Introduction

Turbo-Frodo is a general-purpose molecular modeling environment. It is designed for people who need to model de novo macromolecules, polypeptides, and nucleic acids, by building up these macromolecules from experimental X-ray crystallographic and NMR data and displaying the resulting models in various forms. These forms include Van der Waals and Connolly's molecular dot surfaces as well as spline line surfaces and secondary-structure representation. Compact views are possible, including CPK, icosahedra, and ball-and-stick styles, either individually or in various combinations.

Turbo-Frodo can color molecules, either objectively or subjectively, in order to compare them with existing structures and to evaluate their geometry. Turbo-Frodo can also fit and stack proteins, interactively mutate or chemically modify a protein, and assess the resulting conformational changes.

Turbo-Frodo reads and displays X-ray density maps. Electron density maps are calculated using the Mappage program, which is supplied along with Turbo-Frodo, giving a choice between 5 maps with 5 levels displayed on the screen. Skeletal maps are used as templates to start building up molecules with homologous protein fragments. A set of Crystallize options are available for various purposes, such as fitting molecules into electron density maps, taking the symmetry of molecules into account, and depicting molecular packing with water-accessible surfaces. These options can be used to evaluate the size of crystal channels and to decide whether diffusion or cocrystallization is the most feasible way to fit a ligand into the target protein.

Turbo-Frodo also reads and displays NMR spectra processed on Brucker equipment and displays on the screen any or all of the NMR constraints that were used to calculate a structure as well as the residual violations. A model can be interactively corrected in order to satisfy these NMR constraints while checking whether the geometry is accurate. Each residue, or even each atom, can be colored differently depending on the number of constraints in which it is involved.

With Turbo-Frodo, molecules are displayed full screen, and the functions take the form of pull-down menus, which are displayed only when they are needed. Turbo-Frodo enables the user to interact directly with a molecule via these pull-down menus and dials. Plot files can be generated on Postscript files, which can be printed out on any Postscript laser printer, such as that of a MacIntosh.

This software program is written in C language, making maximum use of the vast computing and graphics potential of the latest graphics workstations based on the Silicon Graphics Library. This is a brand new

program, apart from few routines such as REFINE and RIGID, which were previously part of Frodo. "Frodo" is included in the name of the new program not for code-homology reasons but simply because Turbo-Frodo is able to deal with Frodo style electron-density maps.

Turbo-Frodo is a graphic display program built onto a data base, the Heap File, in which molecules can be stored. The Heap File is a non-finite arborescent molecular data base that is limited in size only by users' disk capacity. Turbo-Frodo users can therefore have access to all the structures stored in the Protein Data Bank (PDB), providing they have enough space on their disk .

What follows are a few notes about symbols and stylistic matters that are used in this manual:

Quotation marks: Whenever you come across quotation marks in the examples given in this manual, this means that you have to replace what is in quotation marks by the name of the file that you would like to use by simply typing the name of your file. For example, if you would like to display the structure of trypsin, the instruction turbo>**load "molecule"** means that you have to type turbo>**load trypsin**.

Bold-faced type: Bold-faced type used in this manual means that you have to type the words, letters, or characters that are in boldface. For example, if instructions in the text tell you to use the turbo> **read pdb** command, this means that you have to type the letters "read pdb" after the prompt sign. Bold-faced type is also used in this manual for names of files, directories, subdirectories, programs, and commands in order to distinguish them from text matter.

Uppercase letters: Words that are written in uppercase letters mainly indicate program options.

Brackets: Brackets indicate the default answer. For example, in the Textport order FREAD PDB [WH], WH is the default answer. Generally speaking, whenever you enter a command, the default answer, if any, is indicated in brackets.

Pound sign: The pound sign (#) generally stands for the data base number of the molecule.

Star sign: Two stars (**) indicate a submenu or sub-submenu. The two stars generally appear in a list of options below a menu title and thus indicate a submenu, or sub-submenu, with options of its own.

Part
I

Getting Started

Part I provides basic operation instructions such as Turbo-Frodo's hardware and memory requirements, the various ways of interacting with the loaded molecule, and how to run Turbo-Frodo, among other necessary operation information. It also covers the various files that need to be prepared, such as the Heap File, before displaying a molecule on the screen.

Operation

Hardware Requirements

The Turbo-Frodo software program will run on Silicon Graphics workstations possibly working under IRIX 4.0 but more likely 5.2 and above. The program uses the RGB color mode and thus can work on an 8-bit graphic board. Nevertheless, for better usage convenience, it requires 2x12 bitplanes and a Z buffer in order to see not only nicer electron density maps but especially accurate and impressive representations of CPK and secondary structure.

Start-Up

Use the following UNIX order from the general directory **TURBO_DIR (ex:/disk_usr/turbo)**
>**tar xvop**
This makes it possible to generate the Turbo-Frodo tree. Modify your **.cshrc** as follows by typing,
>**setenv TURBO_DIR   /disk_user/turbo**
>**alias turbo TURBO_DIR/prog/turbo**
or execute the progam entitled EXECUTE-ME and follow the instructions. Start the program by typing >**turbo** or >**turbo -d** (select a default heap name, i.e., **test.heap**) or >**turbo -t** (cancel the display of the Turbo-Frodo starting logo, which can be useful if working on a non-graphic terminal).

Memory Requirements

For better performance, Turbo-Frodo requires some swap memory. If you have some free disk space, you should therefore try adding some swap memory. For machines running under system 5, you can add swap memory interactively. The following procedure adds 150Mb of swap space to the system by using a file in the **/swap** directory:

> /usr/sbin/mkfile 150m /swap/swap1
> /sbin/swap -a /swap/swap1

To make this swap area permanent (that is, automatically added at boot time), add the following line to **/etc/fstab**:

/swap/swap1 swap swap pri=300

User Interface

There are four ways of interacting with the loaded molecule: via the mouse, via the keyboard by using the Textport commands, via the pull-down menus, or via the button box (or pseudo-button box).

The Mouse

With the mouse one can either access the pull-down menus by clicking the right button or interact with the molecules on the screen by clicking the middle button. The left button comes into use in the atom-addition mode (see this specific function).

The following combinations of mouse buttons can also be used (Active, or clicked-on buttons, are indicated in boldface):

L **M** R: To pick an atom or a bones point.
L M **R**: To activate a pull-down menu option.
L M R: Used in the atom-addition mode.
L M **R**: To pop to the next residue.
L **M** R: To quit display to Textport.
L M R: To shift the structure displayed so that the atom selected is centered on the graphic screen.
L M R: This combination is still not dedicated.

A new feature now exists that helps you walk along the density map and have the closest possible view. You can now center on a density point, in the same manner as you do on an atom, by clicking the left and middle buttons on the mouse. When selecting a point on the screen, the program searches for the highest of the density points found along the current slab and centers on it. This feature is especially useful to look at empty MIR density maps.

Textport Commands

Textport commands consist of orders and options. An order needs a target, whereas an option does not. Both types of commands have to be validated with the ENTER key on your keyboard.

Pull-Down Menus

The pop-up menus are activated when the cursor is driven up to the top of the pop-up menu head. The submenus or the options can be accessed accordingly by dragging the mouse down, keeping the right button depressed. This will make the various options appear in yellow. To choose an option, drive the mouse along the menu until the option required is highlighted in yellow. Releasing the right button on the mouse will activate the option, which will then be highlighted in green. To turn off the option, go to it with the mouse and click the right button again. This will change the color of the option back to yellow and deactivate it.

The Button Box

The button box (or pseudo-button box) makes it possible to move the entire molecule, or part of it, on the screen, in all three dimensions, by performing translations and rotations. It is thus possible to set the distance for stereoscopic viewing. With the button box, one can also display only a section of the molecule and move this section along the Z-axis. If you do not have a button box, the Textport PSEUDO-DIALS option will display on the screen a set of cursors controlling the same options as the button box, via the right button on the mouse. The dial functions are summarized in the display option called HELP DIALS.

Running Turbo-Frodo

To run Turbo-Frodo, just type >**turbo** or >**turbo -t** if you are working on a non-graphic terminal. The **turbo -t** option may be useful if you want to prepare particular files such as the Heap File, or the files necessary for NMR purposes. This will set the program running.

If this procedure fails, and if you are on an appropriate terminal connected to an Iris workstation, ckeck the environment by typing >**printenv**. Check the value of the variable **TERM**, which should be set to **iris-ansi**. Otherwise, correct it by typing >**setenv TERM = iris-ansi**.

Turbo-Frodo will appear on your graphic screen or not, depending on whether you have started the program with or whitout the **-t** option, and a Textport window will open, prompting you to give the name of the Heap File. Enter a name. If the name you entered corresponds to an already existing Heap File, the program will wait for the next command. Otherwise it will ask you whether or not you want to create this file. Answer accordingly.

If your Heap File already contains a structure, it can be loaded by typing turbo >**load "molecule"** or turbo >**load "#",** # being the data base number of the molecule you want to load on the screen. The data base number may be obtained by typing turbo>**list heap**. Conversely, a molecule can be unloaded by typing turbo >**unload "molecule"** or turbo>**unload "#".** Once the molecule has been loaded, if you want to display it, simply type turbo>**go.**

Necessary Files

The Heap File

Before displaying a molecule on the screen, you need to prepare some files. First of all, you have to prepare the data bank in which all your molecules are stored, the Heap File. The Heap File is an infinite file and is limited only by your disk space. It contains data on the molecular structure in a format readable by Turbo-Frodo.

To prepare the Heap file, you have to read an ASCII coordinate file about your molecule. This file may by written either in the Protein Data Bank format (PDB) or in the Wayne Hendrikson format (WH). Accordingly, type either turbo>**read pdb** or turbo>**read wh.**

The program will then prompt you to give the molecule in the Heap File a name. This name should not be longer than 8 digits, otherwise only the first 8 digits will be recorded. The contents of the Heap File can be listed by typing turbo>**list heap.** Whenever you want to save changes in a molecular entry made on the screen, you must save it in the Heap File.

The Heap File is not readable without Turbo-Frodo. Therefore, if you want to use this information with other programs, it is necessary to prepare a file in the PDB, WH, or XPLOR readable PDB formats by typing turbo>**make pdb** or turbo>**make wh** or turbo>**make xpdb**, as appropriate. The program will prompt you to give the name of the output file.

While working within the Heap File, you can duplicate any existing stucture by typing turbo>**duplicate "filename".** The program will then prompt you to give the new name. You should avoid giving the duplicated molecule the same name. Conversely a molecule can be deleted from the Heap File by typing turbo>**delete "filename"**or turbo>**delete "#"**, # being the number attributed to the molecule in the Heap File.

The Surface Files

The Connolly Surface File and the Spline Connolly Surface. These files are located in remote programs, not in Turbo-Frodo itself. These programs are supplied, however, along with Turbo-Frodo and are placed within the /**ms** subdirectory.

The X-Ray Files
The Map File

The Map File has the same format as the classical DSN6 Frodo map. The Map File can be obtained from the Fourrier output of various X-ray packages (CCP4, PROTEIN, and so on) after conversion using the Mappage program, which is provided with Turbo-Frodo. All Mappage files required are placed under the **/map** subdirectory.

Assuming you have already loaded a molecule, you can then load an X-ray file by typing turbo>**load map.** The program will prompt you to give the name of the input file and to define the contour levels, colors, and so on, which are now required.

The Bones File

The Bones File must first be calculated with the supplied program named **bones_turbo**, which is located under the **$TURBO_DIR/map** subdirectory. It accepts maps from either CCP4 or XPLOR softwares. After calculation, the output file must be entered in Turbo-Frodo with the LOAD BONES Textport order.

The NMR Files
The Distance Constraints

This file contains the NMR constraints deduced from a nOe set collected on a NOESY spectrum recorded with a molecule. The format of an NMR file is the same as that used by XPLOR for NMR structure determination and refinement. It contains the name of the two protons between which a nOe has been determined, the distance determinated after calibration of the NOESY spectrum, and the lower- and upper-tolerance values.

In the following example,

assign (resid 1 and name ha) (resid 32 and name hb*) 4.0 2 1.0

a nOe has been observed between the alpha proton of residue 1 and one beta proton of residue 32. Since the methylene protons were not unambiguously assigned, a wildcard is used (hb*). This corresponds to the use of pseudo-atoms in distance geometry calculations. After calibrating the spectrum, the upper limit was set at 4.0 Å. The lower-tolerance limit was set so that the lowest acceptable distance would be the Van der Waals distance. The upper-tolerance limit compatible with the use of a pseudo-atom has been set at 1.0 Å.

Assuming you have already loaded a molecule, you can then load an NMR file by typing turbo>**load nmr**. The program will prompt you to give the name of the input file and then list the atoms present in the nOe file but not in the molecular structure. This often arises when you want to check experimental constraints on a structure originating from a PDB, which does not give the protons. In this case, you will have to generate these hydrogens by using, for example, the XPLOR package. Upon

typing **go**, the structure will appear on the screen with the constraints displayed in different colors, depending on the difference between the constraint given in the input file and the real interatomic distance.

The Proton File

This file contains a list of all the protons in your molecule, including the pseudo-atoms and the NMR chemical shifts coming from your experimental data or from the literature. It can be generated manually, but it will be automatically produced by the EASY NMR assignment software. The format is that used by EASY:

128 4.123 0.002 0.000e+00 HA   14 1

The first number is the arbitrary number of protons given. The second and third numbers are the proton's chemical shift and the degree of uncertainty, if any. The fifth entry is the atom identifying the proton, and the sixth is the number of the residue present in the sequence to which this proton belongs. The fourth and the seventh entries are not used for our present purposes.

Assuming you have loaded your molecule, type turbo>**nmr-spect.** The program will then prompt you to give the proton file, the peaks file, and the reference peak file output names (with **filename.peaks** and **filename.ref** formats). The **filename.peaks** then contains a list of theoretical NOESY peaks, the relative intensities of which are deduced from the interproton distances determined on the loaded structure, assuming the existence of a rigid molecular structure.

This corresponds to a rough calculation of a theoretical spectrum. The peaks file generated can be directly used in EASY to display on a spectrum the location of theoretical NOESY peaks and thus to compare the theoretical and experimental locations. The main advantage of this option is that it helps end the assignment of NOESY spectra during the NMR refinement process.

The DGNL Data Base

The DGNL data base has the same structure as the Heap File. It contains the structural data on several molecules, with which a selected fragment of the loaded molecule can be compared. This data base already exists when you install Turbo-Frodo, but you can also create your own data bank by using the following protocol:

Generate a heap data called **dgnl.heap** by starting Turbo-Frodo. The program will ask you for the name of the Heap File. Type **dgnl.heap**. Then fill up this heap by using the following option, turbo>**read pdb** or turbo>**readwh**. You then have to stop the program. Restart it with another Heap File and type turbo>**make-dgnl**. This option calculates the distance matrix, and the output is the **diag.mtx** file. If you want to use the

latest DGNL data base, you must copy the 2 files (**dgnl.heap** and **diag.mtx**) on the **$TURBO_DIR/dict** directory.

P a r t

II

Textport Commands See also Textport OrdersSee also Textport Options

Textport commands consist of orders and options. An order needs a target, whereas an option does not. Both types of comands have to be validated with the ENTER key on your keyboard.

Philosophy of Use: **Textport commands apply to the most recently selected or most recently loaded molecule. Loading or selecting a molecule will turn the prompt "Turbo>" into "name>", "name" being the name the loaded molecule has in the Heap File.**

As an example, loading or selecting the structure of trypsin will call up the following stretch:

turbo>load trypsin
trypsin>
or
toxin>select trypsin
trypsin>

However, to simplify this manual, the Textport command example will always use the prompt **turbo>,** keeping in mind, though, that it can actually be replaced by the name of the most recently selected or most recently loaded molecule. Generally speaking, whenever you type a command, the default answer, if any, is proposed in brackets.

CHAPTER 3

Textport Orders: Main Commands to Manage the Files

DELETE ATOMS [MOLECULE] RESIDUES.
FREAD PDB [WH].
LIST ATOMS BONES [HEAP] MAP MOLECULE RESIDUES SURFACE.
LOAD BONES MAP [MOLECULE] NMR SEQ SURFACE.
MAKE ACCESS MS PDB SS [WH] XPDB.
RAD BONES MAP NEIGHBOR [SPHERE] SYMMETRY.
READ ACCESS PDB SS [WH].
SELECT [MOLECULE].
UNLOAD BONES MAP [MOLECULE] SEQ SURFACE.

