D to the unit matrix. See Section 3.10.4.

Appendix G

Force Fields

In the GRAF products (BIOGRAF, POLYGRAF, and NMRgraf) the energy for an arbitrary geometry of an arbitrary molecule is written as a superposition of various two-body, three-body, and four-body interactions. We use a simple analytic expression for each such interaction, allowing rapid evaluation of the forces on all particles. From the forces we can evaluate the optimum geometry (conjugate gradient minimization) or the dynamics of motion (Newton's equations). Such procedures are often referred to as molecular mechanics or molecular dynamics, respectively.

The overall description of an N-body system in terms of a superposition of simpler terms is called a *Force Field*. It is an approximation (the exact description would involve solving Schroedinger's Equation for the electronic wavefunctions at each geometry). Accurate force fields are available for many organic and biological systems and are being developed for inorganic and metallic systems.

The program incorporates several standard force fields (Dreiding-I and Dreiding-II¹, AMBER^{2,3}, MM2^{4,5}, and MMP2⁶; CHARMM⁷ and MM3⁸ are being incorporated) and allows general forms for the energy expressions appropriate for new force fields (particularly for inorganic materials) that are under development.

The potential energy is expressed as a sum of valence (or bonded) interactions, nonbonded interactions, a user energy term, plus constraints.

$$E_{total} = E_{valence} + E_{nonbond} + E_{user} + E_{constraint} \quad (G.1)$$

¹ S. L. Mayo, B. D. Olafson and W. A. Goddard III, *J. Phys. Chem.*, **94** (1990) 8897-8909.

² S. J. Weiner *et al.*, *J. Amer. Chem. Soc.*, **106** (1984) 765-784.

³ S. J. Weiner *et al.*, *J. Comp. Chem.*, **106** (1986) 230-252.

⁴ N. L. Allinger, *J. Amer. Chem. Soc.*, **99** (1977) 8127-8134.

⁵ T. Liljefors, J. C. Tai, S. Li and N. A. Allinger, *J. Comp. Chem.*, **8** (1987) 1051-1056.

⁶ J. T. Sprague, J. C. Tai, Y. Yuh and N. L. Allinger, *J. Comp. Chem.*, **8** (1987) 581-603.

⁷ B. R. Brooks *et al.*, *J. Comp. Chem.*, **106** (1983) 187-217.

⁸ N. L. Allinger, Y. H. Yuh, and J. H. Lii, *J. Amer. Chem. Soc.*, **111** (1989) 8551-8566.

The valence interactions consist of diagonal terms [namely, bond stretching (E_b), bond angle bending (E_a), dihedral angle torsion (E_t), and inversion (E_i)], that are in nearly all force fields of covalent systems plus cross terms that are included in more sophisticated force fields developed to produce accurate vibrational frequencies.

$$E_{valence} = \underbrace{E_{bond} + E_{angle} + E_{torsion} + E_{inversion}}_{diagonal} + \underbrace{E_{ba} + E_{bb'} + E_{aa'} + E_{bt} + E_{at} + E_{UB}}_{cross\ terms} \quad (G.2)$$

The nonbonded interactions consist of van der Waals (E_{vdw}), electrostatic ($E_{Coulomb}$) and hydrogen bond (E_{hb}) terms.

$$E_{nonbond} = E_{vdw} + E_{Coulomb} + E_{hbond} \quad (G.3)$$

Optional constraint terms having the form of bond, angle, torsion, and inversion are also allowed (the atoms involved need not be connected).

$$E_{constraint} = E_b + E_a + E_t + E_i \quad (G.4)$$

In addition E_{user} is for specialized potentials to be added by the user. The various forms of the potentials allowed by the program are described in the following sections. The parameters used in various standard force fields are described in Appendix I.

G.1 Topology and Automatic Energy Expressions

G.1.1 Nonperiodic

Both the graphics and energy routines used in the program are organized in terms of bond connections

$$I \text{ CONS}(I,J), J = \text{MAXCON}$$

which for each atom I lists the atom number for all atoms (J) bonded to I. Such bonds are displayed graphically and are also used to automatically generate the Energy Expression (EEX) for the molecule. The bond orders are currently required to be integer and stored as

$$I \text{ ORDER}(I) = 1, 2, 3$$

(I ORDER= 0 is understood as I ORDER= 1).

These arrays are generated automatically during BUILD and CONNECT. For nonperiodic cases, the .bgf output file contains the I CONS and I ORDER information along with the coordinates (but if there are no multiple bonds to an atom, the IORDER information is not written out).

The information in I CONS is used to generate the energy expression (EEX) each time that SETUP EEX is toggled.

G.1.2 Periodic

For periodic systems the atoms are defined only in the unit cell and are understood to be in equivalent positions for all other cells. There are two classes, depending upon whether there are bonds connecting atoms in different unit cells (MOLXTL = FALSE) or not (MOLXTL = TRUE). MOLXTL systems are described just as are finite systems (only the unit cell is displayed).

A. NonMOLXTL

For nonMOLXTL systems a second array is defined

$$I\ CONX(I,J),\ J = 1, MAXCON.$$

For each atom I in the unit cell, ICONX lists the atoms (J) bonded to I but for bonds to atoms outside the unit cell, it lists the translational equivalent atom within the cell. To completely define the atom J there is a second array

$$I\ TRAN(I,J),\ J = 1, MAXCON$$

which indicates the translation required to move the second atom from the unit cell to its bonded position. An I CONS array is created and used for the display of any extended cells. The .bgf output file includes the I CONX, I TRAN, and I ORDER information. (If an atom has no bonds outside the cell the I TRAN information is not saved.)

The information in I CONX and I TRAN is used to generate the energy expression (EEX) each time that SETUP EEX is toggled.

B. MOLXTL

For MOLXTL cases ICONS (and IORDER) are constructed and saved, just as for nonperiodic cases.

G.2 Bond Stretch

The program automatically finds all bonds and associates each with the appropriate force field parameters. If no parameters have been defined for a particular bond, a warning message is displayed and the energy is not calculated until the missing parameter is provided. Toggle MODIFY PARAMS and the program will find and prompt the user for all missing parameters. To turn off all bond interactions, go to DEFAULTS/ENERGY VAR and toggle off BONDS.

The energy for the bonded atoms depends only on the distance R between them. The choices for bond stretch are described below. In most cases $E_b = 0$ when $R = R_0$, so that the value of E_b is a measure of the strain energy (distortion from equilibrium).

a. **Harmonic (Type=1)**

$$E_b = \frac{1}{2} K_b (R - R_0)^2 \quad (\text{G.5})$$

where

R_0 is the bond distance in units of Å.

K_b is the force constant in units of (kcal/mol)/Å² [multiply by 143.88 to convert mdyne/Å to (kcal/mol)/Å²].

The harmonic bond term is simplest and is used in all standard force fields for proteins and nucleic acids. Some papers^{2, 7} write Equation G.5 as

$$E_b = \bar{K} (R - R_0)^2$$

where $\bar{K}$ is half the force constant.

b. **Morse (Type=2)**

$$E_b = D_0 [e^{-\alpha(R - R_0)} - 1]^2 \quad (\text{G.6})$$

where

D_0 is the bond energy in units of kcal/mol.

Here $\alpha = \sqrt{K_b/2D_0}$ is the Morse scaling parameter, which the program calculates from K_b and D_0 .

The Morse potential (see Figure G.1) goes to a finite energy ($E_b = D_0$) for $R = \infty$ and should be used for systems where a covalent bond is allowed to break.

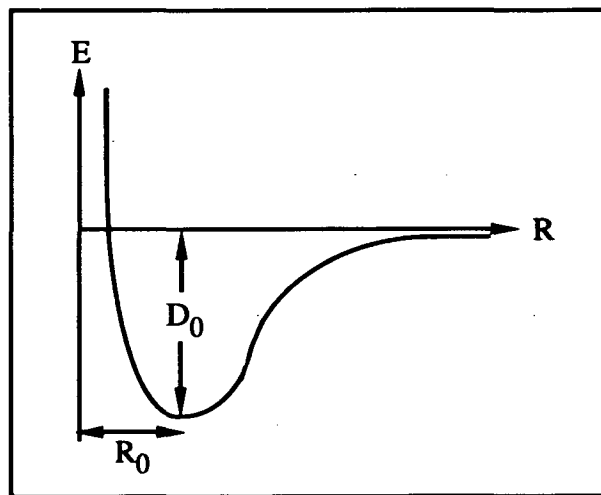


Figure G.1: Morse potential function

Warning: The forces with the Morse term go to 0 as $R \rightarrow \infty$. Consequently, this is not a good form to use when sketching molecules or otherwise starting from a poor geometry.

c. Cubic (Type=3)

$$E_b = \frac{1}{2} K_b (R - R_0)^2 [1 + d(R - R_0)] \quad (\text{G.7})$$

where R_0 and K_b are as above and d is the cubic correction term (default: $d = -2$ for MM2 and MMP2 force fields).

The cubic form is used in the MM2 and MMP2 force fields. As defined, the cubic potential is unsuitable, since $E_b \rightarrow -\infty$ as $R \rightarrow +\infty$ (assuming $d < 0$). Consequently, the program uses the modified form illustrated in Figure G.2. The point A is the maximum for the cubic potential. When $R > R_A$ the sign of dE/dR and d^2E/dR^2 is changed to yield the solid curve.

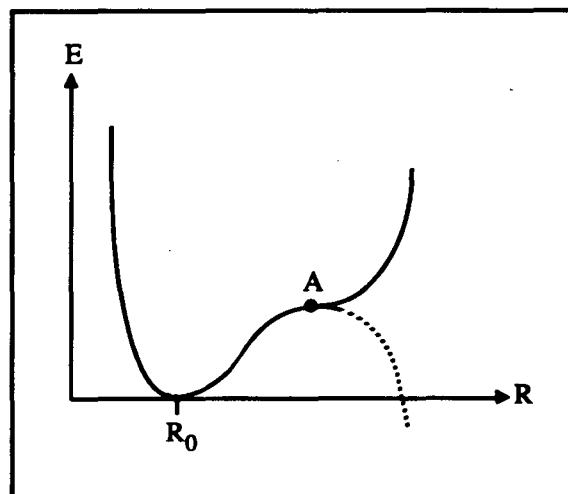


Figure G.2: Modified cubic potential used by the program

d. Polyene (Type=6)

$$E_b = \frac{1}{2} K_2 (R - R_{20})^2 \quad \text{for double bonds} \quad (\text{G.8a})$$

$$E_b = \frac{1}{2} K_1 (R - R_{10})^2 \quad \text{for single bonds} \quad (\text{G.8b})$$

The polyene option is described more fully in Section G.3.3.

e. Pure Repulsive (Type=5) (Note: $E = 0$ for version 3.0)

$$E_b = \frac{De^{-C(R-R_0)}}{R^E} \quad (\text{G.9})$$

Here R_0 is the point at which $E_b = D/R^E$ where E is zero or a positive integer. For inorganic materials where two-body Coulomb and nonbonded interactions are often used between all atoms, an exponential term such as G.9 might be used to modify these interactions for nearest neighbors. If $E = 0$ the repulsive term is pure exponential. If $C = 0$ the remaining term D/R^E is referred to as a Born term (generally $E = 9$ to 14).

f. Nothing (Type=9)

$$E_b = 0 \quad (\text{G.10})$$

The reason for a nothing term is that the program requires a bond term in the EEX for all bonded atoms. Thus, G.10 might be used for an inorganic system where one wants an angle term but does not need a special bond term (one could also just put $K = 0$ for type = 1).

G.3 Bond Angle Bend

Given any two bonds to a common atom (e.g., IJ and JK in Figure G.3), the bond angle interaction is a function of the angle θ between them. Unless the ANGLES switch is turned off [by using the ENERGY VAR (energy variables) menu under DEFAULTS], the program requires force field parameters for all angle terms. If there are no parameters for a particular angle term, a warning message is displayed and an energy is not calculated until these parameters are provided under the MODIFY PARAMS function of the SIMULATE menu.

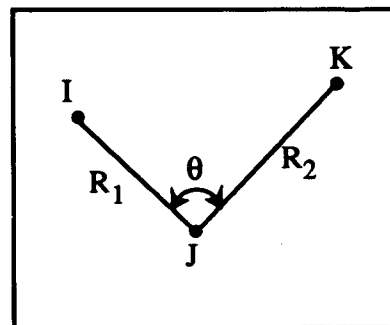


Figure G.3: Bond angle bend

There are two basic expansions for the bond angle terms: cosine θ expansions as described in Section G.3.1 and θ expansions as described in Section G.3.2. In addition there is a special polyene option as described in Section G.3.3.

For an accurate description of the three-body terms (required for good vibrational frequencies), it is important for E_a to depend on the bond distances, $R_1 = R_{IJ}$ and $R_2 = R_{JK}$. This leads to cross terms also discussed in Sections G.3.1 and G.3.2. The θ expansion is used in most force fields for proteins and nucleic acids (AMBER, CHARMM, Dreiding) but does not have the correct behavior for nearly linear bonds. The cosine expansion is more realistic for such cases.

G.3.1 Cosine Expansions

a. Cosine Harmonic (Type=1)

$$E_a = \frac{1}{2} C (\cos\theta - \cos\theta_0)^2 \quad (\text{G.11a})$$

where

$$K_\theta = C \sin^2\theta_0 \quad (\text{G.11b})$$

If $\theta_0 = 180^\circ$, the cosine expansion becomes

$$E_a = K_\theta (1 + \cos\theta) \quad (\text{G.11c})$$

while if $\theta_0 = 0^\circ$, it is

$$E_a = K_\theta (1 - \cos\theta) \quad (\text{G.11d})$$

If $\theta_0 \neq 0^\circ, 180^\circ$, the energy to linearize the molecule is

$$\text{Barrier} = \frac{1}{2} C (1 + \cos\theta_0)^2 = \frac{1}{2} K_\theta \left(\frac{1 + \cos\theta_0}{\sin\theta_0} \right)^2 \quad (\text{G.11e})$$

From Figure G.3 we see that the potential energy should satisfy the relations

$$E(\pi - \theta) = E(\pi + \theta)$$

$$E(-\theta) = E(+\theta)$$

This is always satisfied for the cosine expansion but not for the theta expansion (unless $\theta_0 = 180^\circ$). The input is in terms of θ_0 and K_θ .

b. Cosine Periodic (Type=4)

$$E_a = C [1 - B(-1)^N \cos N\theta] \quad (\text{G.12})$$

This is particularly useful for organometallic systems where $N = 4$ might be used for an octahedral complex and $N = 3$ might be used for a trigonal center.

The parameter B is either

$B = +1$ (leading to a minimum for the linear geometry)

$B = -1$ (leading to a maximum for the linear geometry)

The periodicity N can be $N = 1, 2, 3, 4, 5$, or 6

The constant C is related to the force constant as

$$K_\theta = N^2 C$$

The input is in terms of K_θ , N , and B .

c. Cosine-stretch (Type=11)

To E_a in G.11a we add the cross terms

$$E_{ba} = D(R_1 - R_{10})(\cos\theta - \cos\theta_0) + E(R_2 - R_{20})(\cos\theta - \cos\theta_0) + K_{R_1R_2}(R_1 - R_{10})(R_2 - R_{20}) \quad (\text{G.13a})$$

In the case that $\theta_0 = 180^\circ$ or 0° , the angle-stretch cross terms are zero (by symmetry) and equation G.13a is replaced by

$$E_{ba} = K_{R_1R_2}(R_1 - R_{10})(R_2 - R_{20}) \quad (\text{G.13b})$$

and G.11c is used for E_b . The angle-stretch factors D and E (and $K_{\theta R}$) have units of (kcal/mol)/Å-radian, while the stretch-stretch factor $K_{R_1R_2}$ has units of (kcal/mol)/Å².

From G.13 the angle-stretch force constants are

$$K_{\theta,R_1} = \frac{\partial^2 E}{\partial\theta\partial R_1} = -D \sin\theta \quad (\text{G.14a})$$

$$K_{\theta,R_2} = \frac{\partial^2 E}{\partial\theta\partial R_2} = -E \sin\theta \quad (\text{G.14b})$$

The angle-stretch terms may be input either as D and E or as K_{θ,R_1} and K_{θ,R_2} . If the variable ANGX 2 K in the parameter file is true, then K_{θ,R_1} and K_{θ,R_2} are used. Otherwise D and E are used (this is the default). (ANGX 2 K can be changed using the ANGLE-STRET 2 K option under SIMULATE/DEFAULTS/MISC VAR.)

G.3.2 Theta Expansions

a. Theta Harmonic (Type=21)

$$E_a = \frac{1}{2} K_\theta (\theta - \theta_0)^2 \quad (\text{G.15})$$

b. Angle-stretch (Type=31)

The standard coupling terms for θ expansions have the form

$$E_\theta = K_{\theta,R_1}(R_1 - R_{10})(\theta - \theta_0) + K_{\theta,R_2}(R_2 - R_{20})(\theta - \theta_0) + K_{R_1R_2}(R_1 - R_{10})(R_2 - R_{20}) \quad (\text{G.16})$$

In Equation G.16, D and E have units of (kcal/mol)/(Å-radian).

c. **Harmonic with Linear Theta** (Type=28) (used in CFF of Lifson and Warshel)

$$E_a = \frac{1}{2} K_\theta (\theta - \theta_0)^2 + d K_\theta (\theta - \theta_0) \quad (\text{G.17})$$

This is equivalent to

$$E_a = \frac{1}{2} K_\theta (\theta - \bar{\theta}_0)^2 + C_0 \quad (\text{G.18})$$

where

$$\bar{\theta}_0 = \theta_0 - d$$

and C_0 is a constant.

d. **Quartic Theta Expansion** (Type=29)

$$E_a = \frac{1}{2} K_\theta (\theta - \theta_0)^2 + f K_\theta (\theta - \theta_0) + \frac{1}{2} d K_\theta (\theta - \theta_0)^3 + \frac{1}{2} e K_\theta (\theta - \theta_0)^4 \quad (\text{G.19})$$

The odd powers (f and d) are not allowed for linear molecules.

e. **Schleyer Potential** (Type=24)

The Schleyer potential is a special form of angle expansion potential. It has the form

$$E_\theta = \frac{1}{2} K_\theta (\theta - \theta_0)^2 [1 + d |\theta - \theta_0|] \quad (\text{G.20})$$

where the cubic angle factor $d = -0.55$ in units 1/radian. Note that the cubic term is proportional to the cube of the absolute value of $\theta - \theta_0$.

f. **MM2 Angle Potential** (Type=33)
(also called the Allinger potential)

$$E_a^{MM2} = \frac{1}{2} K_\theta (\theta - \theta_0)^2 [1 + d (\theta - \theta_0)^4] \quad (\text{G.21})$$

where d is the sextic term in units 1/radians⁴. Its default value is

$$0.75437723 = 7 \times 10^{-8} (180/\pi)^4$$

The MM2 angle potential can also include angle-stretch terms,

$$E_{ba}^{MM2} = E(R_1 - R_{10})(\theta - \theta_0) + E(R_2 - R_{20})(\theta - \theta_0) \quad (\text{G.22})$$

The angle-stretch terms are the same as those for G.16, except that in G.22 both bonds have the same angle-stretch factor. In addition with MM2 no angle stretch term is included for bonds to hydrogen atoms.

In the MM2 forcefield certain cases involve a correction to the equilibrium angle θ_0 if one or more hydrogen atoms is bonded to the central atom of the angle. (The correction term depends on whether there are one or two hydrogen atoms.) This is allowed in the program (the F and G input parameters).

G.3.3 Polyene (Type= 6, 16, 26, 36)

For systems such as polyenes, *e.g.*, 1,3-butadiene, the four carbons are all sp^2 and would be considered as C_2 in Dreiding. However, the bond lengths are clearly quite different. In order to handle such cases, the program allows different force constants for different bond orders.

For polyenes the type 6 bond stretch is used, where R_1, K_1 corresponds to the single bond and R_2, K_2 corresponds to the double bond. The angle interactions are described in terms of three sets of parameters as in Figure G.4.

Generally, only case a is of interest, but b is needed for Bucky Balls (C_{60} , C_{70}) and related systems while c is needed by cumulenes. The polyene option can be used with cosine or theta expansions and with or without cross terms.

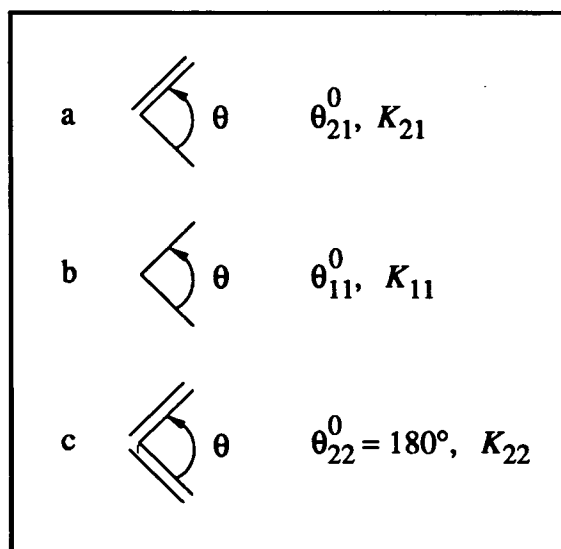


Figure G.4: Angle interactions for the polyene case

G.4 Dihedral Angle Torsions

Given any two bonds IJ and KL attached to a common bond JK (see Figure G.5), the dihedral angle ϕ is defined as the angle between the JKL plane and the IJK plane (see Figure G.6). A positive angle is clockwise when looking from J toward K. $\phi = 0^\circ$ for the *cis* configuration and $\phi = 180^\circ$ for the *trans* configuration. The torsional energy is expanded as

$$E_\theta = \sum_{n=1}^p \frac{1}{2} K_{\theta,n} [1 - d \cos(n\phi)] \quad (\text{G.23})$$

where each $K_{\theta,n}$ is one-half the rotational barrier in kcal/mol, $n = 1, 2, 3, 4, 5, 6$ is the periodicity of the potential and $d = \pm 1$ is the phase factor. For $d = +1$, the *cis* conformation is the minimum while for $d = -1$ the *cis* conformation is the maximum. The input data is K and d for each n .

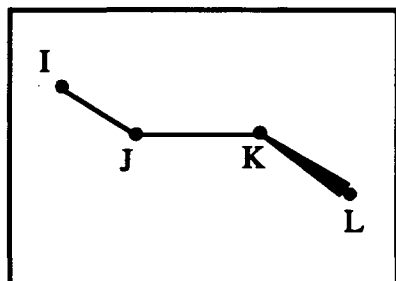


Figure G.5: Dihedral torsion atoms

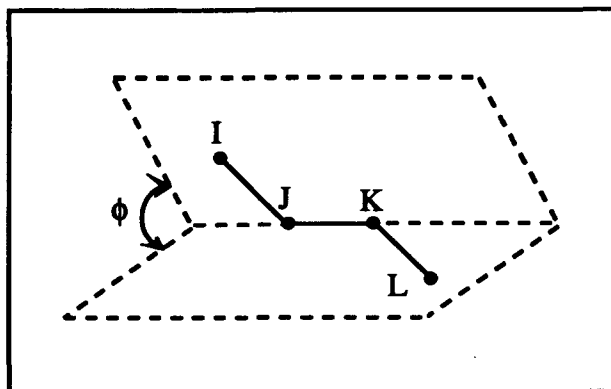


Figure G.6: Definition of torsion angle

For a given bond, the program usually considers all possible torsions (*e.g.*, nine for ethane, four for ethylene), and discards any that are not compatible with the parameter set. With the Dreiding and CHARMM force fields, the K_θ is then rescaled by dividing each K_θ by the number of torsions for the bond. However, with MM2 and AMBER, the K_θ is not rescaled. If USE ALL TORSION on the ENERGY VAR (energy variables) menu is toggled off, then the program uses only the first non-zero torsion found for a particular bond JK (no rescaling).