DELETE: ATOMS [MOLECULE] RESIDUES.
The DELETE option removes data from the disk data base. For example:

del trypsin
del 5 trypsin (where 5 is the number of the molecule)
del H atom all (delete all the H atoms from the selected molecule)
del res 5 12 (delete residues 5 to 12, inclusively)

FREAD: PDB [WH].
The FREAD option reads structural data from ASCII disk files with the Wayne Hendrickson (WH) or Protein Data Bank (PDB) formats when the coordinates are fractional. You have to enter the cell parameters before reading your molecule. For example:

fread wh (hit ENTER)
> enter name of the molecule (8 car): **toxin**
> enter name of input file (.wh): **toxin.wh**
Enter the cell parameters: **a,b,c,**α**,** β**,** δ**,**
…
> enter name of input file (.wh) or '.' to finish: .

LIST: ATOMS BONES [HEAP] MAP MOLECULE RESIDUES SURFACE.
With the LIST option you can generate a list of information from memory data (atoms, residues, molecules), disk data (heap), or file information (bones, map, surface). This option works on the molecule selected. For example:

li heap (give heap contents)
li res RANGE (list the residues of the molecule selected)
li ato RANGE
li map
RANGE=ZONE,ZONE, ZONE (20 times maximum)
RANGE=all (all the residues of the selected molecule)
ZONE=NRES1 NRES2 (stretch from residues NRES1 to NRES2)
ZONE=NRES (stretch with only residue NRES)
NRES=residue name (including chain label if exists)

Ex: 15 25
10
all
10 25, 30, 35, 45
121 a 133 a (a=chain label)

LOAD: BONES MAP [MOLECULE] NMR SEQ SURFACE.
The LOAD option loads molecules (from the heap disk storage), bones, maps, NMR constraints and surface (from the respective files). For example:

lo trypsin
lo 5 (where 5 is the number of the molecule)
lo 4 6, 8 (where 4 6 specifies a range and 8 a single molecule)
lo all (load all the molecules in the heap)
lo map (hit ENTER)
>enter name of the file: **trypsin.map**

MAKE: ACCESS MS PDB SS [WH] XPDB.
With the MAKE option, you can write ASCII files on a disk with Wayne Hendrickson, Protein Data Bank, and Connolly Molecular Surface (MS) formats or with a special PDB format for the XPLOR program (XPDB). The coordinate source is the memory storage. For residue accessibility (ACCESS) and secondary structure (SS), the information is dumped onto the disk under a classical ASCII file. The name is requested on the screen. For example:

make pdb (hit ENTER)
>name of the outpout file: **trypsin.pdb**
>symmetry number (0 for initial molecule): **0**
Do you want selection ? (y/[n]) **n**
Residue range: **RANGE**

The symmetry number is based on the Turbo system of numbering. (Pick any atom belonging to the molecule in question to find out its number.)

RANGE=ZONE, ZONE, ZONE (20 times maximum)
RANGE = all (all the residues in the molecule selected)
ZONE =NRES1 NRES2 (stretch from residues NRES1 to NRES2)

13

ZONE = NRES (zone containing only residue NRES)
NRES=residue name (including chain label if exists)

Ex : 15 25
10
all
15 25, 30, 35, 45
121 a 133 a (a=chain label)

The MAKE XPDB option asks for an additional information to meet the **generate.inp** format requirements of XPLOR, i.e., >**segment name.** If you are not working on a multichain protein, just hit RETURN, otherwise give the segment name (up to 4 characters).

RAD: BONES MAP NEIGHBOR [SPHERE] SYMMETRY.
The RAD option defines the radius of several objects: BONES for the bones map, MAP to display the map around the displayed atoms, NEIGHBOR for the GEOMETRY NEIGHBOR option, SPHERE for the View menu option,and SYMMETRY for the SYMMETRY NEIGHBOR option. For example:

rad 5.5
rad neig 4

READ: ACCESS PDB SS [WH].
The READ option reads structural data from ASCII disk files with Wayne Hendrikson or Protein Data Bank formats. For residue accessibility (ACCESS) and secondary structure (SS), the information can be read from the Turbo ASCII file previously recorded on the disk, or from the ACCESS or SS file with the Turbo format. For Example:

read wh (hit ENTER)
> enter name of the molecule (8 car): **toxin**
> enter name of input file (.wh): **toxin.wh**
VAL 1 N 17.962000 20.219000 1.707000 9.290000

…
HIS 64 N 39.705002 18.393999 17.282000 12.350000
>enter name of the input file (.wh) or '.' to finish:.

SELECT: [MOLECULE].
With the SELECT option you can select the current molecule if it is already loaded in the memory. Several options (LIST,ZONE, among others) make use of the current selected molecule. For example:

sel trypsin
sel 5 (where 5 is the number of the molecule)

UNLOAD: BONES MAP [MOLECULE] SEQ SURFACE.
The UNLOAD option unloads molecules, maps, bones, and surfaces from the memory. For example:

unlo trypsin
unlo 5 (where 5 is the number of the molecule)
unlo 4 6, 8 (where 4 6 specifies a range, and 8 a single molecule)
unlo all (unload all the molecules loaded)
unlo map
>enter name of the file: **trypsin.map**

Textport Options:
48 Options to Parameterize and Display

ACCESS, ACON, BALL & STICKS, B-FACTOR, BFACT-DEF, BF-AVERAGE, CHANGE-NAME, COL_RAMP, COMPARE, COMPSEQ, CONNOLLY, CONTOUR, CP-HEADER, CPK, DEVIATION, DIST-CA, DOMAIN, DUPLICATE, END, GO, GR-ACCESS, GR-BF-DIFF, GR-PC-REFI, GR-RESET, GR-RMS-DEV, HELP, HYDROPATHY, INFOS, MAKE-DGNL, MAKE-DNA, NCS, NEIGHBOR, NMR-RIGID, NMR-SPECT, OCCUPANCY, PSEUDO-DIALS, READMULT, READ-NMRPDB, RENAME, RIGID, SAVE-E, SETLINEWIDTH, SLAB-DISPLAY, SOLVENT, SPLINE, SYMMETRY, SYSTEM, TOLERANCE, TRAJECTORY, TRANSLATE, VDW-DENSITY, WORK-MAP, ZONE.

ACCESS
The ACCESS option calculates the accessibility of each residue for a given protein, with respect to the solvent. You can further color the protein accordingly to the accessibility to the solvent. Acessibility scores are not saved until you activate the display option SAVE. You can keep accessibility scores in an ascii file by using the Textport order MAKE ACCESS (and read them with the Textport order READ ACCESS).

ACON
The ACON option centers the selected atom to display the current molecules, maps, surfaces, and so on. The input is the name of the residue and that of an atom (CA or first atom is by default). For example:

acon 23 (CA, if existing, or first atom by default)
acon 35 od1

BALL & STICKS
The BALL & STICKS option defines zones that are to be displayed as "balls and sticks". The non-selected parts keep the status previously defined, using the COMPACT SELECT screen option in the View menu. For example:

bal RANGE
RANGE = ZONE,ZONE, ZONE (20 times maximum)
RANGE = all (all the residues of the selected molecule)

ZONE = NRES1 NRSE2 (zone from residues NRS1 to NRES2)
ZONE = NRES (zone with only residue NRES)
NRES=residue name (including chain label is exists)

Ex :15 25
all
15 25,30,35 45
121 a 133 a (a=chain label)

B-FACTOR

The B-FACTOR option allows one to change the B-factor values of atoms, for a range of residues. For example:

b-factor
>Give the new B-factor: 36.5
>Residue range: RANGE
RANGE=ZONE, ZONE, ZONE (20 times maximum)
RANGE=all (all the residues of the selected molecule)
ZONE=NRES1 NRES2 (zone from residues NRES1 to NRES2)
ZONE=NRES (zone with only residue NRES)
NRES=residue name (including chain label if exists)

Ex: 15 25
all
15 25, 30, 35 45
121 a 133 a (a=chain label)

BFACT-DEF

The BFACT-DEF option defines the range of B-factors used by the FEATURE BFACT option in the display Color menu: For example:

bfac
Bfactor range min max: 5 25 (Give the minimum and maximum values shown by your protein.)

BF-AVERAGE

The BF-AVERAGE option provides the averages of the B-factors for the whole residues, the backbone, and the side chain. This option works for a range of residues. For example:

bf-aver
> Residue range: Range

Range: Zone, Zone, Zone (10 times maximum)
Range: all (all the residues of the selected molecule)
Zone: NRES1 NRES2 (zone from residues NRES1 to NRES2)
Zone: NRES (zone with only residue NRES)
NRES = residue name (chain label if exists)

17

EX:
15 25
10
all
15 25, 30, 35 45
121 1 133 1 (1 = chain label)

CHANGE-NAME
The CHANGE-NAME option changes the name of the molecular entry in the heap.

COL_RAMP
The COL_RAMP option displays on the screen the color ramp used for the BFACT, HYDROPATHY, or ACCESSIBILITY color features of the display Color menu.

COMPARE
The COMPARE option compares 2 different zones. The molecules required must be loaded before running this option. For example:

comp
> enter source molecule name: **toxin**
> enter target molecule name: **trypsin**
First residue of the zone in the source molecule: **25**
First residue of the zone in target molecule: **35**
number of residues in the zone (< 20): **15**
>enter the name of the deviation file: **test.dev**
> enter the name 1 or 2: **1**

You can display 2 graphics at the same time with the GRAPH option in the Plot menu. You can choose the graphic you want by giving the appropriate graph number. The deviation file gives the RMS deviation per residue. This file can be used as an input in the GR-RMS-DEV option.

COMPSEQ
The COMPSEQ option searches in the selected protein sequence the pattern that you specify. For example:

compseq
Enter the sequence to compare: 3-letter code, end with a period "."
ser
pro
ser
.
searching a sequence of 3 residues
sequence found 1 time
1 start 7A end 9A (indicates the location of the pattern in the protein)

CONNOLLY

The CONNOLLY option reads a Connolly Surface File. Surface points can be attributed to the nearest atom. This option can be therefore rather time consuming. The output file name is **file.ts** if the input file name is **file.srf.** For example:

conn
> enter the name of the input file (.srf): **toxin .srf**
> do you want to link surface to atom (1) or simple calculation (2): **1**
> number of points is 984

In this case, the output file name is **toxin.ts.** This file can be used as an input with the LOAD SURFACE option.

CONTOUR

The CONTOUR option centers the current molecules, maps, surfaces, and so on. The inputs are the XYZ coordinates of the point required. For example:

cont 1.5 3 .75
Center: 1.500   3. 000   0.750

CP-HEADER

The CP-HEADER option copies some information from one molecular entry to another. The former molecular entry can belong to another heap file. You can copy map information, space group and cell parameters, zones, or non-crystallographic symmetry information.

EX:
cp-h
>Do you wish a header from another heap file (y/ [n]) y
>Enter the heap name (<CR> to list): test.heap
>Enter the source molecule (<CR> to list): molecule name
>Do you wish to copy map information? ([y] / n)

CPK

The CPK option defines zones to be displayed as icosahedra. The non-selected parts keep the status previously defined using the COMPACT SELE screen option in the View menu. For example:

cpk RANGE
RANGE = ZONE, ZONE, ZONE (20 times maximum)
RANGE = all (all the residues of the selected molecule)
ZONE = NRES1 NRSE2 (zone from the residue NRS1 to NRES2)
ZONE=NRES (zone containing only residue NRES)
NRES=residue name (including chain label if exists)

Ex: 15 25

19

10
all
15 25,30,35 45
121 a 133 a (a=chain label)

DEVIATION

The DEVIATION option calculates the RMS between two structures without carrying out any fitting on these molecules. The program prompts you to give the file name of the two molecules, as well as the name of the output file and the kind of deviation file you want. Type 1 deviation file contains the RMS values (in Postscript form) and can be used with the GRAPH option. Type 2 file contains the RMS values in the ASCII format, which are thus ready to print.

DIST-CA

The DIST-CA option defines the ideal distance between 2 C Alpha atoms. This is important for the LIST MOL Textport option, the CA display option, and the MAKE FRAGS option from the ab initio construction part. The default is set to 5 angstroms.

DOMAIN

The DOMAIN option allows you to specify the different domains of your molecule, e.g., monomers, fragments, and so on. One domain can include different zones. These domains can be colored with different colors, and they are required for Non-Crystallographic Symmetries. For example:

domain RANGE
>#1 (ask for first domain)
RANGE: DOM, DOM, DOM (20 times maximum)
RANGE: all (all the residues of the selected molecule)
DOM: NRES1, NRES2 (domain from residues NRES1 to NRES2)
DOM: NRES (domain with only residue NRES)
NRES = residue name (chain label if exists)

EX: 15 25
10
all
15, 25, 30, 35, 45
121 a 133 a (a = chain label)

You can modify a previously defined domain by using the following syntax:

Ex: domain = ND DOM (Note that in this particular case you have to type the equal sign.)

ND is the number of the domain that you want to modify. To add a domain in the list, give a new ND number. Be careful with this step because if you

do not accurately use the above-mentioned syntax, previously defined domains will be removed.

DUPLICATE

The DUPLICATE option duplicates a molecular entry from the heap disk data base and files it under another name for further use and modification. For example:

dup toxin
>name of the new molecular entry (8 car)): **tox1**

END

The END option ends a Turbo session.

Ex: end

GO

With the GO option, you can leave Textport and start the interactive session.

Ex: go

GR-ACCESS

The GR-ACCESS option reads an accessibility output file and displays the corresponding graphics. Since Turbo-Frodo can display 2 graphics simultaneously, you must choose the one you want (graph number). For example:

gr-access
> enter the name of the XPLOR acccessiblity output file: **surf.acc**
> graph 1 or 2: **1**

You can display the graphics with the GRAPH option in the Plot menu. Be careful because the XPLOR accessibility output file does not have the same format as the MAKE ACCESS order, since if it had the same format, the MAKE ACCESS output cannot be used here.

GR-BF-DIFF

The GR-BF-DIFF option reads a B-factor diff output file to display the corresponding graphics. Since Turbo-Frodo can display 2 graphics simultaneously, you must choose the one you require (graph number). For example:

gr-bf
> enter the name of the Xplor B-factor diff output file: **bfact.dif**
> graph 1 or 2: **1**

You can display the graphics with the GRAPH option in the Plot menu.

GR-PC-REFI

With the GR-PC-REFI option, you can read an XPLOR PC refinement output file and display the corresponding graphics. Since Turbo-Frodo can display 2 graphics at the same time, you must choose the one required (graph number). For example:

gr-pc
> enter the name of the XPLOR output file: **filter.list**
> data 1 or 2: **1** (you can choose the column before or after refinement)
> graph 1 or 2:1

You can display the graphics with the GRAPH option in the Plot menu.

GR-RESET

The GR-RESET option initializes all the graph.

GR-RMS-DEV

The GR-RMS-DEV option reads a deviation file to display the corresponding RMS deviation graphics. Turbo-Frodo can display 2 graphics simultaneously, and again, you should choose the required graphics (graph number). For example:

gr-rms
> enter the name of the deviation file: **compare1.dev**
>graph 1 or 2: **1**

You can display the graphics with the GRAPH option in the Plot menu.

HELP

The HELP option provides information about all the Textport options. To obtain access to this help, you must have the **doc_text_port.txt** files in the **$TURBO_DIR/doc** subdirectory. For example:

help rigid

HYDROPATHY

The HYDROPATHY option makes it possible to enter the window (number of residues) with which the hydropathy index will be computed. The default is 5 residues.

INFOS

The INFOS option provides access to some predefined parameters. For example:

 info
RADIUS :

BONES 25.00 MAP 2.00 NEIGHBOR 4.00 SPHERE 7.00 SYMMETRY 25.00
B-FACTOR RANGE : min 4.oo max 20 .00

HYDROPATHY :size of the window is 5 residues

VDW DENSITY : 20 dp A2

PSEUDO -DIALS are turned ON

TOLERANCE : 0.2 A acceptable violation of NMR distance constraints.

MAKE-DGNL
The MAKE-DGNL option creates a new DGNL data base. First, you should generate a heap data base (called **dgnl.heap**), selecting your own structures. Second, you should stop the program and start it again with another heap file (for example, **test.heap**). Third, you should activate MAKE-DGNL. The program prompts you to give the name of the heap that contains the new DGNL data base (**dgnl.heap** if you called it so). This option calculates the distance matrix, and the output is **diag.mtx.** To be used in further Turbo-Frodo sessions, the two files, dgnl.heap and diag.mtx, must be moved to the **$TURBO_DIR/dict** directory. For example:

make-dgnl
> enter name of the heap:
DGNL > MOLECULE 0: TRYPSIN

…
DGNL > molecule 17: TOXIN

MAKE-DNA
The MAKE-DNA option generates the DNA molecule corresponding to the user's specifications. For example:

make-dna
> What kind of DNA: a A-DNA, b B-DNA or r RNA? **a**
> Watson-Crick base pairing? (y/n) **y**
> Number of residues in the two strands (n1 n2)? **10 10**
> Sequence 5'-> 3' in the first strand
agagagagag
Standard angle of rotation per residue? (y/n) **y**
> Standard A-DNA -> omega = 32.7 rise =2.60
> Strandard rotation and rise per residue? (y/n) **y**
You are constructing a type A DNA helix
> with variable rotations per residue
>and constant rise per residue?
> Name of the output pdb file dna. pdb: **test-dna.pdb**

NCS

The Non-Crystallographic Symmetries handling option, or NCS, serves to define the symmetry operation between two domains, read a matrix file for a given pair, read the pairs, read rotation matrix and translation from another heap entry, as well as remove previously defined pairs.

1-Define an operation symmetry
You will be asked for a pair of domains. These domains are listed and must be defined with Textport's DOMAIN option. To define the symmetry, you can refer to the RIGID option for help. The rotation matrix and translation, as well as pairs, are systematically saved in the heap.

2-Remove previously defined symmetries
You will be asked for a pair of domains. Note that this pair will be removed from the heap.

3-Read a matrix file
Give the name of the matrix file. The format is the same as the one produced by the display option PRINT NCS. However, you can add spaces, lines, and comments if they are specified by a '#'.

4-Read a matrix from another heap entry
You will be asked for a molecule name. Make sure that defined domains are well defined and relevant.

NEIGHBOR

The NEIGHBOR option generates a file containing all the contacts, below a cutoff value, between 2 zones in 1 or 2 differents structures. For example:

neig
> enter source molecule name: **toxin**
Residue range: **RANGE**
> enter target molecule name (return if the same): **trypsin**
Residue range: **RANGE**
Enter the cutoff: 4
Enter the name of the output file: **tox_tryp.neig**
RANGE = ZONE, ZONE, ZONE (20 times maximum)
RANGE = all (all the residues of the selected molecule)
ZONE = NRES1 NRSE2 (zone from residue NRS1 to NRES2)
ZONE=NRES (zone containing only residue NRES)
NRES=residue name (including chain label if exists)

Ex: 15 25
10
all
15 25,30,35 45
121 a 133 a (a=chain label)

NMR-RIGID

The NMR-RIGID option performs least-square rigid body fitting between a family of structures (source) and a single molecule (target). Please refer to the RIGID option (see below) for instructions about using NMR-RIGID.

NMR-SPECT

The NMR-SPECT option generates **peak** and **ref** files from a **prot** file. These files are used in the NMR assignment program EASY. For example:

nmr
>enter the name of the **prot** input file: **npy.prot**
> enter the name of the peak output file: **npy.peaks2**
> enter the name of **ref** output file: **npy.ref2**

OCCUPANCY

The OCCUPANCY option modifies or not the atom occupancy depending on the value of the B-factor. For example:

occup
>Give the new occupancy: 0.56 (accepts values between 0 and 1)
>to apply for all atoms with a B-factor more than (0 for all atoms): 34.5
(Enter "0" if you want to apply to all atoms no matter what is their B-factor)
>Residue range: RANGE
RANGE=ZONE, ZONE, ZONE (20 times maximum)
RANGE=all (all the residues of the selected molecule)
ZONE=NRES1 NRES2 (zone from residue NRES1 to NRES2)
ZONE=NRES (zone containing only residue NRES)
NRES=residue name (including chain label if exists)

Ex: 15 25
10
all
15 25, 30, 35 45
121 a 133 a (a=chain label)

PSEUDO-DIALS

The PSEUDO-DIALS option turns off the screen pseudo-dials instead of a real dial service. For example:

pseu
> Do you want to use the pseudo-dials ([y]/n)y)? **y**

READMULT

The READMULT option reads several **pdb** files in a unique command. It prompts you to give the name of a text file, which must be created in the following manner:

```
12            (number of files following)
MOL-1      (name of the first molecule in the heap file)
molecule-1.pdb (name of the pdb file)
MOL-2
….
```

READ-NMRPDB

The READ-NMRPDB option reads the special PDB files that contain several NMR structures. Enter a generic name for this set, and the different molecular entries will be named after it. For example:

>Enter name of the molecule (8 char.): TOXIN
>Enter name of input file (.pdb) (<CR> to list): toxine.pdb

The different molecules will be named TOXIN0, TOXIN1, TOXIN2, and so on. You can use the LOAD ALL order to load all the molecules.

RENAME

The RENAME option renames a molecular entry sequence. There are three possibilities: Change the starting count number, add a suffix, or both. You must load and select the molecule of interest before running this option. The changes are made in the memory. To save them, use the SAVE option in the Main menu. For example:

rename RANGE
1 Count   2 Suffix   3 Count + Suffix: **2**
one letter suffix: **a**

RIGID

The RIGID option performs a rigid body superimposition of a molecule (source) on another molecule (target). Both molecules have to be loaded in the memory. Source and target names are requested. At least three fitting points have to be given for the program to work properly. Note that in this unique case, the chain label, if it exists, has to be given before the residue name. For example:

rigid
> enter source molecule name: **crf_deb3**
> enter target molecule name: **crf_fin3**
source [CHAIN] RESIDUE ATOM or XYZ (<CR> to finish): **10 ca**
30.805000 15.312000 21.506001
target [CHAIN] RESIDUE ATOM or XYZ: **10 ca**
30.805000 15.312000 21.506001
source RESIDUE ATOM or XYZ (<CR> to finish): **20 ca**
20.934000 19.79.1000 19.263000
target RESIDUE ATOM or XY: **20 ca**
20.934000 19.791000 19.263000
source RESIDUE ATOM or XYZ (<CR> to finish): **25 ca**
21.493000 10.670000 17.232000

26

target RESIDUE ATOM or XYZ: **25 ca**
21.493000 10.670000 17.232000
source RESIDUE ATOM or XYZ (<CR> to finish) :
21.493000 10.670000 17.232000
matrix and translation applied
rot 1.000000 -0.0000002 0.000001
rot 0.000002 1.000000-0.000002
rot -0.000002 -0.000002 1.000000
tr 0.000002 -0.000002 0.000000
rms: 0.255903

Now comes the refinement process:

>max distance CA selected (<CR> to finish): **10**

The program will fit the two molecules using the alpha carbons of the two molecules that are less than the given cut-off, i.e., 10 Å, apart.

>RMS of 41 CA: **2.65**

The residues of the 41 molecules were closer than the cut-off and were therefore used to fit the two molecules. The resulting RMS value was 2.65 Å.

>max distance CA selected (<CR< to finish): **2**
> RMS of 25 CA: O.80

Another fit has been performed using alpha carbons which at this stage were closer than 2.0 Å (only 25 out of the 41 present in the sequence, which led to a better fit (RMS=0.80 Å).

SAVE-E
The SAVE-E option saves the work environment of the selected molecule, thus saving the different zones defined and those that are activated. This option also saves the orientation matrix, scale, slab, zval, and intensity map. Since different orientation matrices cannot be active simultaneously, the first molecule loaded when running Turbo-Frodo is important: Only its matrix will be applied.

SETLINEWIDTH
The SETLINEWIDTH option allows you to change the width of displayed lines. Note that antialiasing is supported only with a width of 1.

SLAB-DISPLAY
The SLAB-DISPLAY option displays the dimension of the clipping world.

SOLVENT
The SOLVENT option defines a residue with a residue type so that it will not be taken in connectivity calculation (CONNECT).

SPLINE

The SPLINE option reads a Spline Connolly Surface File. Surface points can be to the closest neighboring atoms attributed. This option can be therefore rather time consuming. The output file name is **file.ts** if the input file name is **file.spl**. For example:

spline
> enter the name of the input file (.spl): **toxin.spl**
> Do you want to link surface to atom (1) or simply calculate (2):**1**
> number of points is 984

In this case the output file name is **toxin.ts**. The file can be used as input with the LOAD SURFACE option. To create the **.spl file** you have to use the 2 following programs, **prep_spl** and **make_spl**, which are located under the **$TURBO_DIR/ms** subdirectory. The input of **prep_spl** is a Connolly Surface File (.srf).

SYMMETRY

The SYMMETRY option prompts you to give the molecule space group and cell parameters. Most of the protein space groups have been defined. You need to enter only the name of the space group, not all symmetry operations relative to this space group.

SYSTEM

The SYSTEM option returns to UNIX and executes UNIX orders. To leave this option, enter **lo**. A more efficient way to return to UNIX is to clone the window (that is, open another window).

TOLERANCE

The TOLERANCE option sets the acceptable violation of NMR distance restraints. The default value is set at 0.2 Å. When displaying the NMR restraints using the NMR ON/OFF display option (within the NMR submenu), a distance larger than the restraints but smaller than the restraint plus the tolerance will be displayed in orange, while the distance larger than the restraint plus the tolerance will be displayed in red.

TRAJECTORY

The TRAJECTORY option prompts you to give the name of the trajectory file extracted from the XPLOR dynamic calculation and then permits you to choose any of the conformations displayed. This option works provided the corresponding molecule has been loaded. To calculate conformations with XPLOR you should insert these commands into your XPLOR input file:

ascii =true   ! ASCII file (hexadecimal format by default)
trajectory =traj.hex ! name of the output trajectory file
nsavc = 1000   ! number of steps between conformations

Ex : traj
> enter the name of the file: **traj.hex**
File traj.hex Conformation Number: **90**
> select conformation numbers lower than: **91**
> RANGE
RANGE = ZONE, ZONE, ZONE (10 times maximum)
RANGE = all (all the conformations on the file)
ZONE = NCONF1 NCONF2 STEP (zone from conformations
        NCONF1 to NCONF2) by step of STEP
ZONE = NCONF1 NCONF2 (zone from conformation NCONF1 to NCONF2)
ZONE =NCONF (zone with only conformation NCONF)

Ex : 1 90 10
15 25
all
1 90 10, 15 25, 30,35 45

TRANSLATE
The TRANSLATE option applies a translation (i.e., a lateral shift) to the selected molecule. This option works provided the corresponding molecule has been loaded. This option applies to the memory storage, and you must save the molecule if you want to keep the changes. For example:

trans
> You want to translate the molecule MOL
> Translation X Y Z: 5.5 10 0

VDW-DENSITY
The VDW-DENSITY option chooses Van der Waals surface points with 20, 40 or 80 dots per square angstroms (20 points by default).

WORK-MAP
Use this option to specify the map on which you would like to work. Several options, such as RSR, TORSION DENS SEARCH, DGNL, or MAKE FRAGS, must be told to which map they have to refer. Note that if there is only one map displayed, the options will work on it, and DGNL and MAKE FRAGS may work without map at all.

ZONE
The ZONE option displays a large number of residue zones. These zones can be given different colors. (See the ZONE Option in the Color menu). For example:

zone RANGE
RANGE = ZONE, ZONE, ZONE (20 times maximum)
RANGE = all (all the residues in the selected molecule)
ZONE = NRSE1 NRES2 (zone from residues NRES1 to NRES 2)

ZONE = NRES(zone containing only residue NRES)
NRES=residue name (including chain label if exists)

Ex : 15 25
10
all
15 25, 30, 35 45
121 a 133 a (a=chain label)

You can modify a previously defined zone by using the following syntax:
Ex: zone=NZ RANGE. (Note that in this specific case, you have to type the equal sign.) NZ is the number of the zone that you want to modify. To add a zone in the list, give a new NZ number. Be careful with this step because if you do not accurately use the above-mentioned syntax, previously defined zones will be removed. Use the display Zone menu (located in the Molecule menu) to manage the display of the different zones.

P a r t
III

Display Options

The display options are accessible via the mouse when the Textport command **>go** is used. They are presented in eleven pop-up menus, which in some cases have been divided into submenus. The pop-up menus are activated when the cursor is driven up-screen to the top of the pop-up menu header. The submenus or functions can be accessed by dragging the mouse down, keeping the right button depressed. This will display the various options in yellow. To choose a function, drive the mouse along the menu until the function required is highlighted in yellow. Releasing the right button on the mouse will activate the function, which is then highlighted in green.

To turn off a function, drag the mouse to it and click the right button again. This action will make the function show up in yellow again and deactivate it.

The following eleven menus appear on top of the screen, working from left to right: Help, Main, Color, Geometry, Editing, Modeling, View, Plot, Tools, Surfaces, Molecules.

Philosophy of Use: **The display options apply to the most recently picked molecule.**

Display Options Table

The Help Menu:
Devices for Assistance

HELP MODE, HELP MOUSE, HELP DIALS.

HELP MODE

Three options are available with which on-line information can be obtained. To get this information, choose the HELP MODE option and then click on any other option. Instead of activating this option, this will give information about it. To switch back to the normal mode, choose the HELP MODE option again.

HELP MOUSE

The program makes use of the three mouse buttons, as well as combinations of them, which offers various alternatives. Active buttons are indicated in bold-faced letters.

L **M** R: Press to pick an atom or a bones point.
L M **R**: Press to activate a pull-down menu option.
L M R: Press for a new contour.
L **M R**: Press to quit display to CHAT.
L M R: Press in atom addition mode.
L M **R**: Press to move to the next residue.

HELP DIALS

The program can work without dials. In this case, activate the PSEUDO-DIALS option. The use of dials is summarized in the following scheme:

SCALE	Z PLAN
ROT Z	SLAB
ROT Y	INTENS MAP
ROT X	DIST STEREO

CHAPTER 6

The Main Menu:
Saving and Special Views

CONTOUR, SAVE, SAVE ALL, CLEAR LABELS, NO INFO, CLEAR INFO, ORTHO VIEW, SPLIT STEREO, ALT STEREO, POLAR STEREO.

This menu provides options that are used to manage the items displayed and to save any changes made in the Heap File.

CONTOUR
The CONTOUR option sets the new screen center at the most recently picked atom position. The maps and surfaces are recentered, as well as the molecular data in the sphere mode.

SAVE
The SAVE option transfers the modified molecular data after it has undergone delete, insert, replace, frag-trans, torsion, and move, from the memory storage to the source disk data base. The entry is written above. If several molecules are stored in the memory (and displayed on the screen), only the most recently picked molecule is saved. To save all the molecules use SAVE ALL.

SAVE ALL
The SAVE ALL option transfers the modified molecular data from the memory storage to the source disk data base. This is the same option as SAVE, but in this case all the molecules loaded are saved.

CLEAR LABELS
The CLEAR LABELS options switches off any previously picked atom labels.

NO INFO
The NO INFO option dispenses with the atom information in the upper-left corner of the screen.

CLEAR INFO

The CLEAR INFO option switches off the atom information in the upper-left corner of the screen.

ORTHO VIEW

The ORTHO VIEW option displays an orthogonal view of the current molecule, in the upper-right corner of the screen. It is useful for map fitting if no stereo device is available and for use with the ADD ATOM option.

SPLIT STEREO

The SPLIT STEREO option activates the split stereo mode. You can set the distance between the 2 views with the stereo dial (or pseudo-dial).

ALT STEREO

The ALT STEREO option renders operation in the alternate stereo mode possible. This mode requires special hardware (a stereo graphics board and special glasses). You can set the angle between the 2 views with the stereo dial (or pseudo-dial).

POLAR STEREO

The POLAR STEREO option renders operation in the polarized stereo mode possible. This mode requires very special hardware (a stereo graphics screen). The orientation between the two views can be set with the stereo dial (or pseudo-dial).

The Color Menu:
Coloring the Molecule

SAME, ATOM TYPE, RESIDUE TYPE, SIDE CHAIN, FEATURE**, DEFINE**.

The Color menu and submenus are described on the following page. The Color menu comprises 4 options ,SAME, ATOM TYPE, RESIDUE TYPE, and SIDE-CHAIN, and 2 submenus, FEATURE and DEFINE. All these options can be applied to different molecular entities. In the special case of FEATURE, the choice of entities appears in response to calling one of the FEATURE submenu options. For the other options (SAME, ATOM TYPE, RESIDUE TYPE, SIDE-CHAIN, and the DEFINE submenu), the choice of entities appears immediately. The molecular entities are as follows:

ALL: All the molecules on the screen.
MOLEC: The most recently picked molecule.
ZONE: The most recently picked zone.
RANGE: The program will ask you to state the residue range (first and last residue).
RESIDUE: The last residue to be picked.
ATOM: The last atom to be picked.
DOMAIN: The most recently picked domain.
NOTHING: No changes. This entity is useful to exit the menu if it was mistakently entered.