If none of the dihedrals for a given bond JK have parameters, the program will normally print a warning message and continue. It will calculate energies without further messages. However if the USE ALL TORSION toggle is on (the default), the program will insist on having parameters for each torsion.

If the ETOR SCA parameter is not equal to 1.0 (for example in Dreiding-II), then for sp^2 - sp^2 single-bond torsions that are exocyclic to "aromatic rings" all K_θ values are additionally scaled by the ETOR SCA factor.

G.5 Inversions

Given an atom I having exactly three (and only three) distinct bonds IJ, IK, and IL (see Figure G.7), the force field *may* contain terms affecting the energy involved in planarizing the center atom I. The program allows three ways to do this (corresponding to the common descriptions used in various force fields) as described below.

Since the IL bond is treated differently than the IK and IJ bonds, the program normally does all three cases, treating IL, then IJ, then IK as the special bond and scaling each energy by one-third. If the USE ALL INVERS switch of the ENERGY VAR (energy variables) menu is toggled off, the program will use just the first non-zero inversion term found for a particular I, with no rescaling.

If none of the three inversions for a particular I has parameters, the program will print a warning message and continue. It will calculate energies and forces without further messages. However

if the USE ALL INVERS toggle is on (the default), the program will insist on having parameters for each inversion.

a. Inversion (Type=2)

For the Dreiding-II and spectroscopic force fields, the inversion or umbrella term is defined as

$$E_{\theta} = \frac{1}{2}C(\cos\omega - \cos\omega_0)^2 \quad (\text{G.24})$$

where $K_{\omega} = C \sin^2\omega_0$ is the force constant in (kcal/mol) and ω is the angle between the IL axis and the IJK plane (see Figure G.7). If $\omega_0 = 0^\circ$, the potential term has a minimum for the planar structure. Otherwise it has two minima at $\pm\omega_0$, with a barrier between. This inversion barrier energy is

$$\text{Barrier} = C(1 + \cos\omega_0)^2 = \frac{1}{2}K_{\omega} \left(\frac{1 + \cos\omega_0}{\sin\omega_0} \right)^2 \quad (\text{G.25})$$

If the equilibrium angle ω_0 is 0° (e.g., the atoms bonded to atom I are coplanar), the potential (G.24) is replaced with the form

$$E_{\omega} = K_{\omega}(1 - \cos\omega) \quad (\text{G.26})$$

b. AMBER Improper Torsion JILK (Type=3)

For the Amber force field the inversion term is used only with a few specific cases. It is defined as if it were a torsion (improper torsion JILK) as indicated in Figure G.8. Here

$$E_{\psi} = \frac{1}{2}K_{\psi} \cos[n(\psi - \psi_0)] \quad (\text{G.27})$$

where ψ is the angle between the JIL and KIL planes, K_{ψ} is the rotational barrier in kcal/mol, and n is the periodicity. Here $n = 2$ for planar angles ($\psi_0 = 180^\circ$) and $n = 3$ for tetrahedral angles ($\psi_0 = 120^\circ$). In the AMBER force field $n = 3$ is used for tetrahedral carbon having one implicit hydrogen.

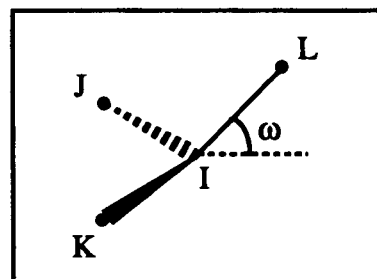


Figure G.7: Umbrella inversion

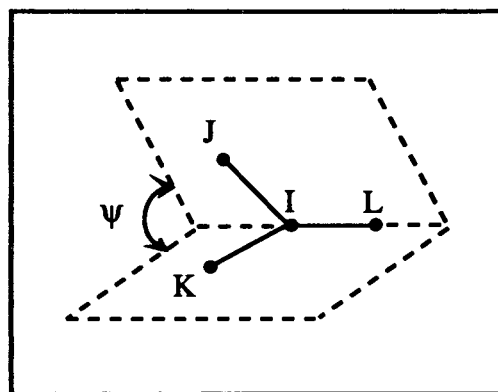


Figure G.8: Amber inversion

c. CHARMM Improper Torsion IJKL (Type=1)

In the CHARMM and Dreiding-I force fields, inversion is defined as if it were a torsion (improper torsion IJKL) as indicated in Figure G.9. Here

$$E_{\phi} = \frac{1}{2}K_{\phi}(\phi - \phi_0)^2 \quad (\text{G.28})$$

where ϕ is the angle between IJK and LJK planes (for a tetrahedral carbon atom with equal bonds this angle is 35.264°) and K_{ϕ} is the force constant in (kcal/mol)/radian².

If ϕ_0 is not zero

$$E(+\phi) \neq E(-\phi)$$

so that this inversion does not have the proper symmetry. This can be used to advantage for enforcing a particular stereochemistry.

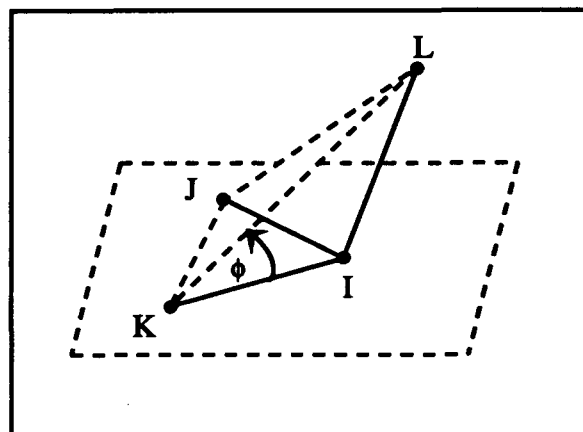


Figure G.9: CHARMM inversion
Taking JK as the axis, ϕ is the angle between the JKI and JKL planes (I is the central atom for inversion)

G.6 Electrostatic Interactions

The electrostatic interactions play a critical role in determining the packing of organic molecules (e.g., drug-receptor interactions) and in the structures of inorganic systems. The total electrostatic energy can be written as

$$E_{\text{coul}} = C_0 \sum_{i>j} \frac{Q_i Q_j}{\epsilon R_{ij}} \quad (\text{G.29})$$

where Q_i and Q_j are the charges (in electron units, thus $Q = +1$ for a proton), R_{ij} is the distance in Å, ϵ is the dielectric constant ($\epsilon = 1$ for a vacuum), and the conversion factor $C_0 = 332.0637$ takes care of the units, giving E_{coul} in kcal/mol.

G.6.1 Charges

The first issue for electrostatics is which charges to use. For small molecules, these charges are sometimes obtained by first calculating the wavefunction [from *ab initio* quantum chemistry (QC) programs such as Gaussian90 or GVB or semiempirical QC programs such as AM1 or MOPAC] and then finding charges that fit the long range electrostatic potentials from the QC. Such charges can be input into the program, however they are available only for selected molecules. More generally useful are the methods which predict the charges given the topology

and electronegativities of the various atoms. There are three such methods built into the program:

- (a) Charge Equilibration developed by Rappé and Goddard⁹ (with support from MSI)
- (b) Gasteiger¹⁰
- (c) del Re¹¹

These methods are discussed further in Section 3.1.13.

For biological systems, the standard force fields (AMBER, CHARMM, Dreiding) have standard charges associated with normal amino acids and bases. For each such force field the program has a conversion file (e.g., *dreidii300.cnv*) which defines the standard charge for each type of atom. Thus reading a Brookhaven file (see Appendix D) automatically leads to the standard charges.

Charges can also be modified by using the CHARGES function on the EDIT ID menu of the Builders (see Section 3.1.13). BioDesign General Format type coordinate files (*.bgf*) include charge explicitly.

G.6.2 Coulomb Summations

For a molecule with N atoms there are $N(N-1)/2$ Coulomb interactions in G.29. Thus for hemoglobin with 6000 atoms there are 18,000,000 interactions to evaluate. In order to decrease the computational costs one would like to consider only the closer interactions (say $R_{ij} < R_{cut}$). However, there are difficulties with this since the Coulomb potential (G.29) is very long range. Thus two unit charges separated by 10 Å have an interaction of 33.2 kcal/mol and charges separated by 100 Å still have 3.3 kcal/mol interaction energy.

The standard approach for handling this problem is to modify G.29, multiplying by the cubic spline cutoff function

$$S(R_{ij}, R_{on}, R_{off})$$

defined to be unity for $R_{ij} \leq R_{on}$, zero for $R_{ij} \geq R_{off}$, and decreasing smoothly in between. [The detailed form of the switching function is presented in Appendix H]. As discussed in Appendix H, the cutoff function insures continuous changes in energy and forces. The idea is that at larger distances, the distribution of positive and negative charges will approximately cancel so that use of the cutoff function will not lead to appreciable errors.

⁹ A.K. Rappé, and W.A. Goddard III, *J. Phys. Chem.* **95** (1991) 3358-3363.

¹⁰ J. Gasteiger, *Tetrahedron*, **36** (1980) 32190-3228.

¹¹ G. del Re, *J. Chem. Soc.*, (1958) 4031-4040.

G.6.3 Nonbond Exclusions

In most covalent force fields, it is considered that the electrostatic (and vdW) interactions of bonded atoms (1-2 interactions) are included in the bond stretch terms and that the electrostatic (and vdW) interactions of atoms bonded to a common atom (1-3 interactions) are included in the bond-angle terms. Thus these terms are generally excluded from G.29. (This is the case for AMBER, CHARMM, and Dreiding). For inorganic systems some force fields do not exclude such interactions and the SIMULATE/DEFAULTS/NONBOND VAR menu allows one to change these conventions [bond (1-2) and angle (1-3)]. The program makes a list of these exclusions each time the energy expression is setup.

Thus the final form of the electrostatic energy is

$$E_{el} = C_0 \sum_{\substack{R_{ij} < R_{cut} \\ [excl(1-2,1-3)]}} \frac{Q_i Q_j}{\epsilon R_{ij}} S(R_{ij}, R_{on}, R_{off}) \quad (G.30)$$

For crystals, the sum in G.29 involves an infinite number of atoms and a different approach, the Ewald procedure, is required for obtaining accurate answers.

For most force fields electrostatic energies are included for i and j that are separated by three bonds (1-4 interactions). The exception is the AMBER force field, where these 1-4 interactions are reduced by a factor of 2. The SCALE 1-4 toggle under SIMULATE/DEFAULTS/NONBOND VAR allows this scaling to be modified. (Setting the scale to zero eliminates 1-4 interactions).

G.6.4 Dielectric Constant

For biological systems in which solvent molecules are *not* included explicitly, it is common to use

$$\epsilon = \epsilon_0 R_{ij} \quad (G.31)$$

where $\epsilon_0 = 1$. This is the default for the AMBER, CHARMM, and Dreiding force fields. To use

$$\epsilon = \epsilon_0 \quad (G.32)$$

toggle off the R*EPS switch on the NONBOND VAR (nonbond variables) menu. The value of ϵ_0 can also be adjusted (default is $\epsilon_0 = 1$).

G.6.5 Nonbond Lists

After setting up the energy expression, the program generates the NBLIST of all nonbond interactions where

$$R_{ij} < R_{cut}$$

and where i and j are not excluded. Nonbond interactions are thus calculated only for pairs within the cutoff distance. During dynamics or minimization this list is updated periodically (usually every 50 to 100 steps of dynamics). This list does not include any 1-2 and 1-3 interactions that are to be excluded (G.6.4).

G.7 van der Waals Interactions

All atoms exhibit a long range attraction proportional to $1/R^6$. This is generally referred to as van der Waals (vdW) attraction since the van der Waals equation of state postulated such a universal attraction. The quantum mechanical origin of this attraction was established by London and it is often referred to as the London force or as the dispersion interaction (the name London used).

At sufficiently short distances the interactions of all atoms are repulsive (due to the overlap of orbitals and the Pauli Principle). Thus quantum mechanics leads to a general nonbonded interaction of the form in Figure G.10. We will refer to this as the van der Waals interaction. In addition to $1/R^6$, the long range potential should also include $1/R^8$, $1/R^{10}$, and higher order terms, but they are generally neglected. The short range potential is often written as e^{-CR} (where $C > 0$) since the overlap has this form. More often, the simple Born repulsions is used, $1/R^N$, (where $N \geq 9$). Although $1/R^6$ falls to zero much more rapidly than the Coulomb interaction, these terms all have the same sign (no cancelling due to opposite charges) and hence the cutoff functions for calculation of accurate vdW interactions is about the same as for Coulomb. Thus we use the same exclusions, cutoffs, and switching functions as for electrostatic interactions. The total van der Waals energy is written as

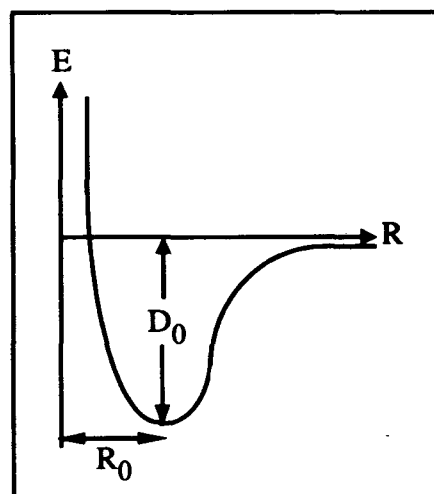


Figure G.10: The form of the van der Waals interaction

$$E_{vdw} = \sum_{\substack{R_{ij} < R_{cut} \\ [excl(1-2,1-3)]}} E_{vdw}(R_{ij}) S(R_{ij}, R_{on}, R_{off}) \quad (G.33)$$

where several forms are allowed for the potential function $E_{vdw}(R_{ij})$. In each case

D_0 = bond strength (well depth) in kcal/mol

R_0 = bond length in Å

a. Lennard-Jones 12-6 (Type=1)

$$E_{vdw}(R) = AR^{-12} - BR^{-6} = D_0 \left\{ \left[\frac{R_0}{R} \right]^{12} - 2 \left[\frac{R_0}{R} \right]^6 \right\} \quad (\text{G.34})$$

The input is in terms of R_0 and v_0 .

b. Exponential-6 (Type=2)

$$E_{vdw}(R) = Ae^{-CR} - BR^{-6} = D_0 \left\{ \left[\left(\frac{6}{\zeta - 6} \right) \exp^{\zeta \left(1 - \frac{R}{R_0} \right)} \right] - \left[\left(\frac{\zeta}{\zeta - 6} \right) \left(\frac{R_0}{R} \right)^6 \right] \right\} \quad (\text{G.35})$$

The default value for the scaling factor γ is 12, which makes the exponential-6 potential approximate the Lennard-Jones 12-6 potential at long distances. (For more details see reference 1.) The input is in terms of R_0 , v_0 , and ζ .

c. Morse Potential (Type=3)

$$E_{vdw}(R) = D_0 [\chi^2 - 2\chi] \quad (\text{G.36})$$

where

$$\chi = e^{-\alpha(R-R_0)} = \exp \left[-\frac{\gamma}{2} \left(\frac{R}{R_0} - 1 \right) \right] \quad (\text{G.37})$$

The input is in terms of R_0 , v_0 , and γ .

The three potentials in a–c all have a repulsive inner wall and attractive long-range behavior. The Lennard-Jones 12-6 potential is the default for the Dreiding-I, Dreiding-II, AMBER, and CHARMM force fields, while exponential-6 is the default for the MM2/MMP2/MM3 force fields. The Morse term is used in some special force functions (H₂O dimer, interstitial electron model of metals¹²).

¹² M. Li and W. A. Goddard III, *Phys. Rev. B.*, **40** (1989) 12155-12163.

The Lennard-Jones 12-6 potential may be too repulsive at short ranges, whereas the exponential-6 potential is expected to have a more reasonable form¹³ (since the overlap changes exponentially). However for sufficiently small R the exponential-6 potential (G.35) goes to $-\infty$. This can cause problems when starting with bad geometries. The program checks for the short range inflection point and keeps the potential repulsive for shorter R in order to avoid this problem (see Appendix I). (A similar problem occurs for the Morse potential.)

In addition to the above potentials the program allows others for special purposes.

d. Lennard-Jones 12-10 (Type=7)

$$E_{vdw}(R) = AR^{-12} - BR^{-10} = D_0 \left\{ 5 \left[\frac{R_0}{R} \right]^{12} - 6 \left[\frac{R_0}{R} \right]^{10} \right\} \quad (\text{G.38})$$

This form is used by AMBER for hydrogen bond interactions. In this case R is the distance between the acceptor atom (N, O, F, S, Cl) and the hydrogen atom of the donor.

e. Hard Exponential-6 (Type=8) (not included in version 3.0)

$$E_{vdw} = \frac{(Ae^{-CR} - B)}{R^6} \quad (\text{G.39})$$

where $A > B$.

This avoids the $-\infty$ singularity of G.35 for small R and may eventually become the standard in place of G.35.

f. Hard Morse (Type=9) (not included in version 3.0)

$$E_{vdw} = D_0 (\chi^2 - 2\chi) / R \quad (\text{G.40})$$

This leads to a repulsive potential of the form

$$\frac{1}{R} D_0 e^{+2\alpha R_0}$$

for small R , which will cancel Coulomb interactions (if $D_0 e^{+2\alpha R_0} > |Q_1 Q_2|$). This is being used for the IEM of metals.¹²

¹³ N. Karasawa and W. A. Goddard III, *J. Phys. Chem.*, (1991) submitted for publication.

e. Pure Repulsive Exponential (Type=5)

$$E_{vdw} = Ae^{-CR(R)} = D_0 e^{\gamma \left(1 - \frac{R}{R_0}\right)} \quad (G.41)$$

where the form is chosen to be consistent with Equation G.35. This is often used for inorganic force fields.

g. Pure R^{-6} Attraction (Type=4)

$$E_{vdw} = -BR^{-6}(R) = -2D_0 \left(\frac{R_0}{R}\right)^6 \quad (G.42)$$

where the form is chosen to be compatible with Equations G.34 and G.35.

h. No van der Waals (Type=6)

The potential may also be set to 0, meaning that van der Waals nonbond interactions between this atom and other atoms are not calculated (there may be nonzero Coulomb interactions).

G.8 Combination Rules (Off Diagonal vdW)

The experimental data on which the van der Waals parameters (R_0 , D_0 , γ) are based is sparse. Thus the parameters for carbon are often based on the interplanar interaction in graphite, and those for P and S can be based on the molecular crystals (P₄, S₈). However the diatomic molecular crystals (H₂, N₂, O₂, F₂, Cl₂, Br₂, I₂) do not mimic the H, N, F, *etc.* in hydrocarbons, fluorocarbons, *etc.* Thus H must be based on hydrocarbons involving simultaneous C·C, C·H, and H·H van der Waals interactions. For most atoms there is no experimental data relating directly to vdW. In order to keep the number of parameters at a minimum, it is generally assumed that the R_0 and D_0 between two different atoms (off diagonal vdW), say C·H, is related to the diagonal values (*e.g.*, C·C and H·H). There are two common combination rules, geometric mean and arithmetic mean.

G.8.1 Geometric Mean

With geometric mean the off diagonal vdW parameters for the Lennard-Jones and exponential-6 potentials G.34 and G.35 are defined as

$$A_{ij} = \sqrt{A_{ii} A_{jj}} \quad B_{ij} = \sqrt{B_{ii} B_{jj}} \quad C_{ij} = \frac{1}{2} (C_{ii} + C_{jj}) \quad (G.43)$$

This allows efficient Ewald procedures¹⁴ and is assumed in some of the advanced methodologies (e.g., CMM method¹⁵) being developed. For Lennard-Jones, G.34, this leads to

$$(D_0)_{ij} = \sqrt{(D_0)_{ii}(D_0)_{jj}} \quad (\text{G.44})$$

$$(R_0)_{ij} = \sqrt{(R_0)_{ii}(R_0)_{jj}} \quad (\text{G.45})$$

For G.35 the use of G.43 does not lead to such simple expressions for D_{0ij} , R_{0ij} , and ζ_{ij} .

G.8.2 Arithmetic Mean

$$(D_0)_{ij} = \sqrt{(D_0)_{ii}(D_0)_{jj}} \quad (\text{G.46})$$

$$(R_0)_{ij} = \frac{1}{2} [(R_0)_{ii} + (R_0)_{jj}] \quad (\text{G.47})$$

G.8.3 Conventions

The default for the CHARMM and Dreiding-I force fields is arithmetic mean. The default for the AMBER and Dreiding-II force fields is geometric mean. The program allows the user to switch between geometric and arithmetic mean with the GEOMETRIC MEAN/ARITHMETIC MEAN toggle under the SIMULATE/DEFAULTS/EWALD VAR or the SIMULATE/DEFAULTS/NONBOND VAR menus.

G.8.4 Nonbond Off-Diagonal

The above combination rules can be overridden by explicitly entering values for the van der Waals type, D_{0ij} , R_{0ij} , and ζ_{ij} for each pair of atoms. Thus with AMBER the Lennard-Jones 12-10 interaction (G.38) is used for H··O while H··H and O··O are described by Lennard Jones 12-6 (G.34).

For inorganic materials, the special nonbond off-diagonal terms are often useful.

¹⁴ N. Karasawa and W. A. Goddard III, *J. Phys. Chem.*, **93** (1989) 7320-7327.

¹⁵ H. Ding, N. Karasawa, and W. A. Goddard III, *PRL*, submitted for publication.