SAME
With the SAME option the molecular entities are displayed with a single color. You can choose the entity to be colored among those on the list (see above). You can choose one color out of 12 appearing in the menu.

ATOM TYPE
The ATOM TYPE option displays atoms in color based on a list defined with the DEFINE submenu in the Color menu. Default colors are Carbon Yellow, Oxygen Red, Nitrogen Blue, Sulfur Green, Hydrogen White, and all other atoms Orange. You can choose the entity to be colored from the list of entities (see above).

Color Menu Table

RESIDUE TYPE

The RESIDUE TYPE option colors residues depending on their type, as defined with the DEFINE submenu in the Color menu. Default colors are Acidic Red, Basic Blue, Aromatic Magenta, Cystein and Cystine Dark Green, Aliphatic Yellow, and Polar Light Green. You can choose the entity to be colored from the list of entities (see above).

SIDE CHAIN

The SIDE CHAIN option colors residue side chains based on a residue color list defined with the DEFINE submenu in the COLOR menu (the backbone color remains unchanged). Default colors are Acidic Red, Basic Blue, Aromatic Magenta, Cystein and Cystine Dark Green, Aliphatic Yellow, and Polar Light Green. You can choose the entity to be colored from the list of entities (see above).

FEATURE (submenu)
B-FACTORS, SECOND STRUCT, ACCESSIBILITY, HYDROPATHY.

B-FACTORS

The B-FACTORS option colors molecular entities depending on their crystallographic temperature factor. The B-factor range and color range can be defined using the B-FACT RANGE option from the DEFINE submenu. The default range is 0-20 but can be also modified with the Textport option BFACT-DEF. You can choose the entity to be colored from the list of entities (see above).

SECOND STRUCT

The SECOND STRUCT option colors secondary structures differently. The secondary structure distribution can be calculated in MENU TOOLS, with the SECOND STRUCT submenu either using the FAST option (fast and dirty) or DSSP, as described by C. Sander and W. Kabsch. The colors option of Helices, Beta-Sheets, Turns, and non-defined structures can be defined in the DEFINE option. When several molecules are displayed, this option works only on the last molecule to be picked (or by default, on the first molecule to be loaded).

ACCESSIBILITY

The ACCESSIBILITY option colors the last-picked molecule based on the residue-accessible surface calculated by either the Textport option ACCESS or the ACCESS CALC option located in the SECOND STRUCT submenu of the Tools menu. Calculations are performed according to C. Sander and W. Kabsch. The accessibility range and color ramp can be defined using the ACCESS RANGE option from the DEFINE submenu.

HYDROPATHY

The HYDROPATHY option colors the last-picked molecule depending on whether it is hydrophobic. For this purpose, calculations are carried out on a zone of the residue with the hydrophobic index given by Kyte and

Doolitle. The size of the computation window can be modified using the HYDROPATHY Textport option. The color ramp can be modified by using the HYDR RANGE option from the DEFINE submenu.

DEFINE (submenu)
ATOM TYPE DEF, RES TYPE DEF, SECOND STRUC, B-FACT RANGE, HYDR RANGE, ACCESS RANGE, DIST COLOR, SAVE DEFINE.

Note that after each definition, color must be applied again to render the new definitions effective.

ATOM TYPE DEF
The ATOM TYPE DEF option defines the atom colors to be used in the ATOMTYPE display and atom radii to be used in CONNECT MOL and CPK display. New atom types can be added with NEW. The default colors can be recovered with DEF.END exits.

RES TYPE DEF
The RES TYPE DEF option defines the residue colors for the RESIDUE TYPE display. New residue types can be added in NEW. Default colors can be restored with DEF.END exits.

SECOND STRUC
The SECOND STRUC option defines the colors in which the structure elements should be shown in the SECOND STRUC display: COIL, COLOR, HELIX COLOR, SHEET COLOR.END exits.

B-FACT RANGE
The B-FACT RANGE option defines the color ramp and the B-factor range.

HYDR RANGE
The HYDR RANGE option defines the color ramp and the hydrophobicity range.

ACCESS RANGE
The ACCESS RANGE option defines the color ramp and the accessibility range.

DIST COLOR
The DIST COLOR option defines the color to be used with the DISTANCE dashed lines.

SAVE DEFINE
The SAVE DEFINE option saves in the HEAP all the modifications set up with DEFINE.

The Geometry Menu: Verifying the Geometry

DISTANCE, CLEAR DIST, ANGLE, CLEAR ANGLE, XYZ, DIHEDRAL, CLEAR ALL, RAMACHANDRAN, NEIGHBOR, H-BONDS, OCCUPANCY.

The options that come with this menu can be used to calculate and display the geometric characteristics of the molecule either on the left part of the screen and/or on the molecule.

DISTANCE
With the DISTANCE option, you can measure the distance between 2 atoms. When the option is activated, the cursor arrow changes to **<1**, meaning that the first atom should be picked. When the first atom has been picked, **<1** is replaced by **<2**, meaning that the second atom should be picked. The option is active until it is switched off. The atoms in each pair are linked by a dashed line, and the distance between the 2 atoms is displayed.

CLEAR DIST
The CLEAR DIST option serves to remove the previously calculated distances and hydrogen bonds.

ANGLE
With the ANGLE option, you can measure the angles between three atoms. When this option is activated, the cursor arrow changes successively to **<1**, **<2**, **<3**, meaning that the first, second, and third atoms should be picked. The option is active until it is selected again, which cancels it. The atoms in the triangle are linked by dashed lines, and the angle value appears between pairs of atoms, opposite to the top angle.

CLEAR ANGLE
The CLEAR ANGLE option serves to remove the previously calculated angles.

XYZ
With the XYZ option, it is possible to display the Cartesian coordinates of the last-picked atom. Pick the option again to cancel it.

DIHEDRAL

The DIHEDRAL option calculates and displays a dihedral angle between 4 atoms. When this option is activated, the cursor arrow changes successively to **<1**, **<2**, **<3**, and **<4**, meaning that the first, second, third, and fourth atoms should be picked. The angle value is displayed in the upper-left information area. Pick the option again to cancel it.

CLEAR ALL

The CLEAR ALL option switches off any previously calculated distances and angles.

RAMACHANDRAN

The RAMACHANDRAN option calculates the phi, psi, and omega conformations and angles and displays the Ramachandran plot corresponding to the last-picked residue. This option does not work with the first and last residues. A general Ramachandran plot is available with the RAM ALL option in the Plot menu. Pick the option again to cancel.

NEIGHBOR

With the NEIGHBOR option it is possible to display all the contacts existing between the last picked atom and its neighbors. An atom is defined as being a neighbor if its distance from the picked atom is lower than the neighbor cutoff. The default cutoff is set at 4 Å but can be modified with the Textport order RAD NEIGHBOR. The contacts are displayed with dashed lines, and the distances are given between 2 atoms. Pick the option again to cancel it.

H-BONDS

The H-BONDS option displays hydrogen bonds as dashed lines on the atoms displayed. Three types of atoms are defined: acceptor (N), donor (O), and acceptor/donor (OH,OG,OW). The distance range is between 2.6 Å and 3.5 Å. Hydrogen bonds are not possible within one residue or between two consecutive residues. This option runs only on molecules containing no hydrogen atoms.

OCCUPANCY

With the OCCUPANCY option, you can set the occupancy of the last picked atom to a new value. Those values are kept in memory, and you should activate SAVE to transfer them to the disk data base.

The Editing Menu:
Editing the Structure

DELETE, INSERT, REPLACE, SHEET/HELIX, ADD ATOM, [WAT]/CA, DEL ATOM, OPEN MODEL**, DATA BANK**, CONNECT MOL.

The Editing menu options can be used to modify the molecule by inserting, deleting, or replacing information. OPEN MODEL** and DATA BANK** mainly serve to construct small molecules, although they can be used on any protein structure.

DELETE
With the DELETE option, it is possible to delete the last-picked residue in the molecular structure. This option works on the memory storage, and you must save the molecule if you want to keep the changes.

INSERT
The INSERT option inserts a residue into a molecular structure. The new residue is inserted either before or after the last-picked residue. The residue must be chosen first on the menu. A Textport appears and a prompt asks for the name of the residue and its location, after or before the picked residue. The type of secondary structure can be chosen by using the SHEET/HELIX option on the same menu. SHEET works by default. The position of the side chain is given by the dictionary. This option works on the memory storage, and you must save the molecule if you want to keep the changes.

REPLACE
The REPLACE option replaces the last-picked residue in the molecular structure. The residue to be replaced must first be chosen on the menu. A Textport appears and a prompt asks for the position of the side chain: the previous one or the one defined by a dictionary. This option works on the memory storage, and you must save the molecule if you want to keep the changes.

SHEET/HELIX
The SHEET/HELIX option specifies the secondary structure of the inserted residue. The residues are inserted in the sheet or helical conformation whether this option is activated or not.

Editing Menu Table

ADD ATOM

The ADD ATOM option adds a new atom which will be treated as a new residue, in the required position. This option is best accomplished with the orthogonal view displayed. With the left button on the mouse, pick the atom position required on the main view (x and y screen coordinates). A line will appear in the orthogonal view. With the left button on the mouse, pick the point on the line where you want to place the atom. Confirm the entry by picking **yes** or **no**. If you pick **yes**, a Textport will appear, and a prompt will ask you for the name of the residue. This option can work in a permanent way. If you choose so, the name of further residues will be set automatically by increasing th first name given. This option also works on the memory storage, and you must save the molecule if you want to keep the changes.

[WAT]/CA

The [WAT]/ CA option serves to choose the type of atom you want to add, using the ADD ATOM option in the same menu. A water molecule or an ALA residue (atom type CA) will be added whether this option is active or not.

DEL ATOM

THE DEL ATOM option deletes an atom in a molecular structure. Be careful with this option and keep in mind a possible dictionary/structure mismatch.

OPEN MODEL (submenu)

OPEN MODEL, CONSTRUCT**, USER DICT, COMMON, SUGARS, AM-ACIDS.

OPEN MODEL

With the OPEN MODEL option, it is possible to obtain access to all the necessary tools for designing a new molecule. The construction takes place step by step by adding successive fragments. A dictionary is available containing some of the basic fragments you may need. To connect a fragment to your molecule, you should select a hydrogen atom on your molecule and a hydrogen atom on the new fragment (or vice versa). You can have your own dictionary (user dictionary), in which you can keep specific molecules (such as your own ones and any small ones which may be useful). When you activate this option for the first time within Turbo, the name of the user dictionary is requested.

CONSTRUCT (sub-submenu)

ADD FRAG, ADD ATOM H, REPLACE ATOM, DEFINE RESIDUE, DEFINE ANGLE, DEFINE DIST, DEFINE ATOM.

ADD FRAG

The ADD FRAG option serves to add a fragment to your molecule. First select a hydrogen atom on your molecule and a hydrogen atom on the

appropriate fragment (or vice versa) to connect the new fragment to your molecule. This option works on the memory storage, and you should save the molecule if you want to keep the changes.

ADD ATOM H

The ADD ATOM H option adds a hydrogen atom to any specific atom in your molecule. Select the atom to which you want to add a hydrogen atom (e.g., a carbon atom) before selecting the option. This option works on the memory storage, and you should save the molecule if you want to keep the changes.

REPLACE ATOM

The REPLACE ATOM option transforms your atom into one of another type. Select first the appropriate atom, before selecting the option. Then choose the new type of atom in keeping with the valence. This option works on the memory storage, and you should save the molecule if you want to keep the changes.

DEFINE RESIDUE

The DEFINE RESIDUE option serves to reorganize your molecule. First, select a set of atoms, using the BREAK BONDS option, then select this set by picking an atom, before selecting the option. Since the new name of the residue will be automatically calculated, the new type of residue is requested. This option works on the memory storage, and you should save the molecule if you want to keep the changes.

DEFINE ANGLE

The DEFINE ANGLE option reshapes an angle. After selecting the option, determine the angle you want to reshape by picking the three atoms. The third atom and all the atoms connected to it will then be shifted to design the new angle. This option works on the memory storage, and you should save the molecule if you want to keep the changes.

DEFINE DIST

The DEFINE DIST option reshapes a distance. After having selected the option, determine the distance you want to reshape by picking the two atoms involved. The second atom and all the atoms connected will then be shifted. This option works on the memory storage, and you should save the molecule if you want to keep the changes.

DEFINE ATOM

The DEFINE ATOM option renames the atom that you have picked.This option works on the memory storage, and you should save the molecule if you want to keep the changes.

DATA BANK (submenu)
SEARCH NAME, SEARCH CHEMIC, USER DICT.

SEARCH NAME

The SEARCH NAME option provides access to the main dictionary. You have to give a string of letters to obtain all the molecules whose name contains these letters.

SEARCH CHEMIC

The SEARCH CHEMIC option grants access to the main directory. Once you have entered the chemical formula (or part of it), you will obtain all the molecules containing these atoms.

USER DICT

The USER DICT option will help you manage your own dictionary, i.e, LIST the contents, CREATE a dictionary, DELETE a dictionary, INSERT a molecule, and DELETE a molecule.

CONNECT-MOL

With the CONNECT-MOL option, the connectivity of the molecule can be recalculated. This step can be performed after FRAG ROT/TRANS, MOVE ATOM, REFI, and BREAK BOND, if necessary.

The Modeling Menu: Designing the Molecule

BOND, BREAK BOND, FRAG ROT/TRANS, MOVE ATOM, REFINE, DGNL, DGNL PICK, TORSION**, RSR-WAT, RSR FULL.

The options of this menu can be used to modify and correct the geometry of the selected molecule.

BOND
The BOND option establishes a bond between two successively picked atoms. This option also works with bones points, but not with a mixture of bones and atoms. When this option has been activated, the cursor arrow changes successively into **<1** and **<2**, which means that the first and second atoms should be picked.

BREAK BOND
The BREAK BOND option removes a bond between two successively picked atoms. This option works also with bones point, but not with a mixture of bones and atoms. When the option is activated, the cursor arrow changes to **<1** and **<2**, which means that the first and second atoms have to be picked.

FRAG ROT/TRANS
The FRAG ROT/TRANS option can rotate and translate a molecular fragment up to 10000 atoms in size. The fragment is defined as a set of connected atoms in the memory. If applied to Carbons Alpha, the rest of the atoms will follow. Therefore, to define a fragment, use the BOND and BREAK BOND options. The changes can be made using the dial (or pseudo-dial). The new position can be confirmed in memory or not by picking **yes** or **no**. The SAVE option in the Main menu should be used to save the new coordinates on the disk, and if necessary, the CONNECT-MOL option to reconnect your fragment.

MOVE ATOM
The MOVE ATOM option serves to move a single atom. The new position can be confirmed or not by picking **yes** or **no**. The SAVE option in the Main menu should be used to save the coordinates on the disk, and if necessary the CONNECT-MOL option to reconnect your atom.

REFINE

The REFINE option geometrically refines a stretch of residues. It uses the former Frodo routine REFI, but with a different user interface. The residues have to be described in the **$TURBO_DIR/dict/dict.pdb** file. You can select the stretch of residues by picking the first and last residues. The first residue cannot be a proline. This triggers REFINE, and each step is displayed on screen. The new position can be confirmed in memory by picking **yes** or refused by picking **no**, or you can start performing new operations on the same stretch by picking CONTINUE. You must use the SAVE option in the MAIN menu to save the coordinates on the disk, and CONNECT-MOL to reconnect your fragment if required.

DGNL

The DGNL option extracts from the data base and display a 3-D molecular fragment that is homologous to a selected fragment of a loaded molecule. This option is based on diagonal distance matrixes and on the DGNL concept by Alwyn Jones, but it has been rewritten. You first have to define the fragment by picking the first and last residues (maximum 20 residues in the fragment). After the calculation, the information fields of the 10 best fragments (lowest in deviation) are displayed on the left side of the screen. These fields contain the name of the source protein, the sequence of the fragment, and the RMS deviation. You can display a fragment by picking its information field. If a map is attached to the molecule, the way in which the fragment fits in the density is displayed on a graph in the lower-left corner. A fragment can be stored in memory by picking **yes**. Picking **no** brings the option to an end. You must use the SAVE option in the Main menu to save the coordinates on disk.

The DGNL option can also be used on bones atoms as long as there is only one molecule entry on the screen. First "dummy" poly-ala entry needs to be prepared by reading the following file, **$TURBO_DIR/dict/poly-ala.wh**. You can then use DGNL in the same way, using REPLACE to obtain the requisite sequence.