G.9 Hydrogen Bond

Some force fields (*e.g.*, CHARMM, Dreiding-I and Dreiding-II) use a "special" hydrogen bond potential to describe the interactions between atoms involved in hydrogen bonds. We let R_{AD} be the radial distance between the donor (D) and acceptor (A) atoms and θ_{AHD} be the bond angle between the acceptor (A), the hydrogen (H) and the donor (D), as in Figure G.11. In the program this hydrogen bond potential has the general form:

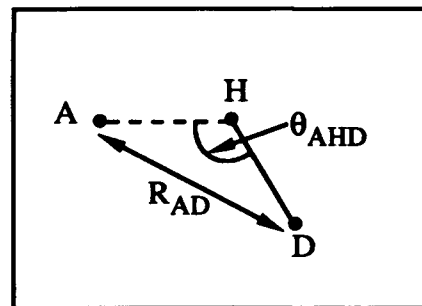


Figure G.11: The geometry for hydrogen bonds

$$E_{hb} = \sum_{R_{AD} < R_{cut}} E^{hb}(R_{AD}) \cos^4(\theta_{AHD}) S(R_{AD}^2, R_{on}^2, R_{off}^2) \times S[\cos^2(\theta_{AHD}), \cos^2(\theta_{on}), \cos^2(\theta_{off})] \quad (G.48)$$

where $\cos^2 \theta_{AHD} > \cos^2 \theta_{cut}$

S is the cubic spline switching function which smoothly cuts off the interaction. (The detailed form of the switching function is presented in Appendix H.) The sum is over all pairs of hydrogen acceptors and donors (A, D) separated by less than a cutoff distance R_{cut} (see Section 5.12.6) and such that the angle has $\cos^2 \theta_{AHD} > \cos^2 \theta_{cut}$ (usually $\theta_{cut} = 90^\circ$, that is $\cos^2 \theta_{cut} = 0$).

a. Lennard-Jones 12-10 (Type=1)

The default term for the Dreiding-I and II force field is the Lennard-Jones 12-10 potential:

$$E^{hb}(R) = AR^{-12} - BR^{-10} = D_0 \left\{ 5 \left[\frac{R_0}{R} \right]^{12} - 6 \left[\frac{R_0}{R} \right]^{10} \right\} \quad (G.49)$$

where D_0 is the hydrogen bond strength in kcal/mol.

b. Morse (Type=2)

A Morse potential can also be used for the hydrogen bond potential:

$$E_{Morse}^{hb} R = D_0 \{ \chi^2 - 2\chi \} \quad (G.50)$$

where

$$\chi = e^{-\alpha(R - R_0)} = e^{-\frac{\gamma}{2} \left(\frac{R}{R_0} - 1 \right)}$$

G.10 Urey-Bradley

An alternative or supplement to bond-angle interactions (G.2) are Urey-Bradley potentials between atom pairs having 1-3 interactions (atoms bonded to a common atom). The Urey-Bradley potential can have the following forms. In each case there are two choices for R_0 . One is to supply a fixed value (Type 1, 2, 3); the other (Type 11, 12, 13) is to always use the instantaneous value of R_{1-3} as R_0 .

a. Harmonic (Type=1, 11)

$$E^{UB} = \frac{1}{2}K_{UB} (R - R_0)^2 + dK_{UB}R_0 (R - R_0) \quad (\text{G.51})$$

where K_{UB} is the Urey-Bradley force constant in (kcal/mol)/Å², and d is the linear factor (default $d = -0.1$). For version 3.0, type=2 is used in place of type=11.

b. Exponential (Type=2, 12) (not included in version 3.0)

$$E^{UB} = \frac{De^{-C(R-R_0)}}{R^E} \quad (\text{G.52})$$

c. Lennard-Jones (Type=3, 13) (not included in version 3.0)

$$E^{UB} = D_0 [(R_0/R)^{12} - 2(R_0/R)^6] \quad (\text{G.53})$$

G.11 Default Potential Terms for the Standard Force Fields

TERM	FORCE FIELD	POTENTIAL	ITYP
Bonds	Dreiding-I	Harmonic	1
	Dreiding-II	Harmonic	1
	MM2	Cubic	3
	AMBER	Harmonic	1
	CHARMM	Harmonic	1
Angles	Dreiding-I	Simple harmonic	21
	Dreiding-II	Simple harmonic	21
	MM2	MM2	33,34,35
	AMBER	Simple harmonic	21
	CHARMM	Simple harmonic	21
Torsions	<i>All</i>	Cosine expansion	
Inversions	Dreiding-I	Improper Torsion IJKL	1
	Dreiding-II	Umbrella	2
	MM2	<i>not used</i>	
	AMBER	Improper Torsion JILK	3
	CHARMM	Improper Torsion IJKL	1
van der Waals	Dreiding-I	Lennard-Jones 12-6	1
	Dreiding-II	Lennard-Jones 12-6	1
	MM2	Exponential-6	2
	AMBER	Lennard-Jones 12-6	1
	CHARMM	Lennard-Jones 12-6	1
Off diagonal nonbond	Dreiding-I	Combination rule	
	Dreiding-II	Combination rule	
	MM2	Combination rule except for C-H interactions (Exponential-6)	
	AMBER	Combination rule except for H··X of H bonds (Lennard-Jones 12-10)	
Hydrogen bond	Dreiding-I	Lennard-Jones 12-10	1
	Dreiding-II	Lennard-Jones 12-10	1
	MM2	<i>not used</i>	
	AMBER	<i>not used</i>	
	CHARMM	Lennard-Jones 12-10	1

Appendix H

Long-Range Cutoffs

The number of nonbond interactions increases as the square of the number of atoms, so that for 3000 atoms there are 4.5 million possible pairwise interactions. To reduce the computational requirements, the program calculates the nonbond interactions only for pairs of atoms within a cutoff distance, R_{cut} . The cutoff distance must be chosen so that the interactions which are being excluded are negligible. The default for R_{cut} is 9.0 Å. During the setup of the energy expression the program calculates a list of atom pairs within the cutoff distance. The list must be updated periodically as atoms move during an energy minimization or dynamics calculation (see Section 5.12.6). For the atom pair in the nonbond list, the program multiplies the interaction energy E_{ij} by a function S which switches off the interactions smoothly so that there are no forces at the boundaries of the cutoff region.

The switching function used by the program is defined below:

$$\begin{aligned}
 S(x, x_{on}, x_{off}) &= 1 && \text{if } x \leq x_{on} \\
 S(x, x_{on}, x_{off}) &= \frac{(x_{off} - x)^2(x_{off} + 2x - 3x_{on})}{(x_{off} - x_{on})^3} && \text{if } x_{on} < x < x_{off} \\
 S(x, x_{on}, x_{off}) &= 0 && \text{if } x_{off} \leq x
 \end{aligned}$$

where

$$\begin{aligned}
 x &= R_{ij}^2 \\
 x_{on} &= R_{on}^2 \\
 x_{off} &= R_{off}^2
 \end{aligned}$$

The default values for R_{on} and R_{off} are 8.0 and 8.5 Å, respectively. (Note: R_{off} should be slightly smaller than R_{cut} since atom pairs which are initially greater than R_{off} apart may move closer than R_{off} during an energy minimization or dynamics calculation.)

The same switching function is used for the hydrogen bond interactions. It is used for both the distance and angle cutoffs. For the angle cutoffs, the program does not use the angle θ_{AHD} between the hydrogen acceptor (A), hydrogen (H), and donor (D) directly but rather the complementary angle $180 - \theta_{AHD}$, which yields the same results since

$$\cos^2 (180 - \theta) = \cos^2 (\theta).$$

Therefore, the program uses

$$\begin{aligned}x &= \cos^2 (180 - \theta) \\x_{on} &= \cos^2 (180 - \theta_{on}) \\x_{off} &= \cos^2 (180 - \theta_{off})\end{aligned}$$

for the angular switching function.

When working with a finite system where the number of atoms is limited (on the order of 10^2), it is more accurate to include *all* the nonbonded interactions in the calculations. This is implemented using the NO LIST method (under NONBOND VAR).

H.1 Periodic Systems and the Ewald Summation Method

The use of a cutoff distance as in the spline switching method shortens the time required to calculate the nonbond energy and forces, but introduces errors. If the interaction potential drops off fast enough, the error is small and this approach is acceptable. But if the interaction potential drops off slowly, the error is large, making this approach less acceptable. This is the case for electrostatic interactions. For instance, to achieve an accuracy of 0.01 kcal/mol in electrostatic energy for crystalline NaCl would require a cutoff distance of more than 300000 Å.

Bulk systems have an infinite number of atoms and must be modeled using periodic boundary conditions. The error associated with the use of a cutoff distance is multiplied in infinite systems making the use of a more accurate method more important. An accurate calculation of nonbond interactions in periodic systems can be achieved using the Ewald Summation method as described by Karasawa and Goddard¹. This method uses an Accuracy Bound criteria for Convergence Acceleration (ABCA) and can be accessed by specifying the EWALD method on the NONBOND VAR menu.

¹ N. Karasawa and W. A. Goddard, III, *J. Phys Chem*, 1989, 93, 7320.

Appendix I

Parameters

The parameters for the force fields described in Appendix G are contained in ASCII files in the *bg_exe* directory. Parameters such as equilibrium bond lengths, bond energies, bond angles and angle force constants, *etc.*, are contained in the parameter files listed in Table I.1 (*.par* is the file type). Each atom is given a 2 to 5 character atom type name, (*e.g.*, C_3) that determines its force field parameters. A conversion file is used to convert atom labels contained in external file formats (*e.g.*, Brookhaven, Cambridge, MOPAC, CHM, MOL, BIG STRAIN) to Dreiding-I or Dreiding-II atom types. The conversion files have the same names as the corresponding parameter files, but the file type is *.cnv* instead of *.par*.

Standard Force Field Parameter and Conversion Files*	
Force Field	UNIX File Name
Dreiding-I	\${BG_EXE}dreidixxx.par \${BG_EXE}dreidixxx.cnv
Dreiding-II	\${BG_EXE}dreidiixxx.par \${BG_EXE}dreidiixxx.cnv
MM2	\${BG_EXE}mm2xxx.par \${BG_EXE}mm2xxx.cnv
MMP2	\${BG_EXE}mmp2xxx.par \${BG_EXE}mmp2xxx.cnv
AMBER	\${BG_EXE}amberxxx.par \${BG_EXE}amberxxx.cnv
AMBER-B	\${BG_EXE}amberbxxx.par \${BG_EXE}amberbxxx.cnv

*xxx = version number $\times$ 100, *e.g.*, 300 = version 3.0

Table I.1: Force Field Parameter and Conversion Files

The conversion files also contain data on atomic charges, the maximal number of covalent bonds, the number of lone pairs, the default atom color for CPK rendering, and the number of three- and four-membered ring parameters. The remainder of the force field parameters are contained in the parameter files.

The parameter file name is assigned to FORTRAN unit 21, and the conversion file name is assigned to unit 22. The user specifies the force field to be used when creating the runtime startup files with the COMSTRUCT program (see the *Molecular Simulations Installation Guide*). The following sections present the file formats for the parameter and conversion files, as well as a list of the default atom types and parameters.

Parameters may be modified interactively (see Section 5.13) or by editing any of the parameter and/or conversion files. The modified parameter file and conversion file can be read into the program using the READ PAR FILE operation (Section 5.13.1). The modified conversion file can also be read using the READ CONV FILE operation (Section 3.13.1). A user can also choose to use his own parameter and conversion files when he creates runtime startup files (UNIX shell scripts) with COMSTRUCT.

I.1 Parameter File Format

The parameter file is an ASCII file, which is divided into 18 sections, each starting with a keyword record. Some of these sections may not be present in all parameter files. Comments beginning with an asterisk in column 1 may be entered within any of the sections after the file header.

PAR	Parameter file header
DEFAU	Defaults parameters
FFLABEL	Force field type labels parameters
ADDED H	Implicit hydrogen relationships parameters
LONE PAIRS	Lone pair types parameters
Q EQUIL	Charge equilibration parameters
DEL RE ATOMS	Del Re charge by element parameters
DEL RE FF	Del Re charge by force field type parameters
GASTEIGER	Gasteiger charge parameters
VDW A	Atomic parameters
AUTOT	Auto type parameters
NONBO	Off-diagonal nonbond parameters
BONDS	Bond stretching parameters
ANGLE	Bond angle bending parameters
TORSI	Dihedral angle torsion parameters
INVER	Inversion (improper torsion) parameters
HBOND	Hydrogen bond parameters
USER	User-supplied parameters

The parameters are assigned in terms of 5-character atom types. The default atom types are listed in the next section. The letter X can serve as a wild card for any atom type. The potentials used in the force field are described in Appendix G. The format for the parameter file is given below.

I.1.1 Parameter File Header**1. PAR**

FORMAT (A3)

File header. Must have the following value: 'PAR'.

2. ICNTRL

FORMAT (16I5)

Control parameters.

Col.

1-5 Version number of the program under which this file was written.

6-10 Force field index of the force field for which this file was written.

FFINDEX	Force Field
1	DREIDING-I or II
2	CHARMM
3	AMBER
4	MM2 or MMP2

3. TITLES

FORMAT (A)

File titles. These are comments that will be printed in the text window when the program reads the parameter file. Multiple records are allowed (up to 10). The list is terminated by a record with only '*' in column 1.

Col.

1-80 Descriptive title for file.

4. FORCE

FORMAT (A5,6X,A8)

Col.

1-5 'FORCE'

12-19 Force field name.

I.1.2 Defaults Parameters (optional)**1. DEFAU**

FORMAT (A5)

Keyword for defaults parameters. Must have the following value: 'DEFAU'.

2. LBOND

FORMAT (A8,2X,L2)

Default for determining whether bond terms will be included in the energy calculations.

Col.

1-8 'LBOND '

11-12 LBOND = True for including bond terms.

3. LANGLE

FORMAT (A8,2X,L2)

Default for determining whether angle terms will be included..

Col.

1-8 'LANGLE '

11-12 LTHETA = True for including angle terms.

4. ANGX 2 K

FORMAT (A8,2X,L2)

Default for determining whether true force constants are used for cosine angle-stretch cross terms.

Col.

1-8 'ANGX 2 K'

11-12 ANGX2K = True for using true force constants for cosine angle-stretch cross terms

5. ANGANGIN

FORMAT (A8,2X,L2)

Default for determining whether angle-angle inversion terms are used.

Col.

1-8 'ANGANGIN'

11-12 LANGANG = True for using angle-angle inversion terms.

6. LINVERSN

FORMAT (A8,2X,L2)

Default for determining whether inversion terms will be included.

Col.

1-8 'LINVERSN'

11-12 LIMPHI = True for including inversion terms.

7. ALL INVE

FORMAT (A8,2X,L2)

Default for determining whether only one inversion term per center is used.

Col.

1-8 'ALL INVE'

11-12 LALLINV = True for using only one inversion term per center in the energy calculations

8. BNDXANG

FORMAT (A8,2X,L2)

Default for determining whether bond cross angle terms are used.

Col.

1-8 'BNDXANG'

11-12 LBNDXANG = True for using bond cross angle terms.

9. ANGXANG

FORMAT (A8,2X,L2)

Default for determining whether angle cross angle terms are used.

Col.

1-8 'ANGXANG '

11-12 LANGXANG = True for using angle cross angle terms.

10. LTORSION

FORMAT (A8,2X,L2)

Default for determining whether torsion terms will be included.

Col.

1-8 'LTORSION'

11-12 LPHI = True for including torsion terms.

11. BNDBNDTO

FORMAT (A8,2X,L2)

Default for determining whether coupling of the 1-2 and 3-4 bonds of torsions is allowed.

Col.

1-8 'BNDBNDTO'

11-12 LBNDTOR = True for allowing coupling of the 1-2 and 3-4 bonds of torsions.

12. ANGANGTO

FORMAT (A8,2X,L2)

Default for determining whether coupling of the 1-2-3 and 2-3-4 angles of torsions is allowed.

Col.

1-8 'ANGANGTO'

11-12 LANGTOR = True for allowing coupling of the 1-2-3 and 2-3-4 angles of torsions.

13. LPITWIST

FORMAT (A8,2X,L2)

Default for determining whether pi twist terms will be included.

Col.

1-8 'LPITWIST'

11-12 LPITWIST = True for using pi twist terms.

14. TORS SCA

FORMAT (A8,2X,L2)

Default for determining torsion summation rule.

Col.

1-8 'TORS SCA'

11-12 TORSCAL = True for summing all torsions about a particular bond and scaling each by the total number of torsions for that bond; false for simply summing all torsions.

15. ALL TORS

FORMAT (A8,2X,L2)

Default for determining whether only one torsion term per center is used.

Col.

1-8 'ALL TORS'

11-12 LALLTOR = True for using only one torsion term per center.

16. ETOR SCA

FORMAT (A8,2X,F10.4)

Default for sp^2 - sp^2 single bond scale factor for exocyclic torsions.

Col.

1-8 'ETOR SCA'

11-20 SCLETOR = Scale factor for single bond sp^2 - sp^2 torsions that are exocyclic to aromatic rings.**17. TORANGSW**

FORMAT (A8,2X,L2)

Default for determining whether the torsion barrier is switched off as the torsion angle becomes linear.

Col.

1-8 'TORANGS'

11-12 TORANGSW = True for switching torsion barrier off as angle becomes linear.

17. TORANGR

FORMAT (A8,2X,2F5.0)

On and off angles for torsion angle switch (see TORANGSW above).

Col.

1-8 'TORANGR '

11-15 TTORSWON = On angle for torsion angle switch

16-20 TTORSWOFF = Off angle for torsion angle switch

18. LNONBOND

FORMAT (A8,2X,L2)

Default for determining whether nonbond terms are included.

Col.

1-8 'LNONBOND'

11-12 LNBOND = True for including nonbond terms.

19. RNB GEOM

FORMAT (A8,2X,L2)

Default rule for determining heteroatomic nonbond interactions.

Col.

1-8 'RNB GEOM'

11-12 RNBGEOM = True for geometric-mean rule for nonbond interactions; false for arithmetic-mean rule.

20. NBEXBND

FORMAT (A8,2X,L2)

Default for determining whether 1-2 terms are excluded from nonbonds.

Col.

1-8 'NBEXBND '

11-12 LNBEXCB = True for excluding 1-2 terms from nonbonds

21. NBEXANG

FORMAT (A8,2X,L2)

Default for determining whether 1-3 terms are excluded from nonbonds.

Col.

1-8 'NBEXANG'

11-12 LNBEXCT = True for excluding 1-3 terms from nonbonds

22. SCAL NB1

FORMAT (A8,2X,F10.4)

Default for 1-4 nonbond scale factor.

Col.

1-8 'SCAL NB1'

11-20 SCLNB14 = Scale factor for 1-4 nonbond interactions.

23. SHRINK C

FORMAT (A8,2X,L2)

Default for determining whether C-H bonds are contracted during calculations.

Col.

1-8 'SHRINK C'

11-12 SHRINKCH = True for contracting C-H bonds during calculations.

24. SHRINK F

FORMAT (A8,2X,F10.4)

Shrink factor for C-H bonds.

Col.

1-8 'SHRINK F'

11-20 TFACTSCH = Shrink factor for C-Hbonds.

25. LCOULMB

FORMAT (A8,2X,L2)

Default for determining whether electrostatic terms are included in the energy calculations.

Col.

1-8 'LCOULMB'

11-12 LCOULOMB = True for including electrostatic terms in the energy calculations.

26. R*EPS

FORMAT (A8,2X,L2)

Default for determining the use of a distant-dependent dielectric constant.

Col.

1-8 'R*EPS '

11-12 LREPS = True for a distant-dependent dielectric constant.

27. LHBOND

FORMAT (A8,2X,L2)

Default for whether explicit hydrogen bond terms will be used.

Col.

1-8 'LHBOND '

11-12 LHBOND = True for explicit hydrogen bond terms.

28. ATM DEFL

FORMAT (A8,2X,A5)

Default atom type for builders.

Col.

1-8 'ATM DEFL'

11-15 DEFATO = Default atom type for builders.

29. MASSZER

FORMAT (A8,2X,L2)

Default for determining whether zero mass is used during the energy calculations.

Col:

1-8 'MASSZER'

11-12 LZERMAS = True for using the zero mass option.

30. POLYENE

FORMAT (A8,2X,L2)

Default for determining whether the polyene option is used during the energy calculations.
(i.e., whether the bond and angle force constants depend on bond order).

Col.

1-3 'POLYENE '

11-12 LPOLYENE = True for using the polyene option.

31. USRENERG

FORMAT (A8,2X,L2)

Default for determining whether the user-defined energy expression is used.

Col.

1-8 'USRENERG'

11-12 LUSERE = True for using the user-defined energy expression.

Terminate with 'VDW', 'ATOM', 'ATOMIC', or 'FFLABEL'.

I.1.3 Force Field Type Labels Parameters**1. FFLABEL**

FORMAT (A7)

Keyword for definition of labels for force field types. Must have the following value:
'FFLABEL'.

2. AI,ATNO,MODIFD MS,CHARG,HYB,BND,CPK,#IH,#LP

FORMAT (A8,2X,I5,F10.5,F6.2,I4)

Definition of force field type label. Terminate with 'ADDED H', 'LONE PAIRS',
'Q EQUIL', 'GASTEIGER', 'DEL RE', or 'VDW A'.

Col.

1-8 AI = Label for atom type.

11-15 ATNO = Atomic number.

16-25 MODIFD MS = Modified atomic mass in atomic mass units (12.000 for ¹²C).

26-31 CHARG = Default atomic charge. Used in charge routines for nonbonded atoms.

32-36 HYB = Hybridization of atom (3 = sp³, 2 = sp², 1 = sp¹, 0 = other).

37-41 BND = Default maximum number of bonds to this atom type.

42-46 CPK = Default color used for CPK rendering of this atom type.

47-51 #IH = Number of implicit hydrogens associated with this atom.

52-56 #LP = Default number of lone pairs on this atom.

I.1.4 Implicit Hydrogen Relationships Parameters**1. ADDED H**

FORMAT (A7)

Keyword for section defining relationship between atom types differing in the number of implicit hydrogen atoms. Must have following value: 'ADDED H'.

2. BASE, HYDROGEN, 1IMPLCTH, 2IMPLCTH, 3IMPLCTH, ...

FORMAT (A8, 2X, A8, 2X, A8, 2X...)

Definition of relationship between atom types differing in the number of implicit hydrogen atoms.

Col.

1-10 BASE = Atom type with no implicit hydrogens (can be blank).

11-20 HYDROGEN = Hydrogen atom type to be used when adding hydrogens (*e.g.*, H₂).

21-30 1IMPLCTH = Atom type with one implicit hydrogen (can be blank).

31-40 2IMPLCTH = Atom type with two implicit hydrogens (can be blank).

:

I.1.5 Lone Pairs Parameters**1. LONE PAIRS**

FORMAT (A10)

Keyword for section defining lone pair types to add for atoms needing lone pairs. Must have the following value: 'LONE PAIRS'.

2. FFCODE, LPDEF

FORMAT (A8, 2X, A8)

Definition of lone pairs.

Col.

1-8 FFCODE = Atom type code.

11-18 LPDEF = Default lone pair type to be used when adding lone pairs.

I.1.6 Charge Equilibration Parameters**1. Q EQUIL**

FORMAT (A7)

Keyword for parameters for charge equilibration algorithm. Must have the following value: 'Q EQUIL'.

2. Element, QuanNo, MinQ, MaxQ, ElectNeg, Hardness, GenRadii

FORMAT (A2, 8X, 3I5, 3F10.5)

Parameters for charge equilibration algorithm.

Col.

1-2 Element = Atomic symbol (H through Lw are permissible).

- 11-15 QuanNo = Principle quantum number for element.
16-20 MinQ = Minimum permissible charge.
21-25 MaxQ = Maximum permissible charge.
26-35 ElectNeg = Atomic electronegativity (modified Mulliken values are provided in standard files).
36-45 Hardness = Atomic hardness (modified Mulliken values are provided in standard files).
46-55 GenRadii = Atomic radii (bond radii from stable molecules are used in standard files, e.g., crystal radii for metallic elements, or diatomic radii for halogens).

I.1.7 Del Re Charge by Element Parameters

1. DEL RE ATOMS

FORMAT (A12)

Keyword for parameters by element for Del Re charge algorithm. Must have the following value: 'DEL RE ATOMS'.

2. IElement,JElement,epsilon,gammaA(B),gammaB(A),delta(A),delta(B)

FORMAT (A2,8X,A2,8X,5F10.5)

Parameters by element for Del Re charge algorithm.

Col.

1-2 IElement = Atomic symbol for element I.

11-12 JElement = Atomic symbol for element J.

21-30 epsilon = See G. Del Re, *J. Chem Soc.*, 4031 (1958) and G. Del Re, B. Pullman and T. Yonezawa, *Biochim. Biophys. Acta*, 75, 153 (1963).