DGNL PICK

The DGNL PICK option is the same as the DGNL option except that it can perform an insertion of some residues. Pick all the consecutive residues that define a first part of a fragment, then specify the number of residues to insert, then pick the consecutive residues that define the last part of the fragment. You must use the SAVE option to save the coordinates in the Heap file.

TORSION (submenu)
TORSION-PICK, TORSION-AUTO, DENS-SEARCH, DIST-SEARCH.

TORSION-PICK

The TORSION-PICK option can modify up to 8 torsion angles defined between two selected atoms. When activating this option, the cursor

becomes **<1** and **<2**, prompting for the selection of the first and last atoms defining the torsion angle(s). The angles are then numbered, and each angle corresponds to one pseudo-dial on the lower-left part of the screen. The new position can be stored or not in memory by picking **yes** or **no**. Use the SAVE option in the MAIN menu to save coordinates on disk.

TORSION-AUTO

The TORSION-AUTO option modifies any of the torsion angles of the last residue picked. The torsion angles are automatically defined on both the main and side chains. The angles are numbered, and each angle corresponds to one pseudo-dial on the lower-left corner of the screen. The new position can be stored or not in memory by picking **yes** or **no**. Use the SAVE option in the MAIN menu to save the new coordinates on screen.

DENS-SEARCH

The DENS-SEARCH option automatically fits amino acids depending on their electron density map. This option runs on the last-picked residue. You can choose from among the twenty best positions. The position chosen can be accepted or refused by picking **yes** or **no**. Use the SAVE option within the MAIN menu to save the coordinates on disk.

DIST-SEARCH

The DIST-SEARCH option automatically fits amino acids depending on their atomic environment. This option runs on the last residue to be picked. You can choose between the twenty best positions. The position chosen can be stored in memory or not by picking **yes** or **no**. Use the SAVE option within the MAIN menu to save the coordinates on disk.

RSR-WAT

The RSR (Real Space Refinement) -WAT option fits an atom into a density map by seeking the nearest local maximum. This option is particularly useful for fitting water molecules. If several density maps are displayed, you will be asked for the one in which you would like to fit the atom.

RSR-FULL

The RSR-FULL option fits a set of atoms (a maximum of 100) into the density by seeking the nearest local maximum. If several density maps are displayed, you will be asked for the one in which you would like to fit the set of atoms. If only one map is displayed, the fitting will be processed on it.

49

The View Menu: Representing the Structure

SPHERE, CA, BACKBONE, CA-SIDE-RES, CA-SIDE-NAME, NO HYDROGEN, SMALL CIRCLE, LABEL ATOM**, COMPACT SELE**, COMPACT-DO-IT, MAP SPHERE.

This menu contains both options and submenus. These options and those contained in the submenus are used to set the type of representation of the selected molecule and the atom labeling.

SPHERE
The SPHERE option displays the molecular contents of a sphere, the size of which is defined by the radius. This radius can be set with the RADIUS SPHERE Textport option.

CA
The CA option selects and links the Alpha Carbon atoms of the last molecule to be picked. The maximum possible distance at which two consecutive atoms can be linked is 6.3 Å.

BACKBONE
The BACKBONE option selects the backbone atoms (main chain) of the last molecule to be picked.

CA-SIDE-RES
The CA-SIDE-RES option selects and displays amino acid side chains when the CA or BACKBONE option has been activated. When one of these options has been activated, a list of amino acids shows up in the pull-down menu. You can then pick the amino acids required and conclude with END.

CA-SIDE-NAME
The CA-SIDE-NAME option selects and displays amino acid side chains in terms of their position in a sequence when either the CA or BACKBONE option has been activated. When one of these options has been activated, a Textport appears from which it is possible to select the residue names. You must end the list by a period (.).

View Menu Table

NO HYDROGEN
The NO HYDROGEN option removes the hydrogen atoms display from molecules. This option does not remove the hydrogens from memory, but you can do that with the DELETE ATOM Textport option.

SMALL CIRCLE
The SMALL CIRCLE option presents atoms in the forms of filled depth-cued circles, the radii of which are 1/10 of the corresponding atom. This mode is a "pseudo balls and sticks" display, which makes interactive handling possible.

LABEL ATOM (submenu)
LABEL TYPE**, DEF FONT, COLOR LABEL

LABEL TYPE (sub-submenu)
The LABEL TYPE option allows one to choose among various modes of displaying atom labels. These modes are as follows :

T1 NAM ATO (the default mode)
T1 NAM
T3 NAM ATO
T3 NAM
MOL T1 ATO
T1 NAM STRUC
MOL NAM ATO
NAM ATO
DOM NAM ATO

T1 / T3 mean one-letter code/ 3 codes giving the residue type. NAM is the name of the residue, ATO is the name of the atom, MOL is the name of the molecular entry, and DOM is the number of the domain to which the residue belongs, if it exists.

DEF FONT
The DEF FONT option modifies the label font. In the Textport that appears, you can choose the size and font names. This option is useful to set pictures or photos of screens.

COLOR LABEL
The COLOR LABEL option can be used to change the color of the labels.

COMPACT SELE (submenu)
CPK, BALL & STICKS, COMPACT ROT, CURVE LINE, CURVE PATCH, STRUC REP, STRUC MENU, RESET STR MENU.

This submenu contains options used when displaying a compact view of the loaded molecule.

CPK

The CPK option displays the last-picked molecule in the form of lighted compact spheres. From the menu which appears, you can choose the complexity (number of faces) and the size of the spheres. The radii of atoms (and so spheres) can be modified with the ATOM TYPE DEF option of the DEFINE submenu, which is located in the Color menu.

BALL & STICKS

This option displays the last-picked molecule as lighted "balls and sticks". In the appearing menu you can choose the type (B & S White, B & S Colored, STICKS ONLY) and the size of the "sticks". The radii of atoms (and so balls) can be modified with the ATOM TYPE DEF option of the DEFINE submenu found in the Color menu.

COMPACT ROT

The COMPACT ROT option serves to work interactively in the compact display mode. When this option has been activated, you can use the dials (or pseudo-dials) to move the molecule around.

CURVE LINE

The CURVE LINE option displays a curved line calculated from the Alpha Carbons positions.

CURVE PATCH

The CURVE PATCH option displays a curved surface calculated from the Alpha Carbons positions.

STRUC REP

With the STRUC REP option, secondary structures can be displayed in the vector mode in the form of cylinders, arrows, and ribbons. The COMPACT-DO-IT option will then transform this view into compact display.

STRUC MENU

The STRUC MENU option sets the compact display features of secondary stuctures (STRUC REP).

SHEET TYP 0: helices as coiled structures.
HEL TYPE 1: helices as spirals.
HEL TYPE 2 : helices as cylinders.
SHEET TYP 0: strands as coiled structures.
SHEET TYP 1: strands as arrows.
SHEET TYP 2: strands as ellipses.
DIM COIL: sets the size of COIL display.
DIM HEL: sets the size of HELIX display.
DIM SHEET: sets the size of STRAND display.
SAVE: saves these settings in the HEAP.

RESET STR MENU

This option recovers the default values of the STRUC MENU features.

COMPACT-DO-IT

The COMPACT-DO-IT option calls up a compact view based on the previously selected modes.

MAP SPHERE

The MAP SPHERE option displays the map only around the displayed atoms. Use the Textport order RADIUS MAP to define the radius you want.

The Plot Menu:
Graphs and Postscripts

WRITE MATRIX, GRAPHICS**, RAMA ALL, ALL MODE, SEQUENCE, SET VIEW**.

This menu's options enable the user to obtain hard copies of the displayed molecule(s) and to display some of the information built up by residue.

Most of these options create Postscripts files (GRAPHICS' options, PLOT, RAMA ALL). For the PLOT option, keep in mind that the thickness of the lines can be adjusted by making color selections using the SAME color map. The uppermost line is the thickest (red very thin, white very thick). These ASCII Postscript files can be named and fed into a Postscript laser printer. To send a file on UNIX workstation to a laser printer, follow this procedure:

cat trypsin.pst>/dev/ttydx (where x is the port number)
or: lp trypsin.pst (if your laser printer is defined as default printer)

To send a file on a VMS workstation to a laser printer:
copy trypsin.pst txa5: (If the printer is on port 5)

These files can be edited and adapted.

WRITE MATRIX
The WRITE MATRIX option writes the current orientation matrix on disk.

GRAPHICS (submenu)
GRAPH B-F ALL, GRAPH B-F M/S, DIFF B-F ALL, DIFF B-F M/S, GRAPH.

GRAPH B-F ALL
The GRAPH B-F ALL option displays the B-factors per residue (all the atoms except hydrogens) of a single molecule on a graph. A Postscript output file can be generated.

Plot Menu Table

GRAPH B-F M/S

The GRAPH B-F M/S option displays the B-factors per residue (main and side chain except hydrogens) of a single molecule on a graph. A Postscript output file can be generated.

DIFF B-F ALL

The DIFF B-F ALL option displays the difference of B-factors per residue (all the atoms except hydrogens) of the last two molecules picked.

DIFF B-F M/S

The DIFF B-F M/S option displays the difference of B-factors per residue (main and side chain except hydrogens) between the last two molecules picked. A Postscript output file can be generated.

GRAPH

The GRAPH option displays the RMS deviation per residue on a graph after superimposition performed by the RIGID2 option (Tools menu) and by the Textport options COMPARE or DEVIATION. While running RIGID2, you will be asked for a deviation file name. This file must be an input to the Textport option GR-RMS-DEV. For the other options, GRAPH displays automatically the result. Two graphs can be displayed simultaneously. The option works with the GR-RMS-DEV Textport option.

RAMA ALL

The RAMA ALL option displays the full Ramachandran plot of the last molecule to be picked. You can generate a Postscript file with this plot. This file can be named and sent to a Postscript laser printer.

ALL MODE

The ALL MODE option gives access to some options for displaying all the molecules that have been loaded. The following display options are available in the ALL MODE: all the options in the Color menu working on whole molecules, CA, BACKBONE, CA-SIDE-RES, CA-SIDE-NAM, NO-HYDROGEN, NO-PSEUDO, PSEUDO-HYDS, RAMA ALL.

SEQUENCE

The SEQUENCE option displays the sequence of the last-picked molecule. You may go backward or forward with the different commands or use the lift to display the entire sequence.

SET VIEW (submenu)

PLOT, PLOT FRAME, TITLE, PHOTO, SCREENSAVE, SCREENSAVE ALL, WHITE BACK

PLOT

The PLOT option writes the model displayed on a Postscript file. The Textport appears and prompts you to give a file name.

Warning: **The size of the SPLIT STEREO view is wrong and does not fit the PLOT FRAME.**

PLOT FRAME
The PLOT FRAME option displays the plotting frame (size of the paper). You can modify the position of the frame by selecting its lower-left corner with the mouse's left button and by moving it interactively.

Warning: The size of the SPLIT STEREO view is wrong and does not fit the PLOT FRAME; it should be set to 2-3 centimeters on screen.

TITLE
The TITLE option displays up to ten titles on the screen. When defining a title, one is asked to give the title itself, the type and size of the font, whether or not the title should be displayed with an arrow, and the width of the arrow. The title and the arrow can be moved separately by dragging on the little green box at the end of the arrow. Please note that some fonts may not be available in the required size. This depends on the fonts provided by your computer. The TITLE option is useful to set pictures or photos of screens.

PHOTO
The PHOTO option removes the boxes from the titles and arrows so that a photo of a screen can be taken.

SCREENSAVE
The SCREENSAVE option captures an image (**rgb** format file) of Turbo. Give a name to your **rgb** file and then prepare the view to capture the image (CPK views, SECOND STRUCT, or whatever you would like). When ready, drag the mouse while keeping the left button depressed to define the area that you want.

SCREENSAVE ALL
The SCREENSAVE ALL option works the same way as SCREENSAVE but captures a view of the entire screen. When ready just activate the left button.

WHITE BACK
The WHITE BACK option lets you work on a white background instead of a black one.

The Tools Menu:
The Workshop

RIGID 2, SCAN, MAPS**, BUILDING**, BONES**, SYMMETRY**, NCS**, NMR**, TRAJECTORY**, SECOND STRUCT**.

This menu contains submenus, which in turn contain options for displaying the electronic density maps, the bone maps, the symmetric molecules and NMR features, and the dynamic trajectories.

RIGID 2
The RIGID 2 option refines the fit between two molecules that are already quite similar. This option works in the same way as RIGID (see this Textport option), except that a rough fit has to already exist.

SCAN
The SCAN option displays the different dynamic conformations one after another. You must first have loaded the conformations, using the TRAJECTORY Textport option. You can set the speed and the way you want to scan molecules (forward, backward, one way, or in a loop) by using the tape-recorder-like commands. Other features are available under the display TRAJECTORY submenu in the Tools menu.

MAPS (submenu)
MAPS

MAPS
With the MAPS option, you can display the maps' management menu. The 5 map entries with their 5 possible levels are listed. You can manage the display of the maps and levels by picking on DEF. In this case, Textport appears and asks you to give the new parameters (the levels, the color associated with each level, the box size, and the contour direction). You can change the MAP COLOR by picking COL in the box on the right.

Tools Menu Table

BUILDING (submenu)
MAKE FRAGS, HIDE FRAGS, ALIGN FRAGS, ADD TPPR, MULTI REPLACE, TOOLS BOX**, CA LABELS.

MAKE FRAGS
The MAKE FRAGS option connects the CA atoms which are less than 5 angstroms apart. This cutoff is the default and can be modified with the DIST CA Textport option. If present, the map will help the process of fragment building. Connections outside the density are not allowed. The density cutoff is chosen as the first displayed level. If several maps are displayed, choose one with the WORK MAP Textport option. Those fragments will be aligned further with the target sequence in the ALIGN FRAGS option.

HIDE FRAGS
The HIDE FRAGS option disables the display of the previously built fragments.

ALIGN FRAGS
The ALIGN FRAGS option performs the alignment of previously built CA fragments on the target sequence. These fragments must have been built with the MAKE FRAGS option, and the target sequence must have been loaded with Textport's LOAD SEQUENCE order. The 10 best alignments are proposed, and results are presented along the sequence with a color code: green for correct matchings, red for wrong matchings, and yellow if both residue types belong to the same group (short, medium, long, or aromatic). To accept an alignment, click CONFIRM, and all residues will be ordered and named after the target sequence. This options works on the memory storage, and you must save the molecules if you want to keep the changes. Note that you can work with either the right or middle mouse buttons. The left button will end the alignment screen.

ADD TPPR
The ADD TPPR option fits a Trigonal Planar Pseudo Residue in the electron density map. After activating the option, activate one of the two BUILD MODE buttons that appear on the right of the screen. Click on the density map where you would like to add the pseudo residue with the left mouse button. The pseudo residue comes at this place and is automatically fitted by a special algorithm of Real Space Refinement. If you choose the BM1 button, RSR will be performed with both rotation and translation. If you choose the BM2 button, RSR will be performed with only a rotation. For the first residue you add, you have to specify the name (suffix included), and the following residue names will be increments of this first one. TPPRs are added on the last molecule. This option Works on the memory storage, and you should save the molecule if you want to keep the changes.

MULTI REPLACE

The MULTI REPLACE option replaces a set of Trigonal Planar Pseudo Residues by real residues. The TPPRs must first have been aligned with the target sequence, so that the atoms are named after N, CA, C, and CB.The first step of the replace algorithm is to create the backbone oxygens. This step is achieved by rotating a virtual peptidic bond on the CA (i) CA (i +1) axis and searching for the best fit in the density map, for the N, C, O of this virtual bond. They are then replaced by residues according to their types. The side-chain orientation is given by the CB position of the former TPPR. One cycle of refine is automatically performed.

If there is one mismatch between the set of residue types and the target sequence, you will be asked to correct the TPPR types. This option is also suitable for Molecular Replacement. You have to change the residue types of your residues with the CA LABELS option, and the side chain will thus be changed. The MULTI REPLACE option works on the memory storage, and you should save the molecule if you want to keep the changes.

TOOLS BOX (sub-submenu)
RESI ROT/TRANS, RSR MAIN, RSR SIDE, ROTAMERS.

RESI ROT/TRANS

The RESI ROT/TRANS option works as the FRAG ROT/TRANS option does, but the whole last-picked residue is automatically selected. The modifications are made by using the pseudo-dials visible on the left side of the screen. This option works on the memory storage, and you should save the molecule if you want to keep it changed.

RSR MAIN

The RSR MAIN option fits the backbone (N, CA, C, O, CB) of the last-picked residue into the density by seeking for the nearest local maximum. The side chain follows the movement of the C Alpha atom. You can choose the map that you would like to fit in with the WORK-MAP Textport option. If only one map is displayed, the fittting will be performed on it.

RSR SIDE

The RSR SIDE option fits the side chain of the last-picked residue into the density by seeking for the nearest local maximum. The main chain remains unchanged. You can choose the map you would like to fit in with the WORK-MAP Textport option. If only one map is displayed, the fitting will be performed on it.

ROTAMERS

The ROTAMERS options proposes all different rotamers of an amino acid as they have been defined by J.W. Ponder and F.M. Richards. This option runs on the last-picked residue. The chosen position can be

accepted or refused in memory by picking YES or NO and can be modified by using the left pseudo-dials. The figure noted for each proposed rotamer is its frequency in the Ponder and Richards protein panel.

CA LABELS
In the label window, you may change the residue type of the last-picked residue by clickingf with the right mouse button on the appropriate place. Note that this type is important for the ALIGN FRAGS option. You can also add a comment to the residue, but the comment must not be longer than 60 characters. Comments may be listed or dumped in an ascii file via the LIST COMMENTS and MAKE COMMENTS Textport orders, respectively.

BONES (submenu)
BONES ON/OFF, 8 CELLS, BONES MAIN, BONES SIDE, UNDEFINED, BONES SAVE, BONES C ALPHA, CREATE CA, CREATE TPPR.

Bones Labels
This is not a menu or submenu but a new feature. A new label window now appears when a bones point is selected. In this label window you may change the name (or label), the bones type, or the residue type of the last-picked bones point by clicking with the right mouse button on the appropriate place. The bones type is one of the four following types: C ALPHA, MAIN, SIDE, UNDEFINED; choose one in the appearing pop-up menu. The residue type is one of the 20 amino acids or one generic type such as Short, Medium, Long, or Aromatic. You can give a name to the C ALPHA bones points. The name consists of a number and a letter, which can be considered as a fragment label. If adjacent and connected C ALPHA points exist, a name will be proposed for the last-picked point. If you would like to keep the changes, you should ask for the BONES SAVE option.

BONES ON/OFF
The BONES ON/OFF option toggles on and off the bones map display.

8 CELLS
The 8 CELLS option calculates and displays the bones map of 8 cells.

BONES MAIN
The BONES MAIN option displays only the bones identified as belonging to the main chain.

BONES SIDE
The BONES SIDE option displays only the bones identified as belonging to the side chains.

UNDEFINED

The UNDEFINED option displays any bones that have not been identified as belonging to either the main chain or the side chains.

BONES SAVE

The BONES SAVE option saves on disk the status and the connectivity of the bones map.

BONES C ALPHA

The BONES C ALPHA option displays only the BONES identified as C ALPHA points. In the label window, you may change this bones type as well as the residue type and label by clicking with the right mouse button on the appropriate place.

CREATE CA

The CREATE CA option creates residues (of only one CA atom) at the location of the bones points identified as C ALPHA. Define a fragment by picking the first and last bones point.

CREATE TPPR

The CREATE TPPR option creates Trigonal Planar Pseudo Residues at the location of the bones points identified as C ALPHA. They are automatically fitted by a special algorithm of Real Space Refinement. Define a fragment by picking the first and last bones point.

SYMMETRY (submenu)

SYM NEIGH, 1 CELL, 27 CELLS, 1 BOX, 27 BOXES, CELL CONTENT, DEL SYM, NO SYMM.

SYM NEIGH

The SYM NEIGH option displays all symmetry and translation-related molecules within radius R from the origin molecule. The R radius can be adjusted with Textport's RAD SYM order.

1 CELL

The 1 CELL option displays all the symmetric related molecule of the defined space group according to the definitions of the international tables. This option requires the CA view mode.

27 CELLS

The 27 CELLS option displays all the symmetric related molecules of the defined space group with all the translations along the cell axis. This option requires the CA view mode.

1 BOX

The 1 BOX option displays a representation of the cell.

27 BOXES

The 27 BOXES option displays 3*3*3 representations of the cell.

CELL CONTENT

The CELL CONTENT option displays all the symmetric related molecules of the defined space group, but relevant translations have been applied to place them into the cell. This options requires the CA view mode.

DEL SYM

The DEL SYM option cancels the display of the most recently picked symmetry-related molecule.

NO SYMM

The NO SYMM option switches off the symmetry display.

NCS (submenu)

APPLY MATRIX, AUTO APPLY, NCS MAPS, PRINT NCS.

APPLY MATRIX

With this option, you can display the Non-Crystallographic Symmetries. You have to define the symmetries with Textport's DOMAIN and NCS options.

AUTO APPLY

With this option, you can also display the Non-Crystallographic Symmetries, and you first have to define the symmetries with Textport's DOMAIN and NCS options. In this option's case, however, the display is automatic: Only symmetries related to the centered residue will be displayed.

NCS MAPS

The NCS MAPS option allows you to diplay the symmetric box of a map if Non-Crystallographic Symmetries are defined for the protein. Symmetries defined for the last picked molecule are applied if APPLY MATRIX is activated. Symmetries defined for the centered molecule are applied if AUTO APPLY is activated. You can define different levels and colors than those in the original map box. Furthermore, you can ask for map difference or average density. The map difference is represented by the absolute value of the difference, and, in addition, only positive values are considered. The average density represents the average between all symmetric boxes and the original one.

PRINT NCS

The PRINT NCS option allows you to display the rotation matrix and translation of the different Non-Crystallographic Symmetries defined, in Textport. Give a file name if you want your matrix to be printed in a file.

NMR (submenu)

NMR ON/OFF, VIOLATIONS, COL REST ATO, COL REST RES, DISPLAY RES, PSEUDO HYDS, NO PSEUDOS.

NMR ON/OFF

The NMR ON/OFF option toggles on and off the constraints display. The constraints are displayed as colored lines according to the following code:

Yellow: too short
Green: in the normal range of values
Orange: maximum value range<value<maximum value range +0.2 Å
Red: value>maximum value range + 0.2 Å

0.2 Å is the default tolerance. You can set it with Textport's TOLERANCE option.

The colors will change interactively when you activate the FRAG ROT TRANS and TORSION options in the MODELING menu.

VIOLATIONS

The VIOLATIONS option displays only the constraints which have been isolated (value> maximum value range +0.2) (Red lines). 0.2 Å is the default tolerance. You can set it with Textport's TOLERANCE option.

COL REST ATO

The COL REST ATO option colors the atoms depending on the constraints. Yellow means that there is no constraint with this atom; the colors range from blue (few constraints) to red (many constraints).

COL REST RES

The COL REST RES option colors the residues depending on the number of contraints. Yellow means that there are no contraints with the residue; the color ranges from blue (few contraints) to red (many constraints).

DISPLAY RES

The DISPLAY RES option displays only the contraints applying to the most recently picked residue.

PSEUDO-HYDS

The PSEUDO-HYDS option generates pseudo-hydrogens. These artificial atoms can be used instead of protons in cases where there is no stereospecific assignment. They can be created with protons previously generated using the nmr paramameter files of XPLOR.

NO PSEUDOS
The NO PSEUDOS option serves to alternatively switch on and off the pseudo-hydrogen display.

TRAJECTORY (submenu)
DEFINITION, STATIC, DYNAMIC.

DEFINITION
The DEFINITION option defines the dynamic conformations you want with the TRAJECTORY option. The program will prompt you to give the name of the trajectory files and the way you want to display the different conformations (see Textport's TRAJECTORY option for details).

STATIC
The STATIC option displays all chosen conformations. When picking an atom, the conformation number is given in the information pad.

DYNAMIC
The DYNAMIC option highlights the molecular dynamic trajectory. When this option is activated, the following command pad apppears in the lower part of the screen:

Where MODE is: step by step / one way /auto reverse / loop, and NUM is the number of the displayed conformation.

SECOND STRUCT (submenu)
HELIX, SHEET, TURN, COIL, FAST CALC, DSSP CALC, ACCESS CALC.

The SECOND STRUCT option displays on the screen only some secondary structure elements (e.g. Helices,Turns, and so on). The default color of these elements can be changed with the display option SECOND STRUCT of the DEFINE submenu (Color menu).

HELIX, SHEET, TURN, COIL
1/ Serves to select these secondary structures, once they have been defined by either FAST CALC or DSSP CALC for display.
2/ If the CHANGE option is selected, the secondary structure type of the most recently picked residue is set to the chosen secondary structure HELIX, SHEET, TURN, or COIL.

FAST CALC
The FAST CALC option calculates the secondary structures using the fast and dirty algorithm based on inter-CA distances. This option works on the memory storage, and you should activate the SAVE option to keep the result.

DSSP CALC

The DSSP CALC option calculates the secondary structures. The algorithm used here is from Kabsch and Sander's DSSP program, with minor modifications. This option works on the memory storage, and you should activate the SAVE option to keep the result.

ACCESS CALC

The ACCESS CALC option calculates the water-accessible surface by residue. The algorithm used is also from Kabsch and Sander's DSSP program, with minor modifications. This option works on the memory storage, and you should activate the SAVE option to keep the result.

The Surface Menu: The Landscape

SURF MENU, NO SURF, SAME COLOR, COLOR SURF, NEW VDW SURF, ADD VDW SURF, VDW MENU, COLOR VDW.

This menu contains options to calculate and display Van der Waals surfaces and submenus in which the different loaded surfaces, Van der Waals, Connolly, and Spline surfaces may be selected or switched on and off. Connolly surfaces must have been previously calculated (with programs supplied in the **$TURBO_DIR/ms** subdirectory) and loaded with the Textport LOAD SURFACE order. Van der Waals surfaces, on the other hand, are interactively calculated.

SURF MENU
The SURF MENU option displays the surface menu. You can switch each loaded molecular surface on/off. You can also select a surface with the menu.

NO SURF
The NO SURF option switches off the surface display.

SAME COLOR
The SAME COLOR option colors the selected surface with a single color. This makes the surface display more interactive. If the surface is already colored, however, and depending on the type of atom or residue, there will be a mismatch when you activate this option.

COLOR SURF
The COLOR SURF option switches the system over to the color surface mode. All the subsequent color options will be applied to the selected surface.

NEW VDW SURF
With the NEW VDW SURF option, a Van der Waals surface can be generated around the defined atoms. Before defining the set of atoms, the first and last residues of the required stretch should be picked.

ADD VDW SURF
The ADD VDW SURF option can add a new Van der Waals surface around the defined atoms. To define the set of atoms, you should pick the

first and last residues of the desired zone. Each VDW surface is numbered and can be switched on or off in the VDW MENU.

VDW MENU
The VDW MENU option switches on or off each Van der Waals surface. You can select a surface within this menu.

COLOR VDW
The COLOR VDW option switches the system over to the color Van der Waals surface mode. All the subsequent color options will be applied to the selected VDW surface.

The Molecule Menu:
Displaying Loaded Molecules

This menu contains the loaded molecules. With this menu, the loaded molecules can be displayed or not depending on whether their name is highlighted or not. There is no limit to the possible number of entries, but only the names of the first eight loaded molecules are displayed. If you have loaded more than eight molecules, you can have access to the others via the "updown" option on this menu. For each molecular entry, you can manage the display of zones by picking the Z-knob.

P a r t

IV

Tutorials

Part IV provides several typical tutorials sessions in order to give practical information to users of the Turbo-Frodo program. These sessions include, for example, a general users' tutorial, which describes how to display and model a protein, a NMR users' tutorial, which shows how to display, correct, and refine a calculated structure, or a Non-Crystallographic Symmetry tutorial, which explains how to define a symmetry operation and display NCS-related molecules and maps, among others.

General Users' Tutorial

This chapter describes how a Turbo-Frodo session for displaying and modeling a protein may be run.

Creating a Heap

The Heap File is the data bank in which all your molecules are stored. It is an infinite file and is limited only by your disk space. It contains the molecular structures in a format that is readable by Turbo-Frodo.

When you start the program by typing>**turbo**, you will be prompted to give the name of the Heap File. If you answer with a non-existing Heap File name, the program will ask you whether you want to create this file or not. Answering **yes** will create an empty Heap File having the given name.

To fill the Heap File, you have to read the coordinate file about your molecule. This file may be in the Protein Data Bank format (PDB) or the Wayne Hendrickson format (WH). Accordingly, type turbo>**read pdb** or turbo>**read wh.**

The program will prompt you to give the name you wish to the molecule in the Heap File. This name should not be longer than 8 letters, otherwise only the first 8 letters will be taken into account.

The contents of the Heap File may be listed by typing turbo>**list heap**.

Displaying a Molecule

You first have to load the molecule by typing turbo>**load "molecule"**, the "molecule" file name being that given in the Heap File.

Alternatively, you can load a molecule by using its number in the Heap File. Once you have typed turbo>**go,** the molecule will appear on the screen in blue, which is the default color. At the same time as the molecule, the name of the program will also appear in the top-left part of the screen and XYZ axes in the top right, while the origin of the axes will be in the center of the screen. Remember that you can return to Textport by simultaneously pressing down the middle and right buttons on the mouse.

Since your molecule is loaded on the basis of its coordinates, it may happen that it does not appear on the screen. To center it, use the Textport option turbo>**acon "residue number"**. This will put the alpha

carbon (default atom) of the chosen residue in the molecule sequence at the center of the screen.

You may also find that only part of the molecule has been displayed. This occurs if the molecule was previously saved with the SPHERE option still active. To restore the display of the entire molecule, disactivate the display SPHERE option in the View menu then drag the mouse up the screen and from left to right. The following menu heads will appear: Help, Main, Color, Geometry, Editing, Modeling, View, plot, Tools, Surfaces, Molecules. (You can refer to Display Options for a systematic description of these menus.)

Dragging the mouse on the Molecule menu will make the name(s) of the loaded molecule(s) appear. The highlighted name(s) corresponds to the displayed molecule(s). If you want to turn off one of these molecules, click on the corresponding name, and the molecule will shadow. Click again on the same name to bring back the molecule.

You can change the color of the molecule by choosing the SAME option in the Color menu. If you have more than one molecule on the screen, you should have previously selected a specific one before by clicking on any of its atoms. A set of colors will appear, from which you can choose by clicking on one. This step may be useful if you want to discriminate between several displayed molecules.

A single-color display of the molecule is certainly not the most descriptive possible. You may therefore decide to color the molecule with various color codes, depending on the atom residue, side chain (in which case, the backbone will keep its original color), hydrophobicity, or on the secondary structure.

Another way of coloring the molecule is to discriminate between different streches by using the ZONE option in the Color menu. First, however, you must set up the stretches by using the following Textport option, turbo>**zone.** The program will prompt you to give the residue ranges. Answer accordingly, e.g., in the case of a molecule having 64 residues:

Residue range: 1 27, 28 56, 57 64
Residue range: 1 27, 27 56, 56 64

In the above example, there are three stretches running from 1 to 27, 28 to 56, and 57 to 64. To color these stretches, activate the ZONE option and successively select the colors chosen. The molecule will be displayed once the last color stretch has been set. Note that in the case of a CA representation, this setting will create gaps between alpha carbons 27 and 28, and 56 and 57. To overcome this problem specify the zones as follows:

Residue range: 1 27, 27, 56, 56, 64

You can decide to simplify the display by turning off all side chains. This procedure can be accomplished by activating the BACKBONE option within the View menu, thus displaying only the alpha carbons (the CA option within the same menu).

Once the CA display is on screen, you may decide to show only some of the side chains, for example to illustrate the active site on an enzyme. This step can be performed with the CA-SIDE-NAME option in the View menu. When this option has been activated, the program prompts you to give the names of the side chains you want to display. You can also display one or several types of side chains with the CA-SIDE-RES option in the View menu. When this option is activated, a list of side chain types appears. Click on the ones you want to display and validate your choice by clicking on "end".

You may also decide to display only part of the molecule. With the SPHERE option you can display a region centered on the most recently picked atom, corresponding to a sphere the radius of which can be changed using the following Textport command: For example, suppose you want a 5 Å radius sphere,

turbo>radius 5

The sphere will shift each time you center a new atom. (To center, simultaneously press down the left and middle buttons on the mouse.)

The last possible way of simplifying the displayed structure is to look at slices. The thickness and Z-screen coordinate of this section can be set up with the dials (or pseudo-dials). The scale can be displayed by using the Textport command turbo>**slab-display,** which helps you to localize this section in the space.

You may then decide to have a closer look at your structure. You can obtain various stereo views depending on what equipment you have. The stereo options are found in the Main menu.