31-40 gammaA(B) = See G. Del Re references (above).

41-50 gammaB(A) = See G. Del Re references (above).

51-60 delta(A) = See G. Del Re references (above).

61-70 delta(B) = See G. Del Re references (above).

I.1.8 Del Re Charge by Force Field Parameters

1. DEL RE FF

FORMAT (A9)

Keyword for parameters by force field type for Del Re charges. These override any atom based parameters. Must have the following value: 'DEL RE FF'.

2. IFFCode,JFFCode,epsilon,gammaA(B),gammaB(A),delta(A),delta(B)

FORMAT (A2,8X,A2,8X,5F10.5)

Parameters by force field type for Del Re charges. These override atom based parameters for the specific force field type.

Col.

1-2 IFFCode = Force field code (label) for atom I.

11-12 JFFCode = Force field code (label) for atom J.

- 21-30 epsilon = See G. Del Re, *J. Chem Soc.*, 4031 (1958) and G. Del Re, B. Pullman and T. Yonezawa, *Biochim. Biophys. Acta*, 75, 153 (1963).
 31-40 gammaA(B) = See G. Del Re references (above).
 41-50 gammaB(A) = See G. Del Re references (above).
 51-60 delta(A) = See G. Del Re references (above).
 61-70 delta(B) = See G. Del Re references (above).

I.1.9 Gasteiger Charge Parameters

1. GASTEIGER

FORMAT (A9)

Keyword introducing parameters for Gasteiger's charge algorithm. Must have the following value: 'GASTEIGER'.

2. FFCODE,A,B,C,CHI

FORMAT (A8,2X,4F10.4)

Col.

1-8 FFCODE = Force field code.

11-20 A = See J. Gasteiger, *Tetrahedron*, 36, 3219 (1980).

21-30 B = See J. Gasteiger reference (above).

31-40 C = See J. Gasteiger reference (above).

41-50 CHI (optional). If CHI is not present, CHI is set to A + B + C.

I.1.10 Atomic Parameters

1. VDW A

FORMAT (A5)

Keyword for atomic parameters. Must have the following value: 'VDW A'.

2. AI,ITY,RNB,DENB,SCALE

FORMAT (A5,I5,4F10.5)

Atomic parameter records. Terminate with 'BOND', 'NONBO', 'AUTOT', or 'END'.

Col.

1-5 AI = Atom type.

6-10 ITY = Code for type of van der Waals potential.

ITY	Potential
1	Lennard-Jones 12-6
2	Exponential-6
3	Morse
4	Pure attractive R ⁻⁶
5	Pure repulsive exponential
6	No van der Waals
7	Lennard Jones 12-10

11-20 RNB = Van der Waals diameter in Å. For atoms *i* and *j* the program uses

$$(R_0)_{ij} = \frac{1}{2} (RE_i + RE_j)$$

21-30 DENB = Van der Waals "bond" energy in kcal/mol. For atoms *i* and *j* the program uses

$$(D_0)_{ij} = \sqrt{DE_i \times DE_j}$$

31-40 SCALE = Scaling coefficient for Morse, exponential-6, and exponential potentials. For atoms *i* and *j* the program uses

$$\gamma_{ij} = \frac{1}{2} \times (SCAL_i + SCAL_j)$$

I.1.11 Auto Type Parameters

1. AUTOT

FORMAT (A5)

Keyword for auto type parameters. Auto type parameters are used to assign atom type codes in the AUTO TYPE option under EDIT in the Builders (see Section 3.1.12). Not all parameter files have an AUTO TYPE section.

Col.

1-5 'AUTOT'

2. FFCODE,ELEMENT,HYBRIDIZATION,RING_SIZE,REQUIREMENTS,FACTOR

FORMAT (A5,8X,A2,4I12) Auto type parameter records. Terminate with 'BOND ' or 'NONBO'.

1-5 FFCODE = Auto atom type force field code (atom type code).

14-15 ELEMENT = Auto atom type element name.

16-27 HYBRIDIZATION = Auto atom type hybridization.

HYBRIDIZATION	Hybridization Type
3	sp3
2	sp2
1	sp1
0	other

28-39 RING_SIZE = Auto atom type ring size (used for special small ring atom types in MMP2 force field).

40-51 REQUIREMENTS = Number of "shell atoms" for this atom type (REQUIREMENTS records following current record contain parameters for atoms bonded to this atom type).

52-63 FACTOR = Not used.

I.1.12 Off-diagonal Nonbond Parameters

1. NONBO

FORMAT (A5)

Keyword for off-diagonal nonbond parameters. Not all parameter files have a 'NONBO' section. Must have the following value: 'NONBO'.

2. AI,AJ,TYPE,RVDW,DVDW,SCALE

FORMAT (A5,'-',A5,I5,3F10.5)

Off-diagonal nonbond parameter records. Terminate with 'BOND ' or 'END '.

1-5 AI = Atom type for atom i.

7-11 AJ = Atom type for atom j.

12-16 TYPE = Code for type of van der Waals potential to use for this interaction (see ITY above under entry 'VDW A').

17-26 RVDW = Van der Waals diameter to use for this interaction (see RNB above under entry 'VDW A').

27-36 DVDW = Van der Waals bond energy to use for this interaction (see DENB above under entry 'VDW A').

37-46 SCALE = Van der Waals scale parameter to use for this interaction (see SCALE above under entry 'VDW A').

I.1.13 Bond Stretching Parameters

1. BONDS

FORMAT (A5)

Keyword for bond stretching parameters. Must have the following value: 'BONDS'.

2. AI,AJ,TYPE,C,B,D,E

FORMAT (A5,1X,A5,I5,6F10.0)

Bond stretching parameter records. Terminate with 'ANGLE' or 'END '.

Col.

1-5 AI = Atom type for atom i.

7-11 AJ = Atom type for atom j.

12-16 TYPE = Code for type of bond potential.

TYPE	Potential
1	Harmonic
2	Morse
3	Cubic
6	Polyene

17-76 C,B,D,E = Parameters for bond potential (see Appendix G).

Harmonic:

C = K_b , the force constant in (kcal/mol)/Å².

B = R_0 , the bond distance in Å.

Morse:

C = K_b , the force constant in (kcal/mol)/Å².

B = R_0 , the bond distance in Å.

D = D_0 , the bond energy in kcal/mol.

Cubic:

C = K_b , the force constant in (kcal/mol)/Å².

B = R_0 , the bond distance in Å.

D = Cubic correction term.

Polyene:

C = K_2 , the force constant in (kcal/mol)/Å² for the double bond case.

B = R_{20} , the bond distance for the double bond case in Å.

D = R_{10} , the bond distance for the single bond case in Å.

E = K_1 , the force constant in (kcal/mol)/Å² for the single bond case.

I.1.14 Bond Angle Bending Parameters

1. ANGLE

FORMAT (A5)

Keyword for bond angle bending parameters. Must have the following value: 'ANGLE'.

2. AI,AJ,AK,TYPE,C,B,D,E,F,G,H,P

FORMAT (A5,1X,A5,1X,A5,I5,9F10.4)

Bond angle bending parameter records. Terminate with 'TORSI' or 'END '.

Col.

1-5 AI = Atom type for atom i.

7-11 AJ = Atom type for atom j.

13-17 AK = Atom type for atom k.

18-22 TYPE = Code for type of bond angle bending potential.

TYPE	Potential
1	Simple cosine expansion
x4	Cosine periodic
x6	Polyene
11	Cosine expansion with angle stretch terms
21	Simple harmonic expansion
24	Schleyer potential
31	Harmonic expansion with angle stretch terms
33	MM2 potential
34	MM2 potential with adjustment of default angle for H atoms
35	MM2 potential

23-112 (All types but polyene) C,B,D,E,F,G,H,P = Parameters for bond angle bending potential (see Appendix G).

C = K_{θ} , the bending force constant in (kcal/mol)/radian². Note that the units for the force constant differ for different potential functions. Units are given for the harmonic expansion (TYPE = 21 or 31).

B = θ_0 , the equilibrium bond angle in degrees; the linear minimum for cosine periodic potential.

D = Force constant for first angle stretch term in (kcal/mol)/(Å-radian).

E = Force constant for second angle stretch term in (kcal/mol)/(Å-radian).

F = Force constant for stretch-stretch term in kcal/(mol-Å²).

G = Equilibrium angle correction factor for first hydrogen bonded to central atom (MM2 potential).

H = Equilibrium angle correction factor for second hydrogen bonded to central atom (MM2 potential).

P = Period for cosine periodic potential

23-112 (Polyene type) C,B,D,E,F,G,H,P = Parameters for bond angle bending potential for polyene type (see Appendix G).

C = K_{21} , bending force constant for double bond/single bond case.

B = θ_{21} , bond angle in degrees for double bond/single bond case.

D = Force constant for angle stretch term for double bond/single bond case, in (kcal/mol)/(Å-radian).

E = Force constant for angle stretch term single bond/single bond case (kcal/mol)/(Å-radian).

F = Force constant for stretch-stretch term in kcal/(mol-Å²).

G = K_{11} , bending force constant for single bond/single bond case.

H = θ_{11} , bond angle in degrees for single bond/single bond case.

I.1.15 Dihedral Angle Torsion Parameters**1. TORSI**

FORMAT (A5)

Keyword for dihedral angle torsion parameters. Must have the following value: 'TORSI'.

2. AI,AJ,AK,AL,CASE,BARRIER,PERIOD,CISMIN(1)

FORMAT (A5,1X,A5,1X,A5,1X,A5,I5,3F10.4)

Dihedral angle torsion parameter records. Terminate with 'INVER' or 'END'.

Col.

1-5 AI = Atom type for atom i.

7-11 AJ = Atom type for atom j.

13-17 AK = Atom type for atom k.

19-23 AL = Atom type for atom l.

24-28 CASE = Sequence number. Several cosine terms may occur for a particular torsion potential. Each will have the same sequence number. The program does *not* use ITYP.

29-58 BARRIER,PERIOD,CISMIN(1) = Parameters for dihedral angle-torsion potential.

BARRIER = $K\phi_n$, the barrier in (kcal/mol).PERIOD = n , periodicity of the potential ($n = 1, 2, 3, 4, 5, 6$).

CISMIN(1), case (+1 if cis geometry is energy minimum, -1 if cis geometry is energy maximum).

I.1.16 Inversion (Improper) Torsion Parameters**1. INVER**

FORMAT (A5)

Keyword for inversion (improper torsion) parameters. Must have the following value: 'INVER'.

2. AI,AJ,AK,AL,TYPE,FRC CNST,EQU ANGL,D,E,F

FORMAT (A5,1X,A5,1X,A5,1X,A5,I5,5F10.0)

Inversion (improper torsion) parameter records. Terminate with 'HBOND' or 'END'.

Col.

1-5 AI = Atom type for atom i.

7-11 AJ = Atom type for atom j.

13-17 AK = Atom type for atom k.

19-23 AL = Atom type for atom l.

24-28 TYPE = Code for type of inversion potential (see Appendix G for explanations).

29-78 FRC CNST,EQU ANGL,D,E,F = Parameters for inversion potential.

FRC CNST = K_ω , the force constant in (kcal/mol)/radian².EQU ANGL = ω_0 , the bond angle in degrees.

TYPE	Potential
1	CHARMM
2	Normal ('umbrella') inversion
3	AMBER

D = Periodicity (AMBER inversion potential) or quartic factor (umbrella potential with planar equilibrium angle).

E = Not used.

F = Not used.

I.1.17 Hydrogen Bond Parameters

1. HBOND

FORMAT (A5)

Keyword for hydrogen bond parameters. Must have the following value: 'HBOND'.

2. AI,AJ,TYPE,DE HB,RE HB

FORMAT (A5,1X,A5,I5,2F10.4)

Hydrogen bond parameter records. Terminate with end-of-file ('END ') or 'USER'.

Col.

1-5 AI = Atom type for atom i.

7-11 AJ = Atom type for atom j.

12-16 TYPE = Code for type of hydrogen bond potential.

TYPE	Potential
1	Lennard-Jones 12-10 (default)
2	Morse

17-36 DE HB, RE HB = Parameters for hydrogen bond potential.

DE HB = $(D_0)_{ij}$, hydrogen bond energy in kcal/mol.

RE HB = $(R_0)_{ij}$, hydrogen bond distance in Å.

I.1.18 User-supplied Parameters

1. USER

FORMAT (A5)

Keyword for user parameters. Not present unless a user routine to read parameters has been linked into the site copy of the program. Must have the following value: 'USER '.

2. User Parameter Records

Contents and format determined by user routine. Terminate with end-of-file ('END ').

I.2 Conversion File Format

The default Dreiding-I and Dreiding-II conversion files are ASCII files which consist of comments and conversion records. Comments are indicated by a 'C' or an '*' in column one. The format for the conversion record is presented below.

RES,OLDTYP,NEWTYP,MB,NL,CHARGE

FORMAT (1X,A3,1X,A7,1X,A7,1X,I2,1X,I2,1X,F8.4)

Record for converting residue and atom label into an atom type.

Col.

2-4 RES = Residue name.

6-12 OLDTYP = Atom label (*e.g.* C8, N3 from Brookhaven file).

14-20 NEWTYP = Atom type (force field type, *e.g.* C_3, N_2).

22-23 MB = Absolute value equals maximal number of covalent bonds. If MB is negative then the atom will be displayed with a 6-point star even if it is connected to other atoms.

25-26 NL = Number of lone pairs. *Note:* If missing hydrogens are added, the number of hydrogens, NH, added will be

$$NH = MB - NL - NB$$

where NB is the number of existing bonds.

28-35 CHARGE = Atomic charge.

The conversion file allows the user to assign different charges, different numbers of lone pairs and different maximal numbers of covalent bonds to atoms with the same atom type.

The default Dreiding-I and Dreiding-II conversion files contain conversions for Brookhaven atom labels and residue names, for CHM and MOL atom labels, and for default BioDesign atom labels. The user is encouraged to examine this file (see Table I.1 for file name), which serves as a guide for adding new conversion records. The charges on the standard amino and nucleic acids are those reported in the CHARMM paper: B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus, *Journal of Computational Chemistry*, 187-217 (1983). The hydrogen charges have been collapsed onto the heavy atom.

In addition conversion files are provided for the MM2, MMP2, and AMBER force fields.

I.3 The Dreiding-I and Dreiding-II Force Fields

I.3.1 Atom Types

The force field parameters are assigned according to a 5-character atom type, the format of which is presented in Table I.2 below.

Char.	Meaning
1-2	Chemical Symbol - left justified
3	Hybridization (3 - tetrahedral, 2 - trigonal, 1 - linear, R - resonant/aromatic)
4	Number of bound implicit hydrogens
5	Alternate parameters indicator (alphabetic)

Table I.2: Dreiding-I and II atom type format

The Dreiding-II atom types are used as the default. Dreiding-II has 60 atom types which include 28 different elements. They are listed in Table I.3 on the following page.

Hydrogen	H__A, H_, H__b
Boron	B_3, B_2
Carbon	C_34, C_33, C_32, C_31, C_3, C_22, C_21, C_2, C_R2, C_R1, C_R, C_11, C_1
Nitrogen	N_33, N_32, N_31, N_3, N_22, N_21, N_2, N_R2, N_R1, N_R, N_1
Oxygen	O_32, O_31, O_3, O_2, O_R1, O_R, O_1
Fluoride ion	F_
Aluminum	Al3
Silicon	Si3
Phosphorus	P_3
Sulfur	S_31, S_3
Chloride ion	Cl
Gallium	Ga3
Germanium	Ge3
Arsenic	As3
Selenium	Se3
Bromide ion	Br
Indium	In3
Tin	Sn3
Antimony	Sb3
Tellurium	Te3
Iodine ion	I_
Sodium ion	Na
Calcium ion	Ca
Titanium	Ti
Iron	Fe
Zinc	Zn
Technitium	Tc
Ruthenium	Ru

Table I.3: Dreiding-II atom types (default)

Appendix J

Dynamics Algorithms

Newton's equation is

$$M_I \frac{d^2 R_{\alpha I}}{dt^2} = F_{\alpha I} \quad (1)$$

where M_I is the mass of particle I
 $R_{\alpha I}$ is the α component (x,y,z) of the position of particle I
 $F_{\alpha I}$ is the α component of the force on I due to all other particles
 t is the time.

J.1 Unit of Time

The following units are used by the program

$R_{\alpha I}$: Angstroms = 10^{-8} cm
 $F_{\alpha I}$: (kcal/mol)/Å
 M_I : atomic mass units (^{12}C has mass 12.0000 amu)

As a result, the unit of time is

$$t_0 = 0.04888742 \text{ ps} = 48.88742 \times 10^{-15} \text{ sec} \quad (2)$$

J.2 Standard Verlet

Eschewing subscripts Newton's equation (1) becomes

$$\ddot{\mathbf{R}} = \mathbf{F}(\mathbf{R}) \quad (3)$$

where $\mathbf{F}(\mathbf{R}) = \frac{1}{M} \mathbf{F}_{\alpha I}$ and $\ddot{\mathbf{R}} = \frac{d^2 \mathbf{R}}{dt^2}$

are continuous functions of time. In a numerical calculation we consider discrete times

$$t_n = nh$$

where t_n is the time at the nth step, and
 h is the time increment.

The time increment is generally taken as $h = 0.002$ ps for systems with hydrogens. The coordinates and derivatives are defined as

$$\begin{aligned} R_n &= R(t_n) \\ \dot{R}_n &= \dot{R}(t_n) \\ F_n &= F(t_n) \end{aligned} \tag{4}$$

J.2.1 2 Point Verlet

From (4)

$$\dot{R}_n \approx [R_{n+1} - R_{n-1}] / 2h \tag{5}$$

$$\ddot{R}_n \approx [R_{n+1} - 2R_n + R_{n-1}] / h^2 \tag{6}$$

Using (6) in (3) leads to the Verlet equation¹

$$R_{n+1} = 2R_n - R_{n-1} + h^2 F_n \tag{7}$$

Thus, given coordinates at steps n and $n-1$ and forces at step n , we calculate coordinates at step $n+1$. In this approximation we define the velocity as

$$v_n = [R_{n+1} - R_{n-1}] / 2h \approx \dot{R}_n \tag{8}$$

J.2.2 1 Point Verlet

To initialize Verlet, we start with the coordinates and velocities at time $n=0$: R_0 and v_0 . [If no previous dynamics is available, the v_0 for various particles are chosen from a Maxwell-Boltzman distribution]. In this case (7) and (8) become

$$R_1 = 2R_0 - R_{-1} + h^2 F_0 \tag{9}$$

$$v_0 = [R_1 - R_{-1}] / 2h \tag{10}$$

Since $R_{-1} = R_1 + 2hv_0$, (9) becomes

¹ L. Verlet, *Phys. Rev.* **159** (1967) 98-103.

$$R_1 = R_0 + hv_0 + \frac{1}{2}h^2F_0 \quad (11)$$

Standard Verlet dynamics uses the 1 point relation (11) for the initial step of a trajectory and switches to the 2 point formula (7) at all subsequent steps. Except for numerical errors, (8) or (11) will lead to energy conservation and we refer to this as *microcanonical dynamics*.

J.3 Summed Verlet

A modification of Verlet dynamics is to define

$$v_{n+\frac{1}{2}} = (R_{n+1} - R_n)/h \quad (14)$$

as the velocity at time step $n+\frac{1}{2}$. Then

$$v_{n+\frac{1}{2}} - v_{n-\frac{1}{2}} = (R_{n+1} + R_{n-1} - 2R_n)/h = h\ddot{R}_n = h_nF_n$$

This leads to the equations

$$v_{n+\frac{1}{2}} = v_{n-\frac{1}{2}} + hF_n \quad (15)$$

$$R_{n+1} = R_n + hv_{n+\frac{1}{2}} \quad (16)$$

which are referred to as the *summed Verlet*² or *leapfrog*³ equations. Each iteration starts with the velocities at $n-\frac{1}{2}$ and the coordinates at R_n and evaluates the forces at R_n . Equation (15) is used to update the velocity to $v_{n+\frac{1}{2}}$ after which (16) is used to update the coordinates to R_{n+1} .

If there are no initial velocities, the initial set (from a Boltzman distribution) can be considered as $v_{-\frac{1}{2}}$ which is used with R_0 to start the iteration.

In order that the kinetic energy and potential energies be evaluated at the same time step, we define

$$v_n = \frac{1}{2}(v_{n+\frac{1}{2}} + v_{n-\frac{1}{2}})$$

and use v_n in (12) to calculate the temperature at step n .

² W. C. Swope *et al.*, *J. Chem. Phys.* **76** (1982) 637-649.

³ Berendsen *et al.*, *J. Chem. Phys.* **81** (1984) 3684-3690.

If the velocities are to be scaled (13), the scaling can be applied to the final velocity

$$\frac{v_{n+\frac{1}{2}}^{\text{new}}}{v_{n+\frac{1}{2}}^{\text{old}}} = \sqrt{\frac{T_0}{T}}$$

In this case, the coordinate update step (16) is redone using the new $v_{n+\frac{1}{2}}$ so that the coordi-

nates of R_{n+1} are consistent with the velocities $v_{n+\frac{1}{2}}^{\text{new}}$.

Summed Verlet and standard Verlet are mathematically equivalent. However, summed Verlet is slightly more accurate, less sensitive to numerical error (Swope *et al* ²) and more efficient when velocities must be adjusted. Therefore, the program uses the summed Verlet algorithm to solve the equations of motion.

Appendix K

Command and Macro Modes

Because the command and macro modes are intended for experienced users, this appendix is relatively terse. The beginning sections describe operating the program in command mode, while the later sections describe the construction and use of macros. This appendix does not address either command or macro philosophy and in general has been provided as a reference for command syntax (see Sample Macros, Section K.4 for examples).

K.1 Command Mode

Command mode can be initiated using several different methods. The most direct method is to click the COMMAND button in the Modules/Messages window after the program has started. Alternatively, command mode can be specified upon start up by either adding a "command mode" line to the defaults file (L_CMND YES) or by specifying a start up macro on program invocation (*e.g.*, `biograf test.macro`). When the program is in command mode a command prompt appears in the text window. The command prompt always indicates the current menu level by including the name of the active Functions menu.

Terminating command mode (*i.e.*, returning to menu mode) can be achieved by entering one of the following case sensitive commands after the command prompt:

```
menu  
quit  
stop
```

K.1.1 Menu Commands

Aside from special commands and macros, command mode is essentially a vehicle that allows the user to exchange typing for picking. In its most elementary form, command mode operation only entails typing the desired menu item after the command prompt. For example, instead of picking the menu item labeled ANALYZE, type the command `analyze` after the command prompt (*Note:* all commands related to menu items are case insensitive). In general, anything

that could have been picked can now be typed. Although this initially sounds trite, the power of the method will become more obvious when we consider the special commands and macros.

K.1.2 Multiple Commands

Up to 25 commands can be entered on a single line (*e.g.*, `build; organic`). Commands on the same line must be separated by a semicolon. If more than one command is entered on a line, the extra commands are placed in a command buffer; additional commands that are required are taken from the command buffer until it is empty at which point the command prompt will reappear.

K.1.3 Continuation Characters

Command lines ending with either the backslash "\ " or dash "- " will cause a continuation prompt to appear on the next line. As many continuations as needed can be used to specify either single or multiple commands.

K.1.4 Abbreviations and Wild Cards

Abbreviations are allowed to the extent that they do not cause ambiguities (*e.g.*, on the IDENTIFY menu under ANALYZE, `atom type` and `atom t` are good commands, but `atom` is an ambiguous command because it could also mean `atom label`, `atom charge`, *etc.*). If ambiguities are present the program will print a warning listing the possible choices. Ambiguous commands also cause the command buffer to be emptied. In addition to simple abbreviations, long commands can be shortened by using "*" for the bulk of the command (*e.g.*, `utilities` can be abbreviated as `u*s`). The "*" wild card is especially useful for commands like `sng f/atm` which can be shortened to `s*f`. Only one wild card can be used per command.

K.1.5 Ambiguous Commands

In reference to ambiguous commands, there are some situations that will always yield ambiguity. For example if two input files have the same file descriptor called `oxidase`, then trying to select one of these files will always generate the warning message associated with ambiguous commands. Situations like these are best avoided by making an effort to keep all group and file names unique. However, this cannot always be done and the symbol "#" can be used to tell the program which item is actually required. In the case above the first file can be referenced as `oxidase#1` and the second file can be referenced as `oxidase#2`.

Ambiguities can also occur when group and function names are equivalent (*e.g.*, the `organic` group and the `organic` function on the BUILD menu). For these cases the group name takes priority; however, the function name can be referenced by using the Functions menu name and slash "/" (*e.g.*, `build/organic`).

K.1.6 Two-State and Cyclic Toggles

There are two additional types of ambiguities that need to be addressed. The first is for two-state toggles like `constrnts` and the second is for cyclic toggles like `method` (minimization method). Cyclic toggles are discussed in the Special Commands Summary section.

The two-state toggle generally presents a problem when entering multiple commands and in macro mode. For example, from the `SIMULATE` menu, assume we want to change the energy variables so that H bond terms are included in the energy expression and the constraint terms are excluded. We might type `simulate; defaults; energy var; h bonds; constraints`. Since we may not know the toggle states of `h bonds` and `constraints` we may or may not get the desired result. The "on" and "off" commands are used to circumvent this problem. The "on" command forces a two state toggle to have the on state, while the "off" command forces a two state toggle to have the off state. Both of these commands take an argument list which allows more than one toggle state to be set simultaneously. For the above example we guarantee the correct result by typing `simulate; defaults; energy var; on=(h bonds); off=(constraints)`. Alternatively, the ">" and "<" symbols could be used (e.g., `constraints<` forces the constraints toggle off, while `h bonds>` forces the H bonds toggle on).

K.1.7 No-opt Commands

The "." command is a no-opt ("no operation", a null function) that is most useful for delimiting lists when more than one command is being entered on a command line. For example, building a tripeptide consisting of alanine, glycine and leucine from the l-amino acid library could be done with the following command `l_lib; ala_1; gly_1; leu_1; .` as opposed to the two following command lines `l_lib; ala_1; gly_1; leu_1;` and `<cr>` -- the carriage return delimits the list as did the "." symbol on the single command line. `<cr>` is the same as picking off of an active menu when delimiting lists. Typing `<cr>` in any other context will only cause the command prompt to echo to the terminal.

K.1.8 Atom Referencing Commands

When atom data is required by the program while in command mode, the ATM REF (atom reference) prompt will appear (see "pick=state" command in Special Commands Summary for mouse/tablet atom picking). Every atom in the program can be addressed either directly or indirectly. To reference an atom directly type `anum/n` where `n` is the atom's absolute sequence number (including consideration for multiple data files).

Indirect referencing involves describing the atom by its file name, chain designator, residue name, residue number, and/or atom label. Any combination of the previous data is allowed. However, the priority of file name (highest priority), chain designator, residue name, residue number, and atom label (lowest priority) is always maintained. File name is specified like `file/string` where `string` is the required file name; chain designator is specified like `chain/string` where `string` is the required chain designator; residue name is specified like `rnam/string` where `string` is the required residue name; residue number is specified like

rnum/n where n is the required residue number; and, the atom label is specified like anam/string where string is the atom label. Additionally, atomid/options can be used where options is any combination of file, residue name, residue number, and/or atom label in any order.

The program will loop through its list of atoms while considering the information provided by the user and will return the *first* atom that satisfies the specified selection criteria. If no atoms meet the specified criteria, then an atom number of zero will be returned (this is equivalent to picking in an area that has no atoms). Combinations of selectors are constructed by separating the individual items with commas. There is no order requirement since the program understands the priority rules. The atom referencing commands are case sensitive and cannot be abbreviated. Examples of atom reference commands :

anum/10	the atom with sequence number 10.
file/leuenk1,rnam/tyr,anam/ca	the first C-alpha carbon atom of the first TYR residue of the LEUENK1 input file.
rnum/9	the first atom of residue 9 of the first input data file that has a residue numbered 9.
atomid/cb phe 17	the C-beta carbon atom of PHE-17 of the first input data file.

K.1.9 Dial Commands

Dial operations can also be issued in command mode (in addition to using a dials box). This feature is actually very useful for animation (especially in a macro/demo) and for generating stereo pictures.

The syntax is "dial value/n" where dial is either "xrot", "yrot", "zrot", "zoom", "xtran", "ytran", "ztran", or "depth"; value is the dial value (in degrees for rotations) and n is the number of increments the value is to be divided into. For example, the command yrot 360/10 will cause a series of 10 y-rotations each of 36.0 degrees for a total of 360.0 degrees. The increment number is optional so zrot 122.4 will cause a z-rotation of 122.4 degrees in one step.

Alternatively each dial can be referenced like "dialn" where n is the dial number (e.g., dial3 360.0/10 would issue a 360.0 degree rotation in 10 increments to what ever was connected to dial3 -- usually z-rotation).

The dial commands are case sensitive and cannot be abbreviated.

K.1.10 Special Commands Summary

The following special commands are all case sensitive and cannot be abbreviated. Also note that the special commands (and the commands presented above) constitute a list of reserved words that users should avoid for group and file names.

!!	This command is used to repeat the last command that was issued in a manner similar to that used in UNIX operating systems.
!string	This command is used to repeat a command in the history stack that starts with string in a manner similar to that used in UNIX operating systems. For example if the history stack looked like <pre> 4 analyze; identify; group; atom type 3 yrot 360/200 2 zoom 100/10 1 xrot 180/100 </pre> and the !yr command were issued, the program would execute command 3 -- "yrot 360/200" (!3 would also produce the same result).
@	This command is used to invoke macros. See documentation below for details.
^	This command is used to return to the entry-level menu from any level submenu and is generally used in combination with "save" and "restore" to move throughout the menu structure quickly.
autopause=state	This command is used to halt the program after every command when the input state is true. The default state is false.
comment "string"	This command is used to place a comment record in the logfile.
cycle	This command is used to set the state of multi-state toggles and is almost exclusively used in conjunction with macro files. The complete syntax is: <pre>cycle=target, state1, state2, ..., stateN</pre> where "target" is the state in which the toggle is to be left, "state1,state2,..." are all of the possible states of the toggle including the "target" state. Generally a macro is constructed in which the "target" state is supplied as input to the macro.

debug=state	This command is used to set the debug state for macro execution. If the input state is <code>true</code> , then the macro debugger will be invoked during macro execution. The default state is <code>false</code> .
echo "string"	This command is used to echo a character string to the terminal. The double quotes are only required when <code>string</code> contains special characters.
entry-level	Same as <code>"^"</code> above.
EXIT	This command is used to exit the program.
history	This command is used to display the command history in a manner similar to that used in UNIX operating systems. The history stack is 50 commands deep.
logmacro=state	This command is used to set the macro logging state. If the input state is <code>true</code> , then commands issued by macros will be logged as any other command. If the input state is <code>false</code> , then only the reference to macros will be included in the logfile. The default state is <code>true</code> .
menu, quit, stop	These commands are used to exit command mode and return to menu mode.
pause	This command is used to halt the program until the user issues a carriage return. It is generally used in combination with macro/demo files.
pick=state	This command is used to set the atom picking mode. If the input state is <code>true</code> , then atom picking defaults to the mouse or tablet. If the input state is <code>false</code> , then atom picking defaults to command mode via atom referencing. The default state is <code>false</code> .
refresh	This command is used to refresh the graphics screen on raster devices like the Titan and the IRIS.
restart	This command is used to restart the program (like PROGRAM RESET in the Utilities menu).
restore	This command is the second half of the "save" command (see below). If more than one state is saved, "restore" will restore the most deeply nested state (<i>i.e.</i> , last in, first out).
save	This command saves the current menu state of the program and in combination with the restore command allows the user to jump from place-to-place in a convenient manner (<i>e.g.</i> , while in the Mechanics menu a macro could be invoked that issues a "save", jumps into the

In/out menu, writes out the current data files, and finally issues a "restore" command to return to the Mechanics menu).

`trace=state`

This command is used to turn macro tracing on and off. If the input state is `true`, then macro commands will be echoed to the screen during macro execution. The default state is `false`.

K.2 Macros

The use of macros greatly expands both the command mode capability as well as the capability of the entire program. In their simplest form, macros can save both time and the number of keystrokes needed to complete a repetitive task. In their full glory, macros can be used to create a personalized command syntax that is customized for a particular user, group, or company. However they are used, macros are fun.

In general, a macro file contains a header record and an end record with commands sandwiched in between. Optionally a macro name can be specified so that multiple macros can be placed in the same file. If one macro in a file contains a macro name, then *all* macros in that file must be named.

```
beginmacro macro_name
    commands
endmacro
```

To start a macro use the "@" command along with the name of the macro file. If the file contains named macros, then specify the name of the macro after the file name by enclosing the macro name in parentheses. No spaces are allowed between the file name and the macro name. If the macro file name looks like "file_name.macro", then only "file_name" is required on the command line.

```
@file_name(macro_name) variable_assignments
```

Assignment of variables on the command line is optional and is explained below. The file finder first looks in the startup directory for the macro file and then looks in *bg_macro* if it does not find it in the local directory. If required, the full path name of the file can be specified.

Both continuation characters, "\" and "-", and comments are allowed in macros. The comment character is "%" and it can be placed anywhere on the command line. Everything after the comment character is ignored by the macro parser.

Nesting of macros is permitted to a level of five. That is, a macro can call a macro, which can in turn call another macro, and so on until five macros are simultaneously open.

The following sections describe the features available in macro mode.

K.2.1 Variables

Currently each of the five possible macros can use up to a maximum of 20 local variables (not including any of the special predefined variables). In addition to the 20 local variables, 20 global variables can be used. The following subsections describe how these variables are used.

Variable Names and Values

The name of a variable can contain up to 20 alphanumeric characters. The use of special characters in variable names is not recommended. If a variable name has at least one lower case alphabetic character, then it is considered to be a local variable. Otherwise it is treated as a global variable. The value of a local variable only has meaning in the context of the macro in which it is created. The value of a global variable has meaning in all of the currently active macros. Variables do not have to be defined since they are automatically defined when they are first used. The value of a variable is essentially unrestricted except that the total length must be less than or equal to 1000 character equivalents. The use of commas "," in variable values indicates a separation of elements. For example in

```
molecule = main chain, side chain
```

the variable "molecule" contains two elements ("main chain" and "side chain").

Assignment of Variables

Variables can be assigned either on the command line that initiates a macro or within the macro itself. Assignment of a variable within a macro follows the syntax

```
variable_name = value
```

where "variable_name" is the name of the variable to be assigned and "value" is the value to be assigned to the variable. The assignment of a variable on a macro initiation command line may take several forms. The following example shows the most direct method:

```
@macro_name value1 value2 ... valueN
```

In this case, the variable values are separated by blank spaces and the variables names are the default input variables input1, input2, ..., inputN. The following example illustrates the use of specific variable names:

```
@macro_name variable_name1 = value1 variable_name2 = value2
```

Even when specific variable names are used, the default input variables are still assigned. If the variable values have special characters, blanks, or multiple elements then the use of parentheses is recommended:

```
@macro_name variable_name1=(element1,element2,element3)
```

Referencing Variables

Variables are referenced by using the following notation:

```
${variable_name}
```

where "variable_name" is the name of the required variable. The macro parser will substitute the value of the variable directly in place of the value reference. If the variable has more than one element and a specific element is required, then use the following extension to the above notation:

```
${variable_name}[element_number]
```

where "element_number" is an integer. If a variable has not been assigned a value or a reference is made to an unassigned element, then the value returned is NULL.

Action Variables

There are special variables called "action variables" that return values by causing the user to take some specified action. Currently only one (prompt) has been implemented.

<code>\${prompt}</code>	This variable reference prints a prompt to the terminal and waits for user input. The value it returns is the terminal input. If there is no terminal input, then the value returned is NULL. The prompt message can be specified by assigning the prompt variable in the following manner: <code>prompt=string1, ..., stringn</code> (where <code>string1</code> through <code>stringn</code> are the prompt strings). The default prompt string is "Prompt:".

Other Macro Variables

The following global variables can be referenced using the `${variable_name}` notation described above.

FIRSTGROUP	This variable is used to request the first used group. And like most of the variables in this section, it can be appended with "+n" and "-n" to indicate an offset, where n is an integer (e.g., <code>\${FIRSTGROUP+2}</code> , will yield Group 3 if the first used group is group 1).
NEXTGROUP	This variable is used to request the next unused group. It can also be appended with "+n" and "-n" to indicate an offset as described above.
LASTGROUP	This variable is used to request the last used group. It can also be appended with "+n" and "-n" to indicate an offset as described above.

ALLGROUPS	This variable is used to request all groups.
FIRSTFILE	This variable is used to request the first used file. It can also be appended with "+n" and "-n" to indicate an offset as described above.
NEXTFILE	This variable is used to request the next unused file. It can also be appended with "+n" and "-n" to indicate an offset as described above.
LASTFILE	This variable is used to request the last used file. It can also be appended with "+n" and "-n" to indicate an offset as described above.
ALLFILES	This variable is used to request all files.

K.2.2 Inquire and Set Functions

The inquire and set functions provide easy access to the program's shared data structure. Most of the inquire functions require a variable name ("variable_name") and an absolute atom number as input. If, however, an atom number is not required, then do not include one in the function reference. If more than one atom number is required, then include it after the first atom number. The set function, in addition to a variable name and an atom number, requires a value for the variable. If more than one value is required (*e.g.*, "coordinates" needs three values), then separate each value with a comma ",".

```
var = inq(variable_name,atom_number)
var = set(variable_name,atom_number,value)
```

Below is a list of the currently supported variables for the inquire and set functions.

active_file	energy_rms_force
atom_label	energy_torsions
bond_order	energy_total
bonds	energy_user
cell_parameters	energy_vdw
cell_stresses	force_field_label
chain	lone_pairs
charge	mass
coordinates	max_bonds
electronegativity	move_status
energy_angles	number_of_atoms
energy_bonds	residue_name
energy_constraints	residue_number
energy_delta	rms_angles
energy_electrostatic	rms_bonds
energy_hbonds	rms_coordinates
energy_inversions	rms_dihedrals

The inquire function returns the value of the item being considered. If an error occurs during the inquiry, "NULL" is returned. The set function returns a completion status: "1" for successful or "0" for unsuccessful completion.

K.2.3 Math Functions

Most of the standard math functions are supported in macro mode. Only one math function can be used per command. The results of functions that generate real numbers are stored with 10 decimal places of precision. In the examples below, "var" is a variable name and "v1" and "v2" are integers or real numbers.

<code>var = v1 + v2</code>	The addition symbol must occur with at least one blank on each side.
<code>var = v1 - v2</code>	The subtraction symbol must occur with at least one blank on each side.
<code>var = v1 * v2</code>	The multiplication symbol must occur with at least one blank on each side.
<code>var = v1 / v2</code>	The division symbol must occur with at least one blank on each side.
<code>var = sin(v1)</code>	The sin function requires input in degrees.
<code>var = asin(v1)</code>	The arc sin function returns the value in degrees.
<code>var = cos(v1)</code>	The cosine function requires input in degrees.
<code>var = acos(v1)</code>	The arc cosine function returns the value in degrees.
<code>var = tan(v1)</code>	The tangent function requires input in degrees.
<code>var = atan(v1)</code>	The arc tangent function returns the value in degrees.
<code>var = exp(v1)</code>	The exponentiation function yields the result of e^{**v1} .
<code>var = pow(v1, v2)</code>	The power function yields the result of $v1^{**v2}$.
<code>var = log10(v1)</code>	The log base 10 function yields the common logarithm of v1.
<code>var = log(v1)</code>	The log function yields the natural logarithm of v1.
<code>var = sqrt(v1)</code>	The square root function yields the square root of v1.
<code>var = ran(seed)</code>	The random number function yields a random number.
<code>var = degrad(v1)</code>	The degrees to radians function converts v1 from degrees to radians.
<code>var = raddeg(v1)</code>	The radians to degrees function converts v1 from radians to degrees.
<code>var = abs(v1)</code>	The absolute value function yields the absolute value of v1.
<code>var = float(v1)</code>	The float function converts v1 to a real number.
<code>var = fix(v1)</code>	The fix function converts v1 to an integer.
<code>var = min(v1, v2)</code>	The min function returns the minimum value of v1 and v2.

`var = max(v1, v2)` The max function returns the maximum value of v1 and v2.

`var = mod(v1, v2)` The mod function returns the modulus of v1 and v2.

K.2.4 Logical Expressions

Logical expressions are used in both "if..then..else" blocks and in "do while" loops. Support is provided for the following expressions:

<code>val1 = val2</code>	Returns true if value 1 ("val1") and value 2 ("val2") are equal.
<code>val1 <> val2</code>	Returns true if value 1 ("val1") and value 2 ("val2") are not equal.
<code>val1 > val2</code>	Returns true if value 1 ("val1") is greater than value 2 ("val2").
<code>val1 >= val2</code>	Returns true if value 1 ("val1") is greater than or equal to value 2 ("val2").
<code>val1 < val2</code>	Returns true if value 1 ("val1") is less than value 2 ("val2").
<code>val1 <= val2</code>	Returns true if value 1 ("val1") is less than or equal to value 2 ("val2").
<code>val1 ~ val2</code>	Returns true if value 1 ("val1") is similar to value 2 ("val2") (i.e., if the characters of value 1 are completely contained in value 2 or if the characters of value 2 are completely contained in value 1).

Certain isolated values also provide legal expressions:

<code>0</code>	The number zero returns false. All other numbers return true.
<code>false</code>	The character string "false" returns false.
<code>FALSE</code>	The character string "FALSE" returns false.
<code>true</code>	The character string "true" returns true.
<code>TRUE</code>	The character string "TRUE" returns true.

In addition to the above expressions, the boolean operations "and" and "or" are supported.

<code>exp1 & exp2</code>	Returns true if both expression 1 ("exp1") and expression 2 ("exp2") are true.
<code>exp1 exp2</code>	Returns true if either expression 1 ("exp1") or expression 2 ("exp2") is true.

More than one "&" or "|" can be used at a time, but "&" and "|" cannot be used together.

K.2.5 If..then..else Blocks

The if..then..else block is used to control the execution of macro commands based on the result of a logical expression. Support for "elseif" is provided along with support for unlimited nesting.

```

if( logical_expression )then
    commands
elseif( logical_expression )then
    commands
else
    commands
endif

```

The "logical_expression" is of the form described in the above subsection. Of course, "elseif" and "else" are optional. In addition, a short hand "if" is also available:

```

if( logical_expression )command

```

K.2.6 Branching and Statement Labels

Branching to a statement label is supported with the "go to statement_label" command:

```

go to statement_label

```

where "statement_label" is a label comprised of no more than 20 alphanumeric characters ending with a semicolon ":":

```

statement_label:

```

The command line containing the statement label must appear by itself on that line. Note that "go to" and "goto" are equivalent. Up to 20 statement labels can be used for each macro.

K.2.7 Loop Constructs

Several loop constructs are available in macro mode: "for variable_name in (list)", "do while (expression)", "do until (expression)", and "do variable_name = start,stop,increment". Each of these will be considered in turn.

The "for" loop is used to incrementally assign a variable with a value from a list of possible values. The list of values is a comma "," delimited list that is usually supplied by using "\${variable}".

```

for variable_name in ( list )
    commands
endfor

```

The value of "variable_name" can be referenced in the loop like any other variable with "\${variable_name}".

The "do while" loop continues execution until the provided logical expression is no longer true.

```
do while ( logical_expression )
    commands
enddo
```

where "logical_expression" is a logical expression like those described above.

The "do until" loop continues execution until the provided logical expression is true (*i.e.*, while the expression is false).

```
do until ( logical_expression )
    commands
enddo
```

where "logical_expression" is a logical expression like those described above.

The "do" loop executes a specified number of times depending on the values of "start", "stop", and "increment".

```
do variable_name = start, stop, increment
    commands
enddo
```

The value of "variable_name" can be referenced in the loop like any other variable with "\${variable_name}". "start", "stop", and "increment" are integers that determine the value of "variable_name" and how many times the loop executes. "increment" is optional with a default value of 1.

The nesting level for loop constructs (in any combination) is 10 for each active macro.

K.2.8 Input/output Statements

Currently, macro mode supports the following input/output commands: "open", "openw", "read", "write", "backspace", "rewind", "close".

```
open      ( file_name, options )
openw     ( file_name, options )
read      ( file_name, options ) variable_name
write     ( file_name, options ) output_data
backspace ( file_name, options )
rewind    ( file_name, options )
close     ( file_name, options )
```

The "open" statement opens an external data file for reading, while the "openw" statement opens an external data file for writing. The "read" statement reads a record from the specified input file into the variable "variable_name" (*i.e.*, the entire record is placed in "variable_name" -- remem-

ber that individual elements of a variable can be referenced using the "\${variable_name}[element_number]" notation). A file must be opened with "open" before it can be read and an open file can be accessed globally by all macros. The "write" statement allows formatted data to be written to an output file. A file must be opened with "openw" before it can be written.

Note: In order for an open file to be accessed by submacros, the file must have been defined as a global variable (all uppercase) or the filename must be passed as an argument when calling the submacro.

The "options" for each i/o statement are listed below (separate options with commas ",", if more than one is used).

```
err=      error_branch
end=      end_branch
iostat=   variable_name
fmt=      [format_statement]
```

The "err=" option specifies a branch location (a statement label) for input/output statements that generate errors. The "end=" option specifies a branch location for end-of-files. The "iostat=" option specifies a variable whose value is set to either "0" for a successful operation, "-1" for an end-of-file, or "+1" for an error. If "iostat=" is used, then both the "err=" and "end=" options are ignored. If an error or end-of-file occurs and an "err=" or "end=" (or "iostat=") option has not been specified, then a fatal error will be generated. The "fmt" option is used for writing formatted data. Only the following simple FORTRAN-like formats are supported: A, I, F, and E. (For example, "fmt=[3A10,2F12.3]" is an acceptable format. However, "fmt=[3(I5,F3.1)]" is not.) Note that there can be no spaces between the option name and the equal "=" symbol. No more than five files (all macros combined) can be open at one time.

K.2.9 Other Macro Functions

This section describes some convenient functions that can be used in macro mode.