Once you have chosen the visual representation you prefer, or diagram, you can animate it. If you have a button box connected to your workstation, you can rotate your molecule along the three axes as well as scale it. If not, you must activate the pseudo-dials box by using Textport's PSEUDO-DIALS option.

Warning: Before making any changes to a structure stocked in the Heap File, it is recommanded to duplicate it, so that you will eventually have one copy with the changes (which you will also save) as well as the original. This duplication can be accomplished by using the Textport option turbo>duplicate molecule. The program will prompt you to give the input and output file names.

The most recently picked molecule (which should not be larger than 1000 residues) will turn green, and additional pseudo-dial menus will appear in the lower-right corner of the screen.

With these pseudo-dials, you can rotate and shift the green structure along all three axes. This step can be done quickly and easily in the stereo mode, but if you are not keen on this mode, you can obtain an orthogonal view by clicking on ORTHO VIEW within the Main menu. This view will appear in an area in the top-right corner of the screen. If you are pleased with the result, pick **yes** to confirm the modifications otherwise pick **no**. If you pick **yes**, the changes will be saved on memory only.

If nothing further is done to save them, these changes will be lost when the session comes to an end. To save these modification in the Heap File, use the SAVE option in the Main menu or the SAVE ALL option if you have made changes to more than one loaded molecule. Do not forget that this will overwrite the entry, so be careful.

You can then transfer the result to a PDB, a WH, or an XYZ file by typing the following Textport options: turbo>**make pdb** or turbo>**make wh** or turbo>**make xyz,** as required. The program will prompt you to give the name of the output file.

You may want to keep a hard copy of your work. Check the position that the structure will have on the paper sheet by activating the PLOT FRAME option within the Plot menu. This will display the plotting frame, whose position can be adjusted by selecting its bottom-left corner with the left button on the mouse. Then the PLOT option within the same menu will prepare a Postcript file. The thickness of the line on the plot will be correlated with the color on the screen: the higher a color is on the color map, the thicker is the line.

There are a number of attractive representations, useful for example in a slide session for demonstrating how remarkable you molecule is. These representations are accessible for the selected molecule by means of the COMPACT SELE option in the View menu. Once you have chosen the representation you want, execute it by clicking on COMPACT DO-IT within the same menu. You can interactively handle a few of the representations, such as "Small Circle". For others, orientate your molecule before implementing the compact view.

Modeling a Structure

You have just purified and sequenced a brand new protein or isolated and sequenced the gene involved. Comparisons between the protein's primary amino acid structure and those available from the protein data banks have shown the existence of a fairly good sequence homology with another protein, whose three-dimensional structure has been previously determinated and whose coordinates are available. You may therefore be anxious to see how singular your protein's structure actually

is compared to the other one, and what structural differences between them might explain, for example, the functional differences observed.

The first thing to do here is to use a structure with known coordinates to model your protein. First, align the two sequences, either manually or with your favorite software, as carefully as possible, as they should not be merged with Turbo-Frodo. Then put the known structure on the screen by entering the following command: turbo>**read pdb** or turbo>**read wh,** as required.

The name of the molecule in the Heap File should be different from that of the coordinate file to avoid any confusion after changes are made in the structure. Then load this molecule by typing turbo>**load "molecule",** the "molecule" filename being that given in the Heap File. Once you have typed turbo>**go,** the molecule will appear on the screen, in blue, which is the default color.

You must then modify the sequence of the loaded molecule to fit the one you want to model. Simplify the displayed molecule by choosing the CA option in the View menu, which will display only alpha carbons, or BACKBONE, which will turn off the side chains. To correct the sequence, click on the alpha carbon of the residue you want to change, then go to the Editing menu and choose the REPLACE option. A window will appear allowing you to choose the new residue to be substituted. Move the mouse within the window until the chosen side chain has become yellow, then click on it and accept it with the previous dihedral angle values.

If the new side chain is longer than the former one, it will be replaced on the basis of the dihedral angles of the previous side chain in an extended geometry, without taking the neighboring into account. You will therefore have to correct this geometry manually in order to reduce as far as possible poor interatomic contacts. This correction can be achieved with the TORSION-AUTO option of the TORSION submenu found in the Modeling menu. Pick the residue you want to correct and you will see all the existing torsion angles displayed on the residue. A pseudo-dial will appear numbered according to the defined torsion angles. Using these pseudo-dials, you can now rotate the angles.

During this procedure, you can check the Van der Waals distance violations by activating the NEIGHBOR option in the Geometry menu, the cutoff being set to 1.8 Å, for example, using the Textport command RAD NEIGHBOR. The contact between the most recently clicked-on atom and any of its neighbors will be displayed until the distance between them has become larger than this cutoff.

A less trivial and often necessary procedure is that required to insert or delete a fragment. One or several residues can be deleted by selecting a residue and then using the DELETE option in the Editing menu. The program will ask you to confirm the deletion. Once you have deleted the loop, reconnect the disrupted backbone by using the BOND option in the

77

Editing menu. Once this option has been activated, the cursor will be replaced by **<1**, which will prompt you to select the first atom in the new bond, and then by **<2** , prompting you for the second atom.

Unfortunately, this brand new bond does not have the right geometry. Its geometry can be corrected, however, by using the REFINE option found in the Modeling menu. Once the option has been activated, the cursor will be replaced by **<1**, which prompts you to select the first residue in the strand to be corrected, and then by **<2** for the last one, then start the refining procedure. The corresponding change of geometry during 20 cycles of refinement will occur on line on the screen. At the end of this procedure, you can either continue for the second refinement run (confirm the change by clicking **yes**) or go back to the former structure by clicking **no**.

On the other hand, you may want to insert a loop somewhere. First, disconnect the two residues between which the insertion will occur by using the BREAK BOND option and then start to insert the new sequence in sequential order by selecting the free carbon atom and then activating the INSERT option in the Editing menu. Select among all the residues the one you want. You will then be prompted to give the name of the residue. One way of naming it is to use the position of the insertion with a suffix: For example, if you want to insert 4 residues between residues 14 and 15, the additional residues will be named 14a, 14b, 14c, and 14d.

There are two possible insertion modes: the extended and helical fold modes. It is possible to switch from one mode to the other by cliking on the SHEET/HELIX option within the Editing menu. When you have inserted all the residues you need, connect the most recent residue to the free nitrogen atom resulting from the previous bond break. Use the REFINE procedure as described above to correct the geometry.

At the end of this correction procedure, and if you are satisfied with the outcome, use the SAVE option in the Main menu: This will ensure that the modified structure is stored in the Heap File. If you have made any insertions or deletions, change the numbering of the amino acids in your protein by means of the RENAME Textport option in order for these changes to be applied to the structure in the Heap File. You should then save the modified structure on disk as a **.pdb** or **.wh** file, using the **make pdb** or **make wh** Textport command.

Designing Small Molecules

Imagine you want to know the structure of a serine-protease inhibitor, i.e.,diethyl paranitrophenyl phosphate. Turbo-Frodo can design this molecule. In order to do so, start the program, selecting your current Heap file. You will then find all the appropriate tools needed to design the molecule de novo.

In the display Editing menu, select OPEN MODEL. The program will prompt you to give the name of your own dictionary. This dictionary is a file where small molecules may be saved. You will see lated the interest of such a file. If you do not have a dictionary, give a name. After giving a name, the program will ask you if you want to create a dictionary or not. A new menu pops up in the lower-right corner of the screen. The CONSTRUCT item is the main menu for construction. COMMON, SUGAR, and AM-ACIDS are the subdirectories of the fragment dictionary.

Select COMMON. In the lower part of the screen, several windows will appear, each containing a fragment. These fragments are basic parts of the puzzle. On the left-hand side are pseudo-dials which can be used to move these fragments. You can screen the fragment dictionary by using the lift located below the windows. You will then be able to browse in this dictionary. If you use the right button on the mouse to select a fragment, the fragment's name will appear. Picking an atom of a fragment with the middle button will give you the name of the atom. SUGARS and AM-ACIDS work in exactly the same way.

Select **"--->USER DICT**".You will see the name of the subdirectories, if any, of your own dictionary. You can go back to the main dictionary by selecting **"--->databank**". Now, start designing the inhibitor. Any atom added will be saved in the Heap File.

Go back to the COMMON directory and find the phosphate group. Then select CONSTRUCT and ADD FRAG. The cursor will change into **<1**, prompting you to give a selected fragment. Select an atom from the phosphate group with the middle button on the mouse.

The program will ask you what name the new molecule should be given in the heap and will display the fragment in red. This fragment will then be treated as a residue by Turbo-Frodo, i.e., like any other molecule. Change the color of the molecule and find the hydrogen atom. The design will proceed by creating covalent bonds between two atoms, each carrying a hydrogen atom.

Select ADD FRAG. The cursor will wait for you to give the first hydrogen atom. Choose that from the phosphate group. Once the **>2** cursor appears, pick a benzene hydrogen. The covalent bond is obtained. SAVE if you want to keep this new fragment in memory. Add the nitrate in the same way to the benzene.

The following step would be used to create an ester bond between an ethyl group and benzene. There is no hydrogen once on the phosphate, but you can add the ethyl groups together by using the ADD ATOM option within the CONSTRUCT menu. After the two ethyl groups have been added, the construction will be complete. Do not forget to save this new molecule, using the SAVE option.

In the CONSTRUCT submenu, you will find other items: DEL ATOM and REPLACE ATOM can be used to delete and replace the most recently picked atom. DEFINE RESIDUE makes it possible to reorganize the molecule under construction .

Use BREAK BOND between the two carbons of an ETHYL group: The second carbon to be picked is that belonging to the methyl group. Then select DEFINE RESIDUE. The program will prompt you to give the type of the residue. Enter **meth**. You will now have 6 residues, including a "meth". If you now delete this residue and add a hydrogen atom to the trivalent carbon group, you will obtain a methyl-ethyl-paranitrophenol phosphate. You can see that the new O-C-H angle is 141°6, which cannot be correct in the case of a methyl group. The DEFINE ANGLE option makes it possible to enter the right value, i.e., 250°5. You can correct a distance in the same way.

The brand new molecule can be used as a basic part for further constructions, provided this molecule has been inserted into your own dictionary: Select DATA BANK within the CONSTRUCT menu, then USER DICT. A window will appear, and you will then be able to play with your dictionary. You can list its contents, create subdirectories, and so on. Select INSERT molecules with the right button on the mouse. If your dictionary is new, you will have to create a directory by using CREATE DIRECTORY.

Go back to the display by using END and use **"---> USER DICT**" to display your own dictionary. You can see that your molecule is now being treated like a fragment in its own right.

X-Ray Users' Tutorial

This chapter describes how a Turbo-Frodo session may be run after obtaining an electron-density map.

Turbo-Frodo uses the classical DSN6 maps from the older Frodo packages. These maps can be calculated from Fourrier maps with the Mappage program, which can read the specific Fourrier map format. Some of the most classical crystallographic packages are CCP4, PROTEIN, and XPLOR.

A Mappage program for Silicon Graphics equipment is provided along with Turbo. This program is located under the **$TURBO_DIR/map** subdirectory.

Whether your map has been produced with XPLOR or CCP4, you should be able to change its format with the supplied Mappage program. For the XPLOR package, there are no particular instructions. For the CCP4 package, however, you should scale the density with the inverse of the rms deviation from mean density. You can see these instructions while running the Mappage program. The output map is the file used by Turbo-Frodo.

Loading the Turbo-Frodo Maps

Since maps are logically linked to a molecule, you need to have one molecule in the heap, loaded in memory. You can load up to five maps for each molecular entry:

turbo> **load "molecule"**
turbo>**load map**
Give the name of the Turbo map file.

You can redefine the map parameters either when loading or afterward from the pull-down MAPS submenu. The default levels are 1 for the first map (usually a 2fo-fc) and 2 for the four others (fo-fc).

CHAPTER 18

NMR Users' Tutorial

This chapter describes how a Turbo-Frodo session may be run after calculating a protein structure using a distance geometry program (DISMAN, for example) or a constrained molecular dynamics program (XPLOR, for example). These programs, using as input a set of distance constraints based on NOESY spectra, or other sources such as the disulfide bonding or hydrogen bonding, have produced an output file containing the coordinates of the molecule studied.

Displaying a Calculated Structure

You first have to check the format of this output file, which has to be under a PDB or WH format. Since this structure is based on NMR data, you do not have to worry about hydrogen generations, which should be present in the coordinate file. Turbo-Frodo can also work with the necessary pseudo-atoms if no stereospecific assignment has been carried out.

The format of the input file containing the distance constraints also has to be written in the format used by XPLOR. Otherwise, you will have to make the corresponding corrections everytime (see the Necessary Files chapter about correct format). You then have to fill up a new or previously existing Heap File with your new structure by using the following commands, turbo>**read pdb** or turbo>**read wh.**

You should then load this molecule by typing turbo>**load "molecule"**, the "molecule" file name being that in the Heap File. Alternatively, you can load a molecule by using the number it has in the Heap File.

Evaluating the Calculated Structure's Quality

To make the best use of the program, you should also load the distance constraints file used for structure calculation or refinement. This step can be achieved by typing turbo>**load nmr.** The program will ask you to give the name of this file. Answer accordingly, assuming this file to have the right format (see the Necessary Files chapter about the correct format). The program will then list the atoms used in the distance constraints file but not in the PDB file that you have used. This problem arises mainly when the names of pseudo-atoms that you have used are not correct according to the XPLOR format.

You can now look at your new structure together with the constraints you have used to calculate it by typing turbo>**go**. The structure will appear in

light blue together with the constraints, which will be colored depending on the violations, if any:

Green: The distance between the two protons (or pseudo-atoms, if any) is lower than (or equal to) the constraints used.
Orange: The constraints used is violated by not more than 0.2 Å.
Red: The constraints used is violated by more than 0.2 Å.

Note: This tolerance of 0. 2 Å can be modified with the TOLERANCE Textport option.

The constraints can be turned off by dragging the mouse through the Tools menu, until NMR is highlighted. If you click on the right button, the NMR submenu will appear. You can then activate the ON/OFF option, which will alternatively turn the NMR constraints display on and off.

Alternatively, you can turn off the molecule display by dragging the mouse to the menu on the very left and clicking (by using the right button) on the highlighted molecule name (the one that is currently displayed). This will make the structure vanish while the distance constraints are still displayed.

Since a standard distance constraints input may contain up to 10 constraints per residue, displaying all of them at once often gives a messy picture. To simplify matters, you can display only the violated constraints by choosing the VIOLATION option in the NMR submenu.

Another way of estimating how the structure was calculated is to draw a picture of the constraints over the molecule sequence. This step can be accomplished by choosing the COL REST RES and COL REST ATO options, which color the residues or atoms, respectively, depending on the number of violations in which they are involved. A yellow residue is an unconstrained one, and the colors range from red (few constraints) to blue (many constraints). With this option, you will immediately see which parts of the molecules have been poorly determined and how flexible they are. This may be roughly compared to the B-factor in X-ray crystallography.

You can also carry out this estimation on one particular residue, using the DISPLAY RES option in the NMR submenu. This will cause the constraints relating to the clicked-on residue to be displayed.

To finally estimate the quality of this initial structure, you have to check its geometrical accuracy by calculating the phi and psi conformational angles. This can be done by picking the RAMA ALL option in the Plot menu.

Correcting the Calculated Structure

Now that you have estimated the quality of the calculated structure in terms of violations, and the geometry of its residues, you can try to correct it manually. As long as there are not many violated distance constraints (otherwise it would be better to sort out the input file constraints first), you can try to satisfy them by modifying the structure while displaying the constraints.

After displaying the constraints (see above), change the geometry of the side chain involved using the TORSION-PICK option in the Geometry menu. You will first have to define the torsion angles by picking successive atoms. Note that four atoms are needed to define the first torsion angle, and then for each additional picked atom, a new angle will be defined. Up to six torsion angles can be defined, and these will be numbered on the displayed structure. Validate your choice by clicking on **yes**, and the pseudo-dial will appear, corresponding to the defined torsion angles. Using these pseudo-dials, you can rotate these angles. The displayed constraints will follow, and their color will change depending on the distance. You should try to keep the modifications to a minimum, and stop rotating, for example, when the constraints' color turns green.

During this procedure, you can check the Van der Waals distance violation by activating the NEIGHBOR option in the Geometry menu, after setting the cutoff to 1.8 Å, for example, by using the RAD NEIGHBOR Textport command. The contact between the most recently picked atom and any of its neighbors will be displayed once the distance between them has become smaller than this cutoff.