<code>var = angle(i, j, k)</code>	Returns angle in degrees.
<code>var = atom_number(atom_reference)</code>	Returns absolute atom number.
<code>var = atom_reference(atom_number)</code>	Returns atom reference.
<code>var = check_rotatable_bond(i, j)</code>	Returns "true" for rotatable bond.
<code>var = check_bond(i, j)</code>	Returns "true" for bonded atoms.
<code>var = color_mode(group_name)</code>	Returns "monochrome" or "half bonded".
<code>var = define_menu(menu_name, entrs)</code>	Returns completion status.
<code>var = dihedral(i, j, k, l)</code>	Returns angle in degrees.
<code>var = display_info(string1, ...)</code>	Returns completion status.

<code>var = display_logo()</code>	Returns completion status.
<code>var = display_menu(menu_name)</code>	Returns completion status.
<code>var = distance(i, j)</code>	Returns distance in Å.
<code>var = file_start(file_name)</code>	Returns absolute atom number.
<code>var = file_stop(file_name)</code>	Returns absolute atom number.
<code>var = fix_angle(i, j, k, val, force)</code>	Returns completion status.
<code>var = fix_dihedral(i, j, k, l, val, force)</code>	Returns completion status.
<code>var = fix_distance(i, j, val, force)</code>	Returns completion status.
<code>var = redraw(group_name)</code>	Returns completion status.
<code>var = regroup(group_name)</code>	Returns completion status.
<code>var = remove_menu(menu_name)</code>	Returns completion status.
<code>var = set_angle(i, j, k, val)</code>	Returns completion status.
<code>var = set_dihedral(i, j, k, l, val)</code>	Returns completion status.
<code>var = set_distance(i, j, val)</code>	Returns completion status.
<code>var = update_ee()</code>	Returns completion status.

If an error is encountered during the execution of one of the above functions, then the returned value is "NULL".

K.2.10 User Routines

The user routines "user1" through "user16" can be invoked by using the "call userN" command where "N" is an integer from 1 to 16.

```
call userN
```

Of course, the required user routines must be linked into the current executable.

K.2.11 Debugger

A rudimentary debugger is provided for use in developing macros. To invoke the debugger, issue the following command before executing the macro (or place it within the macro itself).

```
debug=true
```

To turn the debugger off, use the following command.

```
debug=false
```

Instructions for debugger use can be obtained by typing "help" once the debugger is active.

K.3 Command Summary

Command and Macro Mode:

```
!!  
!string  
@  
^  
autopause=state*  
comment "string"  
cycle  
debug=state*  
echo "string"  
entry-level  
EXIT  
history  
logmacro=state*  
menu, quit, stop  
off  
on  
pause  
pick=state*  
refresh  
restart  
restore  
save  
trace=state*  
xrot, yrot, zrot  
xtran, ytran, ztran  
zoom
```

*state can be true or false

Macro Mode Only:

```
beginmacro, endmacro  
call  
do, do while, do until, enddo, end do  
for, endfor, end for  
goto, go to  
if, then, elseif, else, endif, end if  
open, read, backspace, rewind, close
```


K.4 Sample Macros

Several sample macros have been included in the distribution directory *bg_macro* in order to illustrate the use of the most common macro features.

Appendix L

DREIDING: A Generic Force Field for Molecular Simulations¹

Stephen L. Mayo, Barry D. Olafson, and William A. Goddard III

*BioDesign, Inc.,
199 South Los Robles (Suite 540)
Pasadena, California 91101*

J. Phys. Chem. **94** (1990) 8897-8909

ABSTRACT

We report the parameters for a new generic force field, DREIDING, that we find useful for predicting structures and dynamics of organic, biological, and main group inorganic molecules. The philosophy in DREIDING is to use general force constants and geometry parameters based on simple hybridization considerations rather than individual force constants and geometric parameters that depend on the particular combination of atoms involved in the bond, angle, or torsion terms. Thus all bond distances are derived from atomic radii, and there is only one force constant each for bonds, angles, and inversions and only six different values for torsional barriers. Parameters are defined for all possible combinations of atoms and new atoms can be added to the force field rather simply.

This paper reports the parameters for the "non-metallic" main group elements (B, C, N, O, F columns for the C, Si, Ge, and Sn rows) plus H and a few metals (Na, Ca, Zn, Fe).

The accuracy of the DREIDING force field is tested by comparing with (i) 76 accurately determined crystal structures of organic compounds involving H, C, N, O, F, P, S, Cl, and Br, (ii) ro-

¹ S. L. Mayo, B. D. Olafson and W. A. Goddard III, *J. Phys. Chem.*, **94** (1990) 8897-8909. Only the abstract of this paper is presented here (please refer to *J. Phys. Chem.* for the complete paper).

tational barriers of a number of molecules, and (iii) relative conformational energies and barriers of a number of molecules. We find excellent results for these systems.

Appendix M

Standard Space Group Names

Standard space group names are used to designate one of the 230 space groups. The space group names are based on Hermann-Mauguin short symbols. The allowed space group names are listed below for each of the 230 space groups. The program indicates the corresponding Schoenflies symbol. For more information, see the *International Tables for Crystallography, Volume A*, ed. T. Hahn, (D. Reidel Publishing Co., Boston, 1983). Cases where a space group can have two possible origins, axis geometries, or choices of unique axis are shown.

Note: In addition to the standard number, the internal number which is used by the program is also listed. The cases are also included with these, so there are a total of 274 internal numbers. When inputting data with an input file, the internal number should be included for space groups that have more than one case.

#	Int #	Space Grp	Case	#	Int #	Space Grp	Case
		<u>Triclinic</u>					
1	1	P1		29	42	Pca21	
2	2	P-1		30	43	Pnc2	
		<u>Monoclinic</u>		31	44	Pmn21	
3	3	P2	1 unique c	32	45	Pba2	
	4		2 unique b	33	46	Pna21	
4	5	P21	1 unique c	34	47	Pnn2	
	6		2 unique b	35	48	Cmm2	
5	7	B2		36	49	Cmc21	
	8	C2		37	50	Ccc2	
6	9	Pm	1 unique c	38	51	Amm2	
	10		2 unique b	39	52	Abm2	
7	11	Pb		40	53	Ama2	
	12	Pc		41	54	Aba2	
8	13	Bm		42	55	Fmm2	
	14	Cm		43	56	Fdd2	
9	15	Bb		44	57	Imm2	
	16	Cc		45	58	Iba2	
10	17	P2/m	1 unique c	46	59	ma2	
	18		2 unique b	47	60	Pmmm	
11	19	P21/m	1 unique c	48	61	Pnnn	1 Origin 222
	20		2 unique b		62		2 Origin -1
12	21	B2/m		49	63	Pccm	
	22	C2/m		50	64	Pban	1 Origin 222
13	23	P2/b			65		2 Origin -1
	24	P2/c		51	66	Pmma	
14	25	P21/b		52	67	Pnna	
	26	P21/c		53	68	Pmna	
15	27	B2/b		54	69	Pcca	
	28	C2/c		55	70	Pbam	
		<u>Orthorhombic</u>		56	71	Pccn	
16	29	P222		57	72	Pbcm	
17	30	P2221		58	73	Pnnm	
18	31	P21212		59	74	Pmmn	1 Origin mmm
19	32	P212121			75		2 Origin -1
20	33	C2221		60	76	Pbcn	
21	34	C222		61	77	Pbca	
22	35	F222		62	78	Pnma	
23	36	I222		63	79	Cmcm	
24	37	I212121		64	80	Cmca	
25	38	Pmm2		65	81	Cmmm	
26	39	Pmc21		66	82	Cccm	
27	40	Pcc2		67	83	Cmma	
28	41	Pma2		68	84	Ccca1	1 Origin 222
					85		2 Origin -1 on nca

#	Int #	Space Grp	Case	#	Int #	Space Grp	Case
69	86	Fmmm		106	127	P42bc	
70	87	Fddd	1 Origin 222	107	128	I4mm	
	88		2 Origin -1	108	129	I4cm	
71	89	Immm		109	130	I41md	
72	90	Ibam		110	131	I41cd	
73	91	Ibca		111	132	P-42m	
74	92	Imma		112	133	P-42c	
		<u>Tetragonal</u>		113	134	P-421m	
75	93	P4		114	135	P-421c	
76	94	P41		115	136	P-4m2	
77	95	P42		116	137	P-4c2	
78	96	P43		117	138	P-4b2	
79	97	I4		118	139	P-4n2	
80	98	I41		119	140	I-4m2	
81	99	P-4		120	141	I-4c2	
82	100	I-4		121	142	I-42m	
83	101	P4/m		122	143	I-42d	
84	102	P42/m		123	144	P4/mmm	
85	103	P4/n	1 Origin -4	124	145	P4/mcc	
	104		2 Origin -1	125	146	P4/nbm	1 Origin -4m2
86	105	P42/n	1 Origin -4		147		2 Origin 2/m
	106		2 Origin -1	126	148	P4/nnc	1 Origin 422
87	107	I4/m			149		2 Origin -1
88	108	I41/a	1 Origin -4	127	150	P4/mbm	
	109		2 Origin -1	128	151	P4/mnc	
89	110	P422		129	152	P4/nmm	1 Origin -4m2
90	111	P4212			153		2 Origin 2/m
91	112	P4122		130	154	P4/ncc	1 Origin -4
92	113	P41212			155		2 Origin -1
93	114	P4222		131	156	P42/mmc	
94	115	P42212		132	157	P42/mcm	
95	116	P4322		133	158	P42/nbc	1 Origin -4
96	117	P43212			159		2 Origin -1
97	118	I422		134	160	P42/nnm	1 Origin -42m
98	119	I4122			161		2 Origin 2/m
99	120	P4mm		135	162	P42/mbc	
100	121	P4bm		136	163	P42/mnm	
101	122	P42cm		137	164	P42/nmc	1 Origin -4m2
102	123	P42nm			165		2 Origin -1
103	124	P4cc		138	166	P42/ncm	1 Origin -4
104	125	P4cn			167		2 Origin 2/m
105	126	P42mc		139	168	I4/mmm	

#	Int #	Space Grp	Case	#	Int #	Space Grp	Case
140	169	I4/mcm		173	211	P63	
141	170	I41/amd	1 Origin -4m2	174	212	P-6	
	171		2 Origin 2/m	175	213	P6/m	
142	172	I41/acd	1 Origin -4	176	214	P63/m	
	173		2 Origin -1	177	215	P622	
		<u>Trigonal</u>		178	216	P6122	
143	174	P3		179	217	P6522	
144	175	P31		180	218	P6222	
145	176	P32		181	219	P6422	
146	177	R3	1 rhombohedral	182	220	P6322	
	178		2 hexagonal	183	221	P6mm	
147	179	P-3		184	222	P6cc	
148	180	R-3	1 rhombohedral	185	223	P63cm	
	181		2 hexagonal	186	224	P63mc	
149	182	P312		187	225	P-6m2	
150	183	P321		188	226	P-6c2	
151	184	P3112		189	227	P-62m	
152	185	P3121		190	228	P-62c	
153	186	P3212		191	229	P6/mmm	
154	187	P3221		192	230	P6/mcc	
155	188	R32	1 rhombohedral	193	231	P63/mcm	
	189		2 hexagonal	194	232	P63/mmc	
156	190	P3m1				<u>Cubic</u>	
157	191	P31m		195	233	P23	
158	192	P3c1		196	234	F23	
159	193	P31c		197	235	I23	
160	194	R3m	1 rhombohedral	198	236	P213	
	195		2 hexagonal	199	237	I213	
161	196	R3c	1 rhombohedral	200	238	Pm3	
	197		2 hexagonal	201	239	Pn3	1 Origin 23
162	198	P-31m			240		2 Origin -3
163	199	P-31c		202	241	Fm3	
164	200	P-3m1		203	242	Fd3	1 Origin 23
165	201	P-3c1			243		2 Origin -3
166	202	R-3m	1 rhombohedral	204	244	Im3	
	203		2 hexagonal	205	245	Pa3	
167	204	R-3c	1 rhombohedral	206	246	Ia3	
	205		2 hexagonal	207	247	P432	
		<u>Hexagonal</u>		208	248	P4232	
168	206	P6		209	249	F432	
169	207	P61		210	250	F4132	
170	208	P65		211	251	I432	
171	209	P62		212	252	P4332	
172	210	P64		213	253	P4132	

#	Int #	Space Grp	Case
214	254	I4132	
215	255	P-43m	
216	256	F-43m	
217	257	I-43m	
218	258	P-43c	
219	259	F-43c	
220	260	I-43d	
221	261	Pm3m	
222	262	Pn3n	1 Origin 43
	263		2 Origin -3
223	264	Pm3n	
224	265	Pn3m	1 Origin -43
	266		2 Origin -3m
225	267	Fm3m	
226	268	Fm3c	
227	269	Fd3m	1 Origin -43m
	270		2 Origin -3m
228	271	Fd3c	1 Origin 23
	272		2 Origin -3
229	273	Im3m	
230	274	Ia3d	

Appendix N

Auxiliary Programs

Your Molecular Simulations product distribution includes three auxiliary programs. These are the *rdtrj*, *wrtrj*, and *shtrj* programs. These can be used to examine the information in trajectory files and to transfer trajectory files from one kind of computer to another. These programs are shipped in the same directory as the Molecular Simulations executable programs (see the *Molecular Simulations Installation Guide* for further information). The FORTRAN source code for these programs is included in the product distribution in the *biogv300/source* directory.

Unlike most files used by the program, Molecular Simulations trajectory files are binary format files. It is impossible to examine the information that they contain without using a computer program. We have alleviated this problem by developing the *shtrj* program which provides a quick summary of a trajectory file's contents. *Shtrj* reads a binary format trajectory file and prints selected information about the trajectory in the text window. This program is described in Section N.1 below. Alternatively, the *rdtrj* program can be used to create ASCII format copies of trajectory files. ASCII format is readable without the aid of a computer program, but requires much more disk space than binary format. The *rdtrj* program is described in Section N.2.

A site that uses more than one kind of computer to generate trajectory files may want to transfer trajectory files from one computer to another. Sun, IRIS, RS6000, E&S ESV, Convex, and Alliant computers generate trajectory files that have exactly the same format, and trajectory files generated on any of these computers can be read and used by any other computer of this group. However, trajectory files produced on DEC 5000 and Stardent Titan computers can only be read on the same kind of host, and trajectory files produced on other machines cannot be read on DEC 5000 and Titan computers. The *rdtrj* and *wrtrj* programs were developed to make it possible to copy trajectory files between machines that have incompatible trajectory file formats. An ASCII format copy of the trajectory file is generated using *rdtrj*. A file transfer program is used to transfer the ASCII copy to the target machine. *Wrtrj* is then used on the new machine to convert the ASCII copy to the binary format usable by the program. The ANIMATE functions (see Section 5.14) and the trajectory analysis functions (see Section 6.7) can then be used on the transferred file. The procedure required to transfer trajectory files between computers is outlined in Section N.4.

N.1 SHTRJ

The *shtrj* program allows the user to read a binary format trajectory file and print summary information about the trajectory in the text window. The information that is printed depends on the kind of trajectory file being examined. Generally important energy and temperature information is listed along with the step and time values (see the example shown in Table N.1). Potential energy values can also be listed using the LIST function accessed under ANIMATE (see Section 5.14.3). However, *shtrj* lists more information and is quicker and simpler to use, since *shtrj* does not require the user to start up the program, read in the corresponding coordinate and trajectory files, and use the ANIMATE functions.

User Input:

1. Start the program by typing the command *shtrj* at the system command prompt:
\$ *shtrj*
2. Enter the input trajectory file name, including the full disk/directory path specification if the file is not in your current directory (e.g., */xxxgraf/biogv300/data/filename.trj*).
3. Enter the read increment (1 implies read every point). The default action is to print information for every snapshot in the trajectory file. If you enter a number *n* (*n* must be larger than one), the program will print information every *n*th snapshot.

The program will then read the trajectory file and print summary information about it. The number of snapshots read is included at the end of the listing. For example, when *shtrj* is used to examine the file *filename.trj* using a read increment of 10, it produces output of the form:

25 ps straight dynamics run

STEP	TIME	POTENTLE	INTERNAL	NONBOND	CONSTRAINT	KE(KELVIN)
10	1.0000	109.3001	68.6559	45.7518	0.0000	301.2609
20	2.0000	104.1281	60.9866	44.3569	0.0000	292.6514
30	3.0000	102.1923	75.2188	27.4436	0.0000	345.0107
			:			
			:			

250 snapshots read

Table N.1: Sample *shtrj* summary listing for trajectory file *filename.trj*.

N.2 RDTRJ

The *rdtrj* program is used to make an ASCII format copy of a normal binary format trajectory file produced by a dynamics, constrained minimization, Monte Carlo search, or conformational search calculation. The ASCII copy uses about two and a half times as much disk space as the binary trajectory file, so it is important to verify that there is enough disk space available for the ASCII file. The format of the ASCII copy is the same as that described in Appendix D.5. The program

prints the message *Working....* while it reads the trajectory file and writes the ASCII copy. When the conversion is complete, the message *Binary to ascii conversion completed...* is printed.

User Input:

1. Start the program by typing the command ***rdtrj*** at the system command prompt:
\$ ***rdtrj***
2. Enter the input (binary) trajectory file name, including the full disk/directory path specification if the file is not in your current directory (e.g., */xxxgraf/biogv300/data/filename.trj*).
3. Enter the output (ASCII) trajectory file name you wish to use for the ASCII copy of the trajectory file, including the full disk/directory path specification if you want the copy to be somewhere other than your current directory (e.g., */user/mla/filename.atrj*).

N.3 WRTRJ

The *wrtrj* program is used to convert an ASCII format copy of a trajectory file to a normal binary format trajectory file usable by the program. The message *Working....* appears while the program reads the ASCII copy of the trajectory file and writes the binary file. When the conversion is complete, the message *Ascii to binary conversion completed...* is printed.

User Input:

1. Start the program by typing the command ***wrtrj*** at the system command prompt:
\$ ***wrtrj***
2. Enter the input (ASCII) trajectory file name, including the full disk/directory path specification if the file is not in your current directory (e.g., */user/mla/filename.atrj*).
3. Enter the output (binary) trajectory file name you wish to use for the binary trajectory file, including the full disk/directory path specification if you want it to be somewhere other than your current directory (e.g., */user/mla/filename.trj*).

N.4 Transferring Trajectory Files Between Computers with Incompatible Trajectory File Formats

Sun, IRIS, RS6000, E&S ESV, Convex, and Alliant computers generate trajectory files that have exactly the same format, and trajectory files generated on any of these computers can be read and used by any other computer of this group. However, trajectory files produced on the DEC 5000 and Stardent Titan computers can only be read on the same kind of host, and trajectory files produced on other machines cannot be read on DEC 5000 and Titan computers. The *rdtrj* and *wrtrj* programs make it possible to copy trajectory files between machines that have incompatible trajectory file formats.

The file transfer procedure is as follows. An ASCII format copy of the trajectory file is generated using *rdtrj*. A file transfer program is used to transfer the ASCII copy to the target machine. *Wrtrj* is then used on the new machine to convert the ASCII copy to the binary format usable by the program. The ANIMATE functions (see Section 5.14) and the trajectory analysis functions (see Section 6.7) can then be used on the transferred file.

User Input:

1. After the trajectory file has been created, exit the program and run *rdtrj* on the same computer to generate an ASCII format copy of the trajectory file.
2. Use a file transfer program such as *ftp* or *rcp* to copy the ASCII format trajectory file to the target machine. (This may not be necessary if the target machine is a Titan computer that can access the disk drives of the original computer using NFS). See your systems administrator for further information.
3. Run *wrtrj* on the target computer. Enter the name of the ASCII format trajectory file when *wrtrj* prompts for the name of the input file. The output binary file will be usable by the ANIMATE and ANALYZE/TRAJECTORY functions on the target machine.

Appendix O

Reading and Building Structures with the AMBER Force Field in BIOGRAF

The steps required to read or build RNA, DNA and peptide structures and run energy calculations using the Molecular Simulations implementation of the AMBER force field are described below. Molecular Simulations' implementation is based on the description of the AMBER force field reported by Weiner and coworkers^{1, 2}. These procedures produce structures with atom types that allow usable energy expressions to be set up.

Note: When working with peptides, do not change the atom type of the N-terminal nitrogen from that produced by CONVERT (atom type = N). This atom type is required to do energy calculations when using the AMBER force field.

O.1 Building RNA or DNA

1. Enter the RNA-DNA Builder, SETUP the file and use the SINGLE/DOUBLE/TRIPLE options to create the chain.
2. Return to the RNA-DNA Builder menu and use the BAP toggle to remove the 3' phosphate group and add H bonding hydrogens to the bases. This step is necessary to convert H3 type hydrogens on adenine, guanine, and cytosine amino nitrogens to H2.

Note: If using the DNA-OLD Builder to build double stranded DNA, no special manipulation is necessary.

O.2 Reading DNA BioDesign General Format or Brookhaven Files

1. Read in the file.
2. Run CONVERT on the file.
3. If you have added hetero hydrogens, it is necessary to run CONVERT again.

¹S. J. Weiner, *et al.*, *J. Am. Chem. Soc.* **106** (1984) 765-784.

²S. J. Weiner, *et al.*, *J. Comp. Chem.* **7** (1986) 230-252.

O.3 Building Peptides

No special manipulation is necessary.

O.4 Reading Peptide BioDesign General Format Files (created using the Dreiding force fields)

O.4.1 Files Without Hydrogen Atoms

1. Read the file.
2. Return to the top-level menu, pick BUILD, pick CONVERT, then pick CONVERT on the CONVERT submenu.
3. Enter the Peptide Builder.
4. At the prompt to pick the file you wish to use, pick the file you just read in.
5. Go to the EDIT menu, use SELECT to specify heteroatoms, then choose ADD H.
6. Perform the AUTO TYPE function (pick AUTO TYPE, then RETURN).
7. Modify the atom type of each peptide chain's N terminal nitrogen from NT to N. This can be done using the ATOM TYPE function on the EDIT ID menu. Modify the hydrogens on the arginine side chain nitrogens to atom type H3. It may be easier to find the N terminus (or N termini) and the arginine side chains if you use the MAKE GROUPS menu to create a group consisting only of the first residue and the arginine side chains (see RESIDUE and LIMITS in Sections 4.1.5 and 4.1.6).

O.4.2 Files With Hetero Hydrogens (but no hydrogens on carbons)

1. Read the file.
2. Enter the Peptide Builder.
3. At the prompt to pick the file you wish to use, pick the file you just read in.
4. Perform the AUTO TYPE function on the EDIT ID menu (pick AUTO TYPE, then pick RETURN).
5. Modify the atom type of each peptide chain's N terminal nitrogen from NT to N. This can be done using the ATOM TYPE function on the EDIT ID menu. Modify the hydrogens on the arginine side chain nitrogens to atom type H3. It may be easier to find the N terminus (or N termini) and the arginine side chains if you use the MAKE GROUPS menu to create a group consisting only of the first residue and the arginine side chains (see RESIDUE and LIMITS in Sections 4.1.5 and 4.1.6).

O.4.3 Files With All Their Hydrogens (no implicit hydrogens)

1. Read the file.
2. Enter the Peptide Builder.
3. At the prompt to pick the file you wish to use, pick the file you just read in.
4. Go to the EDIT menu, use SELECT to specify hydrogens, then choose DELETE to delete all hydrogens (the BIOGRAF implementation of AMBER allows only hetero hydrogens).

5. Return to the main BUILD menu, pick CONVERT, then pick CONVERT on the CONVERT submenu.
6. Return to the Peptide Builder.
7. Go to the EDIT menu, use SELECT to specify heteroatoms, then choose ADD H to add back the hetero hydrogens.
8. Perform the AUTO TYPE function (pick AUTO TYPE, then RETURN).
9. Modify the atom type of each peptide chain's N terminal nitrogen from NT to N. This can be done using the ATOM TYPE function on the EDIT ID menu. Modify the hydrogens on the arginine side chain nitrogens to atom type H3. It may be easier to find the N terminus (or N termini) and the arginine side chains if you use the MAKE GROUPS menu to create a group consisting only of the first residue and the arginine side chains (see RESIDUE and LIMITS in Sections 4.1.5 and 4.1.6).

O.5 Reading Peptide Brookhaven Files

1. Read the file.
2. Enter the Peptide Builder.
3. At the prompt to pick the file you wish to use, pick the file you just read in.
4. Go to the EDIT menu, use SELECT to specify heteroatoms, then choose ADD H to add the hetero hydrogens.
5. Perform the AUTO TYPE function (pick AUTO TYPE, then RETURN).
6. Modify the atom type of each peptide chain's N terminal nitrogen from NT to N. This can be done using the ATOM TYPE function on the EDIT ID menu. Modify the hydrogens on the arginine side chain nitrogens to atom type H3. It may be easier to find the N terminus (or N termini) and the arginine side chains if you use the MAKE GROUPS menu to create a group consisting only of the first residue and the arginine side chains (see RESIDUE and LIMITS in Sections 4.1.5 and 4.1.6).

Warning: If there are metallic ligands in your coordinate file, there is a good chance that lone pairs of infinite length will be generated and you will have an unusable structure. This is because AMBER has no atom types for metal ions.

Appendix P

Equilibrium Molecular Dynamics

One of the objectives of molecular dynamics is to obtain the equilibrium thermodynamic properties of a system. If a microscopic dynamic variable A takes on values $A(t)$ along the trajectory, then the following time average

$$A = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(t) dt \quad (\text{P.1})$$

yields the measured thermodynamic value for the selected variable. This dynamic variable can be any function of the coordinates and momenta of the particles of the system. Through time averaging, the first order properties of a system (*e.g.*, internal energy, kinetic energy, pressure, virial) can be calculated. Similarly, using microscopic expressions in the form of fluctuations of these first order properties we can also calculate the thermodynamic response functions of a system (*e.g.*, specific heat, thermal expansion, bulk modulus).

In the thermodynamic limit, the first order properties obtained in one ensemble are equivalent to those obtained in other ensembles (differences are of the order of $1/N$). But second order properties such as specific heats, compressibilities, and elastic constants differ from one ensemble to another. For instance, specific heat at constant pressure differs from specific heat at constant volume. Therefore, it is important to use the appropriate ensemble when performing simulations to obtain these properties. Several forms of molecular dynamics have been proposed to simulate these different ensembles. These include constant pressure-constant enthalpy¹, constant stress-constant enthalpy², constant temperature³, and constant chemical potential⁴ dynamics. These new forms of dynamics are all based on the Hamiltonian formalism and rigorously satisfy the required conditions. These forms of molecular dynamics are briefly described below.

¹ H. C. Andersen, *J. Chem. Phys.* **72** (1980) 2384

² M. Parrinello and A. Rahman, *J. Appl. Phys.* **52** (1981) 7182.

³ S. Nosé, *J. Chem. Phys.* **81** (1984) 511.

⁴ T. Cagin and B. M. Pettitt, *Mol. Phys* **72** (1991) 169.

P.1 Constant Energy Constant Volume Dynamics

Constant energy constant volume dynamics is the traditional form of molecular dynamics. In this form, the Newtonian equations of motion of N particles in a periodic volume V are integrated numerically to obtain the trajectories. The system is described by a classical Hamiltonian:

$$H(x,p) = K + U = E \quad (P.2)$$

where K is the kinetic energy, U is the potential energy through which the particles in the system interact, and E is the total energy of the system (which remains constant). This form of molecular dynamics generates trajectories representative of the classical microcanonical ensemble. The fluctuation expressions can be used to calculate the thermodynamic response functions⁵.

P.2 Constant Enthalpy Constant Pressure Dynamics

In 1980 Andersen¹ introduced volume and the conjugate momentum of volume as additional dynamic variables to perform constant pressure constant enthalpy dynamics. This form of dynamics is particularly useful for isotropic systems. Parrinello and Rahman² extended this method to handle anisotropic solids, where not only the size but also the shape of the system is treated as a dynamic variable. This form of dynamics generates trajectories representative of the constant enthalpy constant pressure/stress ensemble of statistical mechanics (HPN or H σ N). The equations of motion for this form of molecular dynamics are obtained from the following extended system Hamiltonian:

$$H(\rho, \pi, h, \Pi) = \sum \frac{(h^{-1} \pi_a)^t h^{-1} \pi_a}{2m_a} + U(\rho_a, h; a=1, \dots, N) + \frac{\text{Tr}(\Pi^t \Pi)}{2W} - V_0 \text{Tr}(\sigma^t \epsilon) \quad (P.3)$$

Here, h is a 3x3 matrix which is formed by the three vectors defining the simulation cell, Π is the corresponding momenta for the cell variables, t is the matrix transposition, W is the mass parameter, σ is the applied stress, and ϵ is the strain. In this formulation, the coordinates and momenta of the particles appear as the scaled variables ρ and π . They are related to the Cartesian coordinates and momenta as follows:

$$r = h\rho \quad p = (h^{-1})^t \pi \quad (P.4)$$

The natural response functions (specific heat at constant pressure, thermal expansion, adiabatic compressibility, and adiabatic compliance tensor) are obtained from the proper statistical fluctuation expressions of kinetic energy, volume, and strain.⁵

The selection of the mass parameter W is crucial in tuning the fluctuations of the volume/shape matrix⁶.

⁵ J. R. Ray, *Comp. Phys. Rep.* **8** (1988) 109 and references therein.

⁶ S. Nosé and M. L. Klein, *Mol. Phys.*, **50** (1983) 1055.

P.3 Constant Temperature Constant Volume Dynamics

Nosé developed a Hamiltonian formalism to simulate a system at constant temperature:

$$H_N(\rho_a, \pi_a, s, P_s) = \sum \frac{\pi_a^2}{2m_a s^2} + U(\{\rho_a\}) + \frac{P_s^2}{2Q} + (3N+1) \ln s \quad (P.5)$$

Here the additional variables are s and P_s , and Q is a mass-like parameter. Nosé showed that the trajectories generated by this Hamiltonian represent the canonical ensemble distribution. The variables $(\{\rho_a, \pi_a\}, a = 1, \dots, N)$ are the coordinates and momenta of the particles in the extended space representation. Their relation to the physical coordinates and momenta are

$$p_a = \frac{\pi_a}{s} \quad r_a = \rho_a \quad (P.6)$$

A detailed discussion of canonical dynamics, the equations of motion, and the different forms of canonical dynamics available in the program are described below (see Section P.6).

P.4 Constant Temperature Constant Pressure Dynamics

The extension of constant temperature constant volume dynamics to constant temperature constant pressure dynamics is straightforward: the Anderson-Parinello-Rahman Hamiltonian is combined with the Nosé formalism. The resulting Hamiltonian is given by:

$$H(\rho_a, \pi_a, s, P_s, h, \Pi) = \sum \frac{(h \pi_a)^t h \pi_a}{2m_a s^2} + U(h, \{\rho_a\}) + \frac{P_s^2}{2Q} \\ + (3N+1) \ln s + \frac{\text{Tr}(\Pi^t \Pi)}{2W} - V_o \text{Tr}(\sigma^t \epsilon) \quad (P.7)$$

P.5 Non-Hamiltonian Forms of Constant Temperature Dynamics

In addition to the extended system Hamiltonian approach to constant temperature molecular dynamics, coupling to a thermal bath can be achieved through other methods. These include the gaussian constraint method⁷ and coupling schemes based on Brownian-Langevin dynamics⁸. The gaussian constraint method generates a canonical distribution in configuration space but in momentum space it generates an isokinetic distribution. The Berendsen *et al* approach⁹ uses a weak coupling scheme based on scaling the velocities of atoms at every step:

⁷ D. J. Evans and G. P. Morriss, *Comput. Phys. Rep.* **1** (1984) 297.

⁸ G. Ciccotti and A. Tenenbaum, *J. Stat. Phys.* **23** (1980) 767.

⁹ H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.* **81** (1984) 3684.

$$\chi = \left[1 + \frac{\delta t}{\tau} \left(\frac{T_b}{T_i} - 1 \right) \right]^{1/2} \quad (\text{P.8})$$

where T_b is the bath temperature, T_i is the instantaneous temperature, τ is a preset damping constant and δt is the integration time step used in the simulation. The scaling introduced above creates a tendency to move towards the bath temperature while slightly changing the forces on each atom. Equation P.8 can be rewritten by using the first two terms of the binomial expansion of the square root

$$\chi = 1 + \frac{\delta t}{2\tau} \left(\frac{T_b}{T_i} - 1 \right) \quad (\text{P.9})$$

The constant τ is 0.1 – 0.4 ps.

P.6 Nosé-Hoover Canonical Dynamics

The conventional form of molecular dynamics solves the Newtonian equations of motion numerically

$$F_{a\alpha} = M_{a\alpha} A_{a\alpha} \quad a = 1, 2, \dots, N, \quad \alpha = x, y, z \quad (\text{P.10})$$

to obtain the trajectory of the system. If there are no external forces applied to the simulated system, the total energy, linear momentum and angular momentum of the system are conserved. The statistical distribution of states of such a physical system is represented by the microcanonical distribution function

$$\rho(E, V, N) = \frac{\delta(H - E)}{\omega(E, V, N)} \quad (\text{P.11})$$

where H is the Hamiltonian (total energy expression) and E is its constant value. In the above expression, $\omega(E, V, N)$ is the microcanonical partition function and is defined as

$$\omega(E, V, N) = C_N \int \delta(H - E) \prod dp_a dx_a \quad (\text{P.12})$$

Here, C_N is merely a normalization constant and $d\tau$ is the $6N$ dimensional elemental volume. The expression above suggests that the trajectories of the system are constrained to a surface in a multi-dimensional phase space. Such a system is completely isolated from the environment. However, most natural phenomena occur under conditions where the system exchanges energy with its environment (in the form of heat) or where work is performed on or by the system due to external pressure and/or stress. Matter may even diffuse in or out of the system through the container.

In the early 1970's the compelling need for isothermal dynamics prompted the practitioners in the field to apply adhoc velocity scaling in molecular dynamics runs in order to keep the temperature of the system constant

$$v' = v \sqrt{\frac{T_0}{T_i}} \quad (\text{P.13})$$

However, the velocity scaling only suppresses the natural fluctuations in the system. For instance, in the canonical ensemble the kinetic energy of the system has the following fluctuation property, with the angular brackets representing the ensemble average of the quantities

$$(\Delta K)^2 = \langle K^2 \rangle - \langle K \rangle^2 = \frac{3N(kT)^2}{2} \quad (\text{P.14})$$

Therefore, such strict scaling schemes in no way generate a canonical distribution of states. Later on several other temperature constraining schemes were proposed. The Nosé method is based on extending the degrees of freedom by adding a dynamic variable to mock the interaction of the system with the heat bath. The final extended system is represented by a Hamiltonian structure and guarantees a canonical distribution of states in the phase space. The Hamiltonian of the system in the extended phase space is given by

$$H_N(\rho_a, \pi_a, s, P_s) = \sum \frac{\pi_a^2}{2m_a s^2} + U(\{\rho_a\}) + \frac{P_s^2}{2Q} + (3N+1) \ln s \quad (\text{P.15})$$

In the above expression s is the additional heat bath variable and P_s is the corresponding momentum, the variables $(\{\rho_a, \pi_a\}, a = 1, 2, \dots, N)$ are the coordinates and momenta of the particles in the extended space representation. Their relation to physical coordinates and momenta are

$$p_a = \frac{\pi_a}{s} \quad r_a = \rho_a \quad (\text{P.16})$$

The relation between the velocity in terms of variables in the extended system and the physical velocity is given by

$$v_a = s \dot{\rho}_a \quad \text{or} \quad \frac{dx_a}{dt} = s \frac{d\rho_a}{d\tau} \quad (\text{P.17})$$

The s variable appears as the scaling factor for the velocities. Upon inspection of the second relation the physical role of the s variable is actually scaling time. The equations of motion satisfied by the extended system variables are

$$m_a s^2 \frac{d^2 \rho_a}{d\tau^2} = - \frac{\partial U}{\partial \rho_a} - 2 \left(\frac{ds}{d\tau} \right) \left(\frac{d\rho_a}{d\tau} \right) \quad (\text{P.18})$$

$$Q \frac{d^2 s}{d\tau^2} = \frac{(2K - gkT_0)}{Qs} \quad (\text{P.19})$$

where K is the kinetic energy of the system, g is the number of degrees of freedom, and T_0 is the bath temperature. The time evolution of the scaling factor, s , is determined dynamically from the second equation.

Since the extended system equations of motion are derived from a Hamiltonian that does not explicitly depend on time, the total extended system energy, E_N is conserved. The equations of motion generate trajectories representing a microcanonical distribution in this extended phase space of $6N+2$ variables (ρ_a, π_a, s, P_s). The probability density of this extended system is given by

$$\rho_N(E_N, V, N) = \frac{\delta(E_N - H_N(\rho_a, \pi_a, s, P_s))}{\omega_N(E_N, V, N)} \quad (\text{P.20})$$

and normalized as

$$\int \rho_N(E_N, V, N) \left(\prod d\rho_a d\pi_a \right) ds dP_s = 1 \quad (\text{P.21})$$

Starting from the extended phase space partition function of

$$\omega_N(E_N, V, N) = \int \delta(E_N - H_N) \left(\prod d\rho_a d\pi_a \right) ds dP_s \quad (\text{P.22})$$

it can be shown that, upon projecting to the physical phase space, this partition function reduces to a canonical partition function³. The procedure is to replace the extended system variables by physical system variables and then to integrate the dependencies on s and P_s . The resulting expression for the partition function is

$$\omega_N(E_N, V, N) = C(E_N, T) \int \exp\left[-\frac{H(\rho_a, x_a)}{kT}\right] \prod d\rho_a dx_a \quad (\text{P.23})$$

which is equivalent to the canonical partition function,

$$Z(T, V, N) = C \int \exp\left[-\frac{H(\rho_a, x_a)}{kT}\right] \prod d\rho_a dx_a \quad (\text{P.24})$$

In the above formulation, the scaling variable s introduces time scaling which is not desirable. The time scaling can be eliminated as in Hoover's reformulation¹⁰ of Nosé dynamics. Using the

¹⁰ W. H. Hoover, *Phys. Rev. A.*, **31** (1985) 1695.

relation between physical and virtual time, $dt = d\tau / s$, one can write the equations of motion in terms of real time derivatives and eliminate the explicit dependence on s .

$$\frac{d^2 r_a}{dt^2} = -\frac{1}{m_a} \left(\frac{\partial U}{\partial r_a} \right) - \zeta \frac{dr_a}{dt} \quad (\text{P.25})$$

Here ζ is given by

$$\zeta = \frac{\dot{s}}{s} = \frac{P_s}{Q} \quad (\text{P.26})$$

and acts as a friction coefficient. The time evolution of the friction coefficient is given by

$$Q \frac{d\zeta}{dt} = \sum \frac{p_a^2}{m_a} - gkT_0 \quad (\text{P.27})$$

These equations of motion are time reversible and generate a canonical distribution in phase space. Also, the trajectories obtained from the above equations are obtained in real time. This is particularly important when calculating the dynamic properties of a system which requires that the trajectory points be evenly spaced. The original formulation does not provide evenly spaced trajectory points.

The parameter Q is an adjustable parameter which has moment of inertia units ML^2 or ET^2 . The selection of this parameter is based on the stability and precision of the numerical solution to the s - or ζ equation of motion and the vibrational frequencies involved in the system studied. The choice for this parameter should therefore be based on the balance between the stability of the solution and the highest frequency motions of the system.

We can rewrite the s -equation of motion in terms of the deviation of s from the average, $s = \langle s \rangle + \delta s$ as follows:

$$\frac{d^2}{d\tau^2} \delta s = -\frac{2gkT_0 \delta s}{Q \langle s \rangle^2} \quad (\text{P.28})$$

This equation represents a harmonic oscillator with frequency, ω

$$\omega^2 = \frac{2gkT_0}{Q \langle s \rangle^2} \quad (\text{P.29})$$

Assuming $\langle s \rangle^2$ is about 1, therefore Q should be chosen as

$$Q = \frac{2gkT_0}{\omega^2} = \frac{gkT_0 \tau^2}{2\pi^2} \quad (\text{P.30})$$

where τ is the fundamental period for the system. For simple fluids, τ can be chosen as the second moment of the velocity autocorrelation function. For covalently bonded systems, however, it is better to choose τ based on the frequencies involved in the system. If we assume τ_{vib} is the smallest time scale in the system, τ values smaller than τ_{vib} will result in inaccurate solutions for the s or ζ variables. An accurate solution to these equations will require very small time steps. On the other hand, larger values for τ will result in stiffness in the equations of motion, *i.e.* they will inhibit proper canonical fluctuations. A rule of thumb therefore is to choose τ in the same order as τ_{vib} .

In the program, both formulations of canonical dynamics are implemented. In these implementations, the s or ζ equations of motion are written in terms of instantaneous temperature T_i and bath temperature T_0 :

$$\frac{d\zeta}{dt} = \frac{T_i - 1}{q}, \text{ and } \frac{d^2s}{d\tau^2} = \frac{T_i - 1}{qs} \quad (\text{P.31})$$

Thus, the dependence on degrees of freedom and the temperature of the system are eliminated. Here, q is given as

$$q = \frac{2}{\omega^2} \text{ or } \frac{\tau_{\text{vib}}^2}{2\pi^2} \quad (\text{P.32})$$

Index

- 1 Point Verlet, J-2
- 1-4 NONBOND, 6-24, 6-25
- 2 Point Verlet, J-2
- Abbreviations and Wild Cards, K-2
- Action Variables, K-9
- ADD H, 3-15
- ADD ALL H, 3-88
- ADD HET H, 3-88
- ADIABATIC Dynamics, 5-26
- ALL, 5-67, 5-70
- ALL CELLS/UNIT CELL ONLY, 3-94
- ALL CIS, 5-11
- ALL D, 5-12
- ALL L, 5-12
- ALL TRANS, 5-11
- ALPHA HELIX, PAR BETA, EXTENDED,
-50 -40, -75 -20, 50 105, 100 0, 3-31
- Ambiguous Commands, K-2
- ANALYZE, 3-27, 3-36, 3-40, 3-45, 3-47,
3-51, 3-58, 3-68
- ANALYZE BY FF, 5-16
- ANALYZE BY GEOM, 5-16
- Analyze Operations, 6-1
- ANG-ANG INVERS, 5-46, 5-64
- ANG-ANG TORSION, 5-46
- ANGLE (1-3), 5-57
- ANGLE, 3-17, 4-29, 5-9, 5-64, 6-8
- ANGLE-STRET 2 K, 5-59
- ANGLES, 5-17, 5-46
- ANIMATE, 5-65, 5-72
- ANNEAL FREQ, 5-52
- ANNEAL INCREMENT, 5-50
- ANNEALED Dynamics, 5-28
- APPEND TRAJECT, 5-21
- Arithmetic Mean, G-20
- ASSIGN CHARGES, 3-25
- Assignment of Variables, K-8
- Atom and Residue Synonyms File, D-38
- ATOM CHARGE, 6-3
- ATOM LABEL, 6-4
- ATOM LABELS, 3-21
- Atom Labels, D-8
- ATOM MASS, 6-4
- ATOM NUMBER, 6-4
- ATOM RADIUS, 6-4
- Atom Referencing Commands, K-3
- ATOM TYPE, 3-21, 4-21, 6-3
- Atom Types, I-19
- Atomic Parameters, I-12
- AUTO DISTANCE, 3-84
- AUTO ENERGY, 3-84
- AUTO GROUP, 2-20, 3-73, 5-71
- AUTO LABELS, 3-22
- AUTO ROTATE, 4-32
- AUTO SCAN, 5-73
- AUTO TYPE, 3-18, 3-21
- Auto Type Parameters, I-13
- Auxiliary Programs, N-1
- AVERAGE COORDS, 5-75
- BALL & STICK, 4-12
- BALL SCALE, 4-13
- BAP, 3-44
- BASE FRAGMENT, 3-6
- Basic Conventions for Using Menus, 1-19
- Basic Operations, 1-22
- BEST, 5-38, 5-67
- BETTER, 5-67, 5-70
- BG_CARBO, B-6
- BG_DNA_DOUBLE, B-7
- BG_DNA_SINGLE, B-7
- BG_DNA_TRIPLE, B-8
- BG_LIPID, B-5
- BG_ORG_FRG, B-2
- BG_ORG_FUN, B-2
- BG_PEP_D, B-3
- BG_PEP_L, B-3
- BG_SOLVENT, B-3
- BIG STRAIN 2, 2-10

Index-2

- BIODESIGN, 2-3
- BioDesign General Format, D-3
- BND-BND TORSION, 5-46
- BND-X-ANG INV, 5-46
- BOND, 3-16, 5-64
- BOND (1-2), 5-56
- Bond Angle Bend, G-6
- Bond Angle Bending Parameters, I-15
- Bond Stretch, G-3
- Bond Stretching Parameters, I-14
- BONDS, 5-17, 5-46
- BOND-X-ANGLE, 5-64
- BOX SOLVENT, 3-73
- Branching and Statement Labels, K-13
- BROOKHAVEN, 2-3, 2-8
- Brookhaven, D-1
- Brookhaven Connect Files, D-37
- BROWSER, 7-3
- Browser Box, 1-16
- BUILD, 3-36, 3-49
- Building Peptides, O-2
- Building RNA or DNA, O-1
- Build Operations, 3-1
- BY ENERGY, 5-68, 5-70
- BY FILE, 2-21
- BY GROUP, 2-21
- BY RMS, 5-68, 5-70
- C OFFSET, 3-37, 3-40
- CALC LEVEL, 3-38
- CALCULATE, 4-16
- CALCULATE GRID, 6-51
- CAMBRIDGE, 2-4
- CANONICAL (TVN) Dynamics, 5-30
- CAPTION, 2-13
- CARBOHYDRATE, 3-52
- CARBO LIBRARY, 3-55
- CENTER, 3-18, 4-24, 5-41
- CENTER NEXT/CENTER ALL/NO
 CENTER, 2-21
- CENTROIDS, 6-29
- CHAIN, 3-22, 6-4
- CHAINS, 4-4
- CHANGE COLORS, 4-25
- CHANGE FONT, 2-17
- CHARGE, 4-21, 5-64
- CHARGE CONVERGE, 3-24
- Charge Equilibration Parameters, I-10
- CHARGES, 3-23, 3-68, 5-13
- Charges, G-13
- CHECK, 5-63
- CHEMLAB CHM, 2-4, 2-9
- CLEAN, 3-18
- CLEAR, 3-78, 5-73, 6-5, 6-13, 6-24, 6-26
- COLOR, 2-15, 3-67
- COLOR BY ATOM/ELEMENT/FORCE
 FIELD, 3-67
- COLOR BY ENERGY, 5-73
- COLOR BY RMS, 5-73
- COLOR BY SEQNCE, 5-73
- COLOR INCREMENT, 5-74
- COLORS, 6-44
- COLUMNS, 6-40
- Combination Rules (Off Diagonal vdW),
 G-19
- Command and Macro Modes, K-1
- Command Mode, K-1
- Command Summary, K-17
- Communicating with the User, 1-14
- CONDENSE, 3-8
- CONNECT, 3-65, 3-66, 3-93
- CONNECT ALL/NOT CONNECTED, 3-93
- CONNECT FRAG, 3-9
- Constant Energy Constant Volume
 Dynamics, P-2
- Constant Enthalpy Constant Pressure
 Dynamics, P-2
- Constant Temperature Constant Pressure
 Dynamics, P-3
- Constant Temperature Constant Volume
 Dynamics, P-3
- CONSTRAINED MIN, 5-32
- CONSTRAINTS ON/OFF, 5-37
- CONSTRAINTS, 5-8, 5-46
- Constraints Files, D-19
- Contact Check Analysis Variables and
 Switches, D-36
- CONTACTS, 6-25
- Continuation Characters, K-2
- CONTOUR, 6-42
- Conventions, G-20
- Conversion File Format, I-18
- CONVERT, 3-91
- COPY, 3-8
- COPY FILE, 2-12