At the end of this correction procedure, if you are satisfied with the outcome, confirm by picking **yes** or go back to the original structure by picking **no**. This will save the changes made during the open Turbo-Frodo session only. If you want to store the result on disk, activate the SAVE option in the Main menu. This will save the modified structure in your Heap File.

Refining the Structure

You can now submit the corrrected structure to whatever energy minimization or simulated annealing procedure you want. However, you will first have to produce the corresponding coordinate file, by using the following Textport command, turbo>**make pdb** or turbo>**make xpdb**, if you want to refine the structure with XPLOR.

Alternatively, Turbo-Frodo makes it possible to calculate a list of theorical nOe peaks, provided you have the list of protons in your protein. At present, this is still only a rough procedure, consisting of extracting couples of protons separated by 5 Å or less and calculating a peak volume without taking any relaxation matrixes into account. However, you can load a list of the peak on your spectrum, either manually or

automatically, provided you use the EASY program as the assignment software. You will thus be able to label previously unassigned NOESY peaks.

For this purpose, use the following Textport command, turbo>**nmr-spect.** The program will prompt you to give the input proton list, which has to be written in the EASY format or directly produced by it (see the Necessary Files chapter for an example of this format).

The output files are a list peak and a .ref file, which have not yet been used. The list contains a list of peaks for each peak from left to right: the number of the peak, the chemical shift of the first proton, the chemical shift of the second proton, the color of the nOe peak (if displayed using EASY), and a roughly calculated volume (without taking any relaxation matrixes into account). The other entries have not yet been dedicated.

Playing with Surfaces Tutorial

Turbo-Frodo offers several ways of displaying the surface of your protein.

The Van der Waals Surface
Once you have loaded the molecule, you can calculate the Van der Waals volumes of all or part(s) of the molecule. This procedure can be accomplished by activating the ADD VDW SURF option in the Surface menu. The cursor arrow will be replaced by **<1**, prompting you to give the first residue, in the stretch of which you want to calculate the surface. After this selection, **<2** will appear, prompting you to give the last residue in the stretch. During the calculation of surface, the Textport window will appear, giving you information about what is happening. Note that the Van der Waals surface is calculated for all the heavy atoms in the stretch, no matter which mode is used for the molecule. This surface can be colored differently by activating the COLOR VDW option in the Surface menu, and then choosing the color within the Color menu.

You can add another Van der Waals surface by choosing the ADD VDW SURF in the Surface menu, or start another calculation by activating NEW VDW SURF. If you have several calculated surfaces on screen, you can display one or several of them by selecting them in VDW MENU. Those that are selected will be highlighted in yellow.

You can choose the number of dots per square angstrom via the following Textport option, turbo>**vdw-density**, and answer with what follows:
1 for 20 dots/Å^2
2 for 40 dots/Å^2
3 for 80 dots/Å^2

The Connolly Surface
This is the surface of a molecule accessible to a water molecule. It can be calculated with the **ms** program in **TURBO-DIR/ms**, after updating the **ms.in** and **ms.com** files. Copy the **ms.in** and **ms.com** files in your working directory before updating them.

Once the molecule that needs to be calculated has been loaded, go to the Textport window and use the >**make ms** command. The program will then prompt you to name the output file. Give the output file a name with the **.atm**. extension. This file is the coordinate input file of the **ms** program.

Then, outside Turbo, for example by opening another window, edit the **ms.com** file, which is in the **TURBO-DIR/ms** subdirectory and substitute **your_molecule.srf** file with the name of your molecule, keeping the extension as it is. The **your_molecule.srf** file is the output surface file.

Edit the **ms.in** file and change the coordinate input file name **your-molecule.atm** into the name previously given as the output surface file. Then run the **ms** program:>**ms.com.** Since this is fairly time consuming, it is best to run this program as a batch job.

The output surface file of **ms.com** has the **.srf** extension and will be used if you type the turbo>**connolly** Textport command. You will be asked whether you want a simple calculation or to link the points to atoms. In the first case, you do not need to have a loaded molecule, but this will produce a single object, and Turbo will be unable to color it depending on the residue type or the side chain.

This will prepare the real surface file having the **.ts** extension, which can be displayed by typing turbo<**load surface**. You will then be asked to name the surface file (the one with the **.ts** extension). The program knows what kind of surface file is being used. If the surface is to be calculated including the atom dependence, you must have previously loaded the molecule.

The Spline Surface
This is the surface calculated from the Connolly surface which is displayed in the form of lines obtained by interpolating the Connolly surface points.

Using the Connolly Surface File as above (the one obtained with **ms.com**, having the **.srf** extensions), run the **prep_spl** programs followed by **make_spl**, which are to be found in the **TURBO-DIR/ms** directory.

>prep_spl

You will be asked to give the thickness of the slice and the name of the output file which is used as the input file for the program.

>make_spl

Once again, give the name of the output file thus created. Once you have obtained this file, use the following Textport command to read the Spline Surface File turbo>**spline.** The program will then produce an output file with the **.ts** extension. To display this file, use the Textport command turbo>**load surface**.

Warning: If you calculate both the Connolly and the Spline surfaces of the same molecule, you will produce the "your_molecule.srf" and "your_molecule.spl" files. In this case, the program will transform both files into "your _molecule.ts", thus overwriting the first file to be produced. To overcome this problem, give the two surface files different names.

These surfaces can be colored differently by activating the COLOR SURF option of the Surface menu, then choosing the color from the Color menu. You can switch off the surface(s) by activating the NO SURF option within this same menu. Assuming you have calculated and loaded several surfaces, you can display one or several of them with the SURF MENU option, by selecting the required surface(s).

Non-Crystallographic Symmetry Tutorial

Non-Crystallographic Symmetry (NCS) in Turbo-Frodo relies upon four successive steps: definition of the domains related by the symmetry, definition of the symmetry operation, display of the NCS-related molecules, and display of the NCS-related maps. These different steps are managed by the Textport options DOMAIN and NCS, respectively, and by the displayed NCS submenu.

Definition of the Related Domains

You must first be able to define the different domains of the molecule. A domain can be one entire molecule if you have several in the asymmetric unit. (In this case, all molecules of the asymmetric unit must be in the same heap entry.) A domain can also be a part of a molecule if this molecule has an internal symmetry. You can define a domain by using the Textport option DOMAIN. One domain can include several zones and is automatically saved in the heap.

Definition of the Symmetry Operation

Turbo-Frodo considers pairs of domains, and thus one complete symmetry operation (rotation matrix and translation) between two domains. If you have three related domains, you have to define the three related pairs (two may be enough, depending on how you want to work).

You have three different ways of indicating the related pair and its symmetry operation: You can either read it from a file, read it from another heap entry, or calculate it. All of these options are managed by the Textport NCS handling.

In case you have to calculate the rotation and translation, choose the following sentence, **1-Define an operation symmetry,** in the NCS handling. This calculation works as the RIGID option does: It is a rigid body superposition of the two domains. You must therefore indicate some of the atoms from the two domains, which fit together. If you need help or more information on the rigid body superposition, you can refer to the RIGID help. Once defined, the pair and symmetry operation are saved in the heap, so you do not need to save all over again. The inverse operation is also calculated.

If you have a new heap entry, but NCS has already been defined for another one, you can just copy NCS from one molecule to the other. In

order to do so, choose the following sentence, **4-Read a matrix from another heap entry,** in the NCS handling. (The source molecule has to be loaded.) For this option, both domain definitions and symmetry operations are copied. You can also specify a name file that contains the symmetry operations (**3-Read a matrix file**). This file must have the same format as the one dumped by the display option PRINT NCS. It is a way of copying the symmetry from one molecule to another but not the domains. The domains of both molecules can be different.

Display of the NCS-Related Molecules

When domains and symmetry operations are defined, you can work with the Non-Crystallographic Symmetry. The domains can be separately colored by using the numerous options in the Color menu. To display the related molecule, you must use one of the two options of the NCS submenu (Tools menu): APPLY MATRIX or AUTO APPLY.

For APPLY MATRIX, the different related pairs are proposed, and you have to select one or several pairs. You can note that the inverse operations are also proposed. On display, both domains of a related pair are superimposed. This is only a display operation that does not affect the coordinates of the transferred domain. (A mistaken SAVE will not have any consequences.)

If you choose the AUTO APPLY option, all the domains related to the centered one (CONTOUR residue) will be transferred. The symmetry applied may change if you CONTOUR on another residue (belonging to another domain).

Display of the NCS-Related Maps

You may also be able to apply the symmetry to the displayed maps. You can display a map box of a domain with the NCS MAPS submenu. The new box has at first the same parameters as the original map: same levels and colors. These parameters can be changed by choosing the **OP** knot of the NCS MAP menu. You can ask for different levels and colors, dashed lines or solid ones, but the box width remains unchanged, as does the axis contour.

You will also be asked for the type of map that you would like to be displayed. Indeed, not only the map difference can be displayed, but also the average density. The map difference is represented by the absolute value of the difference; furthermore, only positive values are considered. The average density represents the average between all symmetric boxes and the original one. These operations are applied between the original map box and the transferred ones. If you work with a three-fold axis, the two transferred map boxes will be superimposed to the original map box, except if you choose average density, where you see only one superimposed average map.

Ab Initio Building Tutorial

Ab initio building is a new and main feature of Turbo-Frodo, and from now on, you will be able to work with an MIR map. We propose here a new way of building a protein de novo, with the presentation of a new tool: the Trigonal Planar Pseudo Residues (TPPRs).

While moving along your new MIR electron density map, you will see a stretch of helix, or ß strand. You can place C Alpha atoms in a very easy way by adding some TPPRs anywhere you wish on the density map. By asking to connect these TPPRs, you will spot the CB atoms and will be able to describe the side-chain density. When asking to align those fragments with your protein sequence, you will orientate the fragment and place the N and C atoms. Finally, a special algorithm of replacement will fit the O carbonyl and the whole peptidic bond in the density map. All these features can be found under the displayed BUILDING and BONES submenus. The few required parameters may be set via the Textport options.

Moving Along the Density Map

You are now able to move along an empty density map. As you contour an atom, press down the left and middle mouse buttons anywhere on the map, and the whole image will center on the point that you have selected. When selecting a point on the screeen, you in fact select a line that is orthogonal to the screeen and to the displayed slab length. The new center is the highest density point found along this line. When several maps are displayed, a research is carried out among all the maps. If you do not get the expected results, you can decrease the slab, which may cause the selection to become more accurate.

Adding New Entities in the Density Map

To build ab initio in the electron density map, the strategic choice relies on the idea of TPPRs. A TPPR is constituted of four atoms: CA, C1, C2, and C3. The CA atom is in the center of the trigonal plan formed by the three other atoms. The aim is to find the best position for the CA atom, and in order to do so, we chose to handle a TPPR that is easy to place in the crossings of the density that you can spot. This is a planar residue because when looking at an empty density map, evidence is not always available to orientate the residue (or to decide which atom is N and which one is C) and therefore to orientate a real tetragonal carbon. When you click in the density, the position of the TPPR is defined with the same algorithm as that used when centering in the density map. Therefore, a

Real Space Refinement of the only orientation places the three branches of the TPPR in the density map.

Localization of the TPPR center   Real Space Refinement of
on the highest density point   the orientation to determine
                                                                     the best density fit

Sometimes it can be useful to refine the center position, and therefore a Real Space Refinement that performs translations can also be added to the previous one.

By selecting the ADD TPPR option of the BUILDING submenu, you will be able to add a TPPR on the map. Select one of the following two options, Build Mode 1 (BM1) or Build Mode 2 (BM2), at the right of the screen, then select a point in the density by clicking on the left mouse button. The first option (BM1) allows an RSR of the TPPR position by translation, while the BM2 option does not. BM1 and BM2 are useful, depending on the shape of the electron density. When several maps are displayed, you must choose the working map with Textport's WORK-MAP option in order for RSR to be performed on this particular map.

For the first TPPR added, the program will prompt you for a residue name, suffix included (e.g., 24A). The following residue names will be automatic increments of this first one (e.g., 25A, 26A, and so on). The TPPRs are created with a TPPR residue type, which can be changed when activating the CA LABELS option of the BUILDING submenu. A label window appears at the lower-left corner of the screen. If you select the residue type, you may be able to choose another type in the appearing submenu. Residue types include Short, Medium, Long, and Aromatic and can serve as a way to describe the shape of the density, since it is difficult to recognize the residue type in an MIR map. If you select the comment item of this label window, you can attach a comment to your residue. The comments can be listed or dumped in an ascii file by using the following Textport orders, LIST COMMENTS and MAKE COMMENTS, respectively.

Managing the Connectivity

Once you have added some pseudo residues in the density, you have to connect them. While working, you may have added the pseudo residues in a random order, and when you list all the residues or ask for a CA trace, the residues do not have the good connectivity. You can create this connectivity with the MAKE FRAGS option of the BUILDING submenu. This step will connect the CA atoms that are 5 angstroms apart, but this cutoff distance can be changed by using the DIST-CA Textport option. Furthermore, the consistency of the connectivity will be checked. A bond that crosses the density (that is, goes through an empty area) does not exist, although the distance can be below the cutoff. For each CA, the program will choose the two bonds whose density fittings are the highest among the allowed neighbors (distance smaller than the cutoff). The density cutoff is set to the first displayed level of the working map.

When proceeding in this manner, a distance cutoff set between 4.5 and 5.5 is recommended but is also highly dependent on the electron density map that you have. Sometimes, for example, you will find that the distance cutoff must be diminished when connecting residues in helix conformation, otherwise the two neighbors chosen for a given CA will not be the two expected ones.

Identifying the Pseudo Residues

The fragments that you have just assembled are neither orientated nor placed along the target sequence. This target sequence (that is, the currently built protein sequence) must be loaded with Textport's LOAD SEQUENCE order, and it will be attached to the molecular entry and automatically saved in the heap file. Go through the alignment screen by activating the ALIGN FRAGS option of the BUILDING submenu. You will discover in the upper part of the screen the fragments that you established before, and in the lower part the target sequence. (Pressing down the left mouse button will end the alignment screen and return you to the molecule).

When selecting a fragment with either the right or middle mouse button, the alignment is performed, and the ten best solutions are proposed: Just choose one of the ten buttons in the upper-right corner of the screen. The fragment is aligned under the displayed target sequence with a color code: green (10-point score) for correct matching, red (0 points) for bad matching, and yellow (5-point score) if both residue types belong to the same group (Short, Medium, Long, or Aromatic). For each proposed solution, the score (or sum of residue scores) is given, as well as the orientation of the fragment. You therefore realize that the residue type that you gave to the TPPR is extremely important and that the shape of the density must be well described.

If you are satisfied with a proposed solution, you can confirm it with the following consequences: The fragment will disappear from the pool of

fragments, the pseudo residues will be renamed after $x (x depending on the place the pseudo residues now have in the target sequence), and finally the atoms will be named after C, N, and CB as the fragment is now orientated. Overlapping is managed and you can decide to put a fragment back in the pool if you confirm another one at its place.

Having Real Residues

You now need to replace the connected TPPRs with real residues, and this according to the target sequence. The MULTI REPLACE option of the BUILDING submenu will perform this replacement and will replace as many pseudo residues as you would like. Simply select the first and last TPPRs of a fragment. If you wish, all the pseudo residues will be replaced by a real residue according to the target sequence. If you do no wish to do so, answer no to the appearing question; the residues whose type does not correspond to the target sequence (e.g., Short, Aromatic, or anything else) will not be substituted.

The MULTI REPLACE option is not as simple as the REPLACE RES option of the Modeling menu. You noted that up to nown, the O carbonyl (which is as important as the side chain) is missing. This option is in fact an option to place the O, and therefore the whole peptidic bond, but according to the density map. An imaginary peptidic plan (CA_i, C_i, O_i, N_{i+1}, CA_{i+1}) is rotated on the CA_i-CA_{i+1} axis of two TPPRs, and the best density fit is kept. This implies that C_i of the $TPPR_{(i)}$ is moved, as well as the N_{i+1} of the $TPPR_{(i+1)}$ to build a real peptidic plan with the best density fit. The CB of the former $TPPR_{(i)}$ is not moved and is the starting point for the side chain which comes in the most frequent conformation, according to the rotamer library of J.W. Ponder and F.M. Richards. It goes on building peptidic plans and side chains until the last residue you asked for. For this last residue, the oxygen is placed without the help of a map.

Once the completed fragment is built, 20 cycles of refinement are performed (with the REFI algorithm) to correct the geometry. Note that this option does not work only with the TPPR, but is also quite performant to correct a peptidic bond between 2 residues, even in the case of the Molecular