- Cosine Expansions, G-6
- COULOMB, 5-46, 5-56
- Coulomb Summations, G-14
- CREATE, 4-9
- CREATE PS FILE, 2-19
- CRYSTAL, 3-59
- CSSR, 2-8
- CUBIC VOLUME, 6-54
- CURRENT, 4-20
- Custom Plots, 6-34
- CYLINDER SCALE, 4-13
- CYLINDERS, 4-12
- D_LIBRARY, 3-29, 3-36
- Data Files, C-1
- Debugger, K-16
- DEFAULT ATOM, 3-11, 3-19
- DEFAULT COLOR, 2-17
- DEFAULT FONT, 2-17
- Default Potential Terms for the Standard Force Fields, G-23
- DEFAULTS, 3-19, 3-66, 5-16, 5-38, 5-40, 5-44, 6-22, 6-52
- Defaults File, D-31
- Defaults Parameters (optional), I-3
- DEFINE, 2-14, 3-56, 6-31, 6-38
- DEFINE ANGLE, 3-31, 3-32, 3-35, 3-36, 3-54
- DEFINE TRAJECT, 5-66
- DEL RE, 3-24
- Del Re Charge by Element Parameters, I-11
- Del Re Charge by Force Field Parameters, I-11
- DELETE, 3-15, 3-39
- DELETE ALL H, 3-89
- DELETE BOND, 3-17
- DELETE BONDS, 3-95
- DENSITY, 4-16, 6-52
- Dial Commands, K-4
- Dialog and Alert Boxes, 1-15
- Dials Unit, 1-4
- Dielectric Constant, G-15
- DIFFERENCE, 6-12
- DIHEDRAL, 3-17, 6-9
- DIHEDRAL (1-4), 5-57
- Dihedral Angle Torsion Parameters, I-17
- Dihedral Angle Torsions, G-10
- DIRECTION START, 5-29, 5-32
- DIRECTION STOP, 5-29, 5-32
- DISPLAY DISTANCE, 6-26
- DISPLAY ENERGY, 6-24, 6-26
- DISPLAY FORCE, 3-85, 6-24
- DISPLAY GEOM, 6-24
- DISPLAY, 2-15, 5-72, 6-45, 6-54
- DISPLAY, TOP MENU, and COMMAND Buttons, 1-13
- DISTANCE, 3-16, 4-8, 5-8, 6-8
- DISTANCE CUTOFF, 3-38, 3-40, 3-85
- DISULFIDE, 6-11
- DISULFIDE BONDS, 6-49
- DIVISIONS, 4-14
- DNA-OLD, 3-46
- DOCK, 3-80
- DOCK DEFAULTS, 3-85
- DOT, 2-6
- DOUBLE, 3-42
- DRAG, 3-14
- DRAG-CLEAN, 3-14
- DRAW, 3-9, 3-16
- DRAW FREQUENCY, 5-47, 5-51
- DRAW ON/OFF, 5-37
- DREIDING: A Generic Force Field for Molecular Simulations, L-1
- Dreiding-I and Dreiding-II Force Fields, I-19
- DYNAMICS, 5-19
- Dynamics Algorithms, J-1
- Dynamics Switches, D-33
- DYNAMICS VAR, 5-21, 5-23, 5-51
- E CONVERGE, 5-33
- EDIT, 3-12, 3-36, 3-40, 3-45, 3-46, 3-51, 3-58, 3-68
- EDIT CHARGES, 3-25
- EDIT ID, 3-20, 3-40, 3-45, 3-46, 3-51, 3-58, 3-36
- EDIT RESIDUE, 3-91
- EEX OPTIONS, 5-3
- ELEC ENERGY CUT, 6-23
- ELECTROSTATIC, 4-21
- Electrostatic Interactions, G-13
- Energy and Force Analysis Variables, D-35
- ENERGY DIFF, 5-68, 5-70, 5-74
- ENERGY DOCK (Interactive Docking), 3-83
- ENERGY-FORCE, 6-17

- Energy Parameters, D-32
- Energy Switches, D-33
- ENERGY VAR, 3-86, 5-46, 6-23
- Entering File Names, 1-6
- EPSILON, 5-56
- Equilibrium Molecular Dynamics, P-1
- EXAMINE, 5-36
- EXCLUDE RANGE, 5-5
- EXCLUDE SELF, 3-73
- EXCLUDE SINGLE, 5-5
- EXECUTE, 3-24, 3-32, 3-34, 3-40, 3-74,
3-77, 3-81, 3-83, 3-95, 4-17, 4-21, 5-21,
5-25, 5-34, 5-38, 5-43, 6-26, 6-29
- Exit Operations, 8-1
- EXOCYC TORSION, 5-59
- EXTEND, 3-50, 3-57
- EXTEND CELL, 3-64
- EXTRACT, 5-71, 6-36
- EXTRACT MOLEC, 3-65
- FAST SETUP, 5-3
- FF PARAMETERS, 2-5, 2-10
- File Formats, D-1
- FILE ORIGIN, 3-78, 6-37, 6-47
- FILE ORIGIN, GROUP ORIGIN, PICK
ORIGIN, 3-87
- FILE ORIGIN, GROUP ORIGIN, PICK
ORIGIN, ZERO ORIGIN, 3-82, 4-24
- FILES, 4-3
- Files With All Their Hydrogens (no implicit
hydrogens), O-2
- Files With Hetero Hydrogens (but no
hydrogens on carbons), O-2
- Files Without Hydrogen Atoms, O-2
- FINAL TEMPER, 5-49
- FIX ALL, 5-5
- FIX FORCE, 6-24
- FIX GROUPS, 5-6
- FIX RANGE, 5-5
- FIX SELECTED, 3-19
- FIX SINGLE, 5-5
- Force Field Type Labels Parameters, I-9
- Force Fields, G-1
- FORCE SCALE, 3-85
- FORGET RANGE, 4-8
- FORWARD, 3-78
- Fragment Libraries, B-1
- FULL FILE, 4-3
- Function Keys, 1-8
- FUNCTIONAL GRP, 3-6
- Functions Menus, 1-18
- Functions on All Dynamics Menus, 5-21
- GASTEIGER, 3-24
- Gasteiger Charge Parameters, I-12
- General Defaults , D-32
- Geometric Mean, G-19
- GEOMETRIC MEAN/ARITHMETIC
MEAN, 5-55
- GEOMETRY, 6-6
- GEOMETRY PROP, 5-16
- Graphical User Interface (GUI), 1-11
- Graphics Parameters, D-31
- GRAPHICS, TEXT, and PLOT Buttons,
1-13
- Graphics Window, 1-12
- GRAY/BLACK&WHITE/COLOR, 2-19
- GRID SEARCH, 5-35
- GRID SPACE, 3-72
- GRID TYPE, 3-72
- GROSS CHARGE, 4-19
- GROUP, 2-5, 2-10, 6-3, 6-7
- Group Files, D-21
- GROUP ORIGIN, 3-79, 6-14, 6-24, 6-37,
6-47
- GROUPS, 4-18, 6-28
- H BOND DON/ACC, 4-21
- H BOND ON/OFF, 5-37
- H BONDS, 4-10, 5-46
- H BOND UPDATE, 5-42
- H BOND VAR, 5-58
- H MASS=H, H MASS=D, H MASS=T,
3-89
- HALF BOND, 2-20, 3-66, 3-73, 4-13
- HB ANGLE CUT, 5-58
- HB ANGLE OFF, 5-58
- HB ANGLE ON, 5-58
- HB DIST CUT, 5-58
- HB DIST OFF, 5-58
- HB DIST ON, 5-58
- HEAT EQUIL FREQ, 5-52
- HEAT INCREMENT, 5-50
- HOOVER/NOSE, 5-30
- Hydrogen Bond, G-21
- HYDROGEN BOND, 5-64
- Hydrogen Bond Parameters, I-18

- Hydrogen Bond Variables, D-33
- HYDROGEN BONDS, 5-17
- IDENTIFY, 6-2, 6-12, 6-36, 6-46
- If..then..else Blocks, K-12
- Implicit Hydrogen Relationships
 - Parameters, I-10
- IMPULSE Dynamics, 5-29
- In-Out Operations, 2-1
- INCREMENT, 5-32
- INFO LEVEL, 5-49, 5-52
- INITIAL TEMPER, 5-49
- INNER CUTOFF, 3-72
- INPUT COORD, 3-61
- INPUT FILE, 3-60
- Input/output Statements, K-14
- Inquire and Set Functions, K-10
- INSERT, 3-35
- INSERT AFTER RES, 3-35
- INTERSECTION, 6-53
- Introduction, 1-1
- INVERSE, 6-53
- INVERSION, 5-9, 5-64, 6-9
- Inversion (Improper) Torsion Parameters,
 - I-17
- INVERSIONS, 5-17
- Inversions, G-11
- INVERT, 3-15
- K(ANGLE), 5-10
- K(DISTANCE), 5-10
- K(INVERSION), 5-10, 5-12
- K(TORSION), 5-10, 5-11
- KEEP L/D, 5-12
- KEEP RANGE, 4-8
- KEEP TRANS/CIS, 5-11
- KEY ONE RES, 4-7
- KEY RES RANGE, 4-7
- L_LIBRARY, 3-29, 3-36
- LABEL COLOR, 7-2
- LIBRARY, 3-33
- LIMITS, 4-5
- LINE, 5-32
- LINE WIDTH, 6-45
- LINKAGE, 3-53, 6-10
- LIPID, 3-48
- LIST, 5-67
- Logical Expressions, K-12
- Lone Pairs Parameters, I-10
- Long-Range Cutoffs, H-1
- Loop Constructs, K-13
- M-CARLO SEARCH, 5-40
- Macro Files, C-3
- MACROMODEL, 2-4, 2-9
- Macros, K-7
- MAKE GROUPS, 4-2
- MAKE PLOT, 6-33
- Manipulating the Graphics Display Image,
 - 1-4
- MATCH, 6-28
- Math Functions, K-11
- MAX CHARGE, 4-19
- MAX DISTANCE, 3-94
- MAX ELEC VALUE, 4-19
- MAXIMUM DIST, 5-48
- MECHANICS, 3-27, 3-36, 3-40, 3-45, 3-47,
 - 3-51, 3-58, 5-14
- Menu Commands, K-1
- Menus, 1-12
- MERGE, 6-46
- METHOD (ATOMS), 5-47
- MIN DISTANCE, 3-94
- MINIMIZE, 3-86, 5-15
- MINIMIZE VAR, 5-47
- Miscellaneous Energy Variables, D-34
- MISC VAR, 5-59
- MISSING PARAM, 5-63
- MODE, 4-28
- MODIFY, 3-32, 3-36, 5-63
- MODIFY CARBO, 3-54
- MODIFY COORDS, 3-81, 3-84
- MODIFY EEX, 5-5
- MODIFY H, 3-88
- MODIFY PARAMS, 5-61
- MODIFY SUGAR, 3-45
- Modules/Messages Window, 1-12
- MOLEDIT, 2-4, 2-9
- MOLFILE, 2-4, 2-9
- MOMENT, 3-90
- MOPAC, 2-4, 2-9
- Mouse, 1-6
- MOVE, 3-14
- MOVE ALL, 5-5
- MOVE RANGE, 5-5
- MOVE SINGLE, 5-5
- Multiple Commands, K-2

- N OFFSET, 3-37, 3-40
- NB CUTOFF DIST, 6-23
- NBLST CUT, 5-55
- NBOND ON/OFF, 5-37
- NBOND UPDATE, 5-41
- NBONDS, 3-77
- NBRANCHES, 3-77
- NEGATIVE/POSITIVE, 2-19
- NEW ATOM, 5-64
- NEW ATOM IDS, 3-5
- NEW GROUP, 6-5, 6-13, 6-24, 6-26, 6-38
- NEW RESIDUE, 3-34
- New Dir and New File Buttons, 1-17
- No-opt Commands, K-3
- NONBOND VAR, 5-54
- Nonbond Exclusions, G-15
- Nonbond Lists, G-16
- Nonbond Off-Diagonal, G-20
- NONBONDS, 5-17, 5-46
- Nonbond Switches, D-33
- Nonbond Variables, D-33
- Non-Hamiltonian Forms of Constant Temperature Dynamics, P-3
- Nonperiodic, G-2
- Nosé-Hoover Canonical Dynamics, P-4
- NSTEP, 3-86
- NTORSIONS, 5-36
- Off-diagonal Nonbond Parameters, I-14
- OFFDIAG NONBND, 5-64
- OFFSET, 4-29
- OLD RESIDUE, 3-34
- OMEGA, PHI, PSI, CHI1, CHI2, CHI3, CHI4, 3-31
- ON/OFF, 4-28
- ONE ENERGY, 5-15, 6-16
- ORGANIC, 3-3
- ORIENT, 2-16
- ORIGIN, 6-42
- Other Macro Functions, K-15
- Other Macro Variables, K-9
- OUTER CUTOFF, 3-72
- PAIRS, 6-28
- PARAM INOUT, 5-62
- Parameter File Format, I-2
- Parameter File Header, I-3
- Parameters, I-1
- PEPTIDE, 3-28
- PEPTIDE LINKAGE, 6-48
- Periodic, G-3
- Personal Stereo Viewer for the IRIS, 1-10
- PHENYL CUTOFF, 6-11
- PHENYL RINGS, 6-11
- PHI-PSI ANGLES, 6-48
- Picking a List of Atoms, 1-15
- Picking a Range of Atoms, 1-14
- PICK ONE RES, 4-8
- PICK ORIGIN, 3-79, 6-15, 6-24, 6-37, 6-47
- PICK START RES, PICK STOP RES, 3-32, 3-39, 4-7
- PLANE, 5-32, 6-9
- PLOT, 6-38
- PLOT DEFAULTS, 6-33, 6-39
- Plot Files, D-30
- PLOT TYPE, 6-40
- POINT, 5-32
- POINTS, 5-33
- POLYENE, 5-59
- Polyene, G-10
- PRIMITIVES, 4-6
- PRINT SCREEN, 2-19
- PROBE RADIUS, 4-17, 6-52
- PROGRAM RESET, 7-2
- PROP DEFAULTS, 4-19
- PROPERTY, 4-20
- PROTEIN, 6-48
- PROTEIN CONSTRNT, 5-11
- Q EQUILIBRATE, 3-23
- QUENCH STEP, 5-27
- QUENCHED Dynamics, 5-27
- R*EPSILON, 5-56
- RCUTOFF, 5-17
- RDTRJ, N-2
- READ, 2-2, 4-17, 5-69, 6-55
- READ CONSTRAINT, 5-10
- READ CONV FILE, 3-91
- READ FRAGMENT, 3-7
- READ INCREMENT, 6-32
- Reading and Building Structures with the AMBER Force Field in BIOGRAF, O-1
- Reading DNA BioDesign General Format or Brookhaven Files, O-1
- Reading Peptide BioDesign General Format Files (created using the Dreiding force fields), O-2

- Reading Peptide Brookhaven Files, O-3
- READ LIBRARY, 3-5, 3-30, 3-33, 3-42, 3-48, 3-53
- READ PAR FILE, 5-62
- READ PART, 4-17
- READ PLOT, 5-39, 6-33
- REBUILD, 3-37
- Referencing Variables, K-9
- RENAME FILE, 4-9
- RENAME GROUP, 4-9
- RENDER GROUPS, 4-12
- RENUMBER, 3-36, 3-39
- REPEAT, 3-56
- REPLACE, 3-8, 3-33
- REPLOT, 6-35, 6-45
- RESET ALL, 3-82, 4-24
- RESET ALL PARAM, 5-62
- RESET BEST, 5-38
- RESET DIALS, 3-82, 4-24
- RESET INVERSION, 5-12
- RESET TORSION, 5-11
- RESET VELOCITY, 5-21, 5-24
- RESET, 5-10, 5-43, 6-45
- Resetting the Program, 1-3
- RESIDUAL, 6-53
- RESIDUE, 3-22
- RESIDUE NAME, 6-4
- RESIDUES, 4-4
- RETURN, 3-36
- REVERSE, 3-78
- RGB VALUE, 4-25
- RIBBONS, 4-10
- RIGID MOLECULE, 5-3
- RIGID/QUATERN, RIGID/EXP(ROT), 5-60
- RINGS, 5-41
- RMS COORDS, 5-76
- RMS DIFFERENCE, 5-68, 5-70, 5-75
- RMS FORCE CVRG, 5-16, 5-48
- RNA-DNA, 3-41
- ROCK, 4-31
- ROTATE ALL, 3-82, 3-84
- ROTATE BOND, 3-16, 3-75, 3-87
- ROTATE/TRANSLATE, 2-5, 2-11
- Rotations, F-1
- Running the Program in the Background, 1-21
- SALT BR CUTOFF, 6-12
- SALT BRIDGES, 6-11
- Sample Macros, K-18
- SAVE AVG COORD, 5-53
- SCALE, 3-67, 3-94
- SCALE 1-4, 5-57
- SCALE FACTOR, 6-23
- SELECT, 3-13
- SELECT OBJECTS, 2-14
- Selecting Menu Boxes and Buttons and Picking Atoms on the Display, 1-14
- SELECTORS, 4-7
- SEPARATION, 3-37, 4-29
- SET BEST, 5-38
- SET BOND, 3-16
- SET ORIGIN, 3-82, 4-24
- SETUP EEX, 5-4
- SETUP, 3-4, 3-23, 3-28, 3-42, 3-46, 3-48, 3-52, 3-60, 3-70, 3-76, 3-93, 4-3, 4-18, 5-36
- SHADING TYPE, 4-14
- SHOW, 3-13
- SHOW ALL, 5-73
- SHOW ALL/SHOW CURRENT, 5-11
- SHOW EEX, 5-7
- SHTRJ, N-2
- SIMPLE DYNAMIC, 3-86
- SIMPLE DYNAMICS, 5-15
- SIMULATE, 3-27, 3-36, 3-38, 3-40, 3-45, 3-47, 3-51, 3-58
- Simulate Operations, 5-1
- SINGLE, 3-42, 6-6
- SINGLE ATOM, 6-2
- SINGLE ENERGY, 6-18
- SNGL FORCE/ATOM, 6-21
- SOLVATE, 3-69
- SOLVENT, 3-71
- SP2-SP2, 5-41
- SPACE GROUP, 3-68
- Special Commands Summary, K-5
- SPECULAR INTEN, 4-14
- SPHERES, 4-12
- SPLIN OFF, 5-55
- SPLIN ON, 5-55
- Standard File Name Extensions, D-9
- Standard Group Elements, A-1
- Standard Plots, 6-34

Index-8

- Standard Residues, A-1
- Standard Space Group Names, M-1
- Standard Structural Descriptors (Selectors), A-3
- Standard Structural Elements (Primitives), A-2
- Standard Verlet, J-1
- START, 5-66, 6-32
- Starting BIOGRAF, 1-2
- STEREO, 4-27
- STEREOCHEMISTRY, 6-48
- STOP, 5-66, 6-32
- Summed Verlet, J-3
- Supplementary Files, C-1
- SURFACE, 4-15
- SURFACE AREA, 6-54
- SURFACE ONLY, 6-53
- Surface Display Functions, 4-18
- Surface File Format, 4-22
- SYMBOL, 3-66, 6-43
- TAU (CAN), 5-30
- TAU (DMP), 5-31
- TEMPDAMP (TVN) Dynamics, 5-31
- TEMPER WINDOW, 5-50
- TEMPERATUR VAR, 5-21, 5-22, 5-49
- Theta Expansions, G-8
- TIME, 5-21; 5-23
- TIME STEP, 5-51
- Titan's Stereo Monitor, 1-10
- Topology and Automatic Energy
 - Expressions, G-2
- TORSION, 5-9, 5-64
- TORSIONS, 5-17, 5-46
- TOTAL CHARGE, 3-23
- TOTAL COLORS, 4-20
- TRAJ DEFAULTS, 6-32
- TRAJECTORY, 6-31
- Trajectory Files, D-10
- Transferring Trajectory Files Between
 - Computers with Incompatible Trajectory File Formats, N-3
- TRANSLAT TEMPER, 5-29
- Two-State and Cyclic Toggles, K-3
- UNEXTEND CELL, 3-64
- UNION, 6-53
- Unit of Time, J-1
- Units, E-1
- UP CHRG EACH ITR, 5-46
- UPDATE, 3-78
- UPDATE FREQ, 5-33, 5-52
- UPDATE LIST, 3-87
- UREY-BRADLEY, 5-64
- Urey-Bradley, G-22
- USE ALL INVERS, 5-46
- USE ALL TORSION, 5-46
- USE VELOCITY, 5-21
- USER, 2-5, 2-10, 3-44, 7-1
- USER BASE FRAG, 3-6
- USER ENERGY, 5-46
- User Files, C-4
- USER FUNCT GRP, 3-6
- USER LIBRARY, 3-29, 3-33, 3-36, 3-56
- USER NONBOND, 5-55
- User Routines, K-16
- USER TRAJECT, 5-76
- User-supplied Parameters, I-18
- Using the Menus, 1-18
- Utilities, 7-1
- VAN DER WAALS, 5-64
- van der Waals Interactions, G-16
- Variable Names and Values, K-8
- Variables, K-8
- VDW ENERGY CUT, 6-23, 6-26
- VDW SCALE, 4-17, 5-42, 6-52
- VECTOR, 2-6
- VECTOR WIDTH, 4-13
- VECTORS, 4-12
- Viewing the Image in Three Dimensions, 1-10
- Visualize Operations, 4-1
- VOLUMES, 6-50
- WINDOW, 6-40
- WITHIN, 5-67, 5-70
- WRITE, 2-7, 4-22, 6-29, 6-55
- WRITE BGF, 6-37
- WRITE CONSTRAINT, 5-10
- WRITE FILE, 6-13
- WRITE PAR FILE, 5-62
- WRITE PART, 4-22
- WRITE PLOT, 6-36, 6-47
- WRITE TRAJECT, 5-21, 5-24, 5-42
- WRITE TRJ FREQ, 5-52
- WRITE VELOCITY, 5-21
- WRTRJ, N-3

XYZ RESTORE, 3-87, 5-18
XYZ TEMP SAVE, 3-87, 5-18
ZERO CHARGES, 3-26
ZERO ORIGIN, 3-79, 6-37, 6-47
-,0,+, 4-21

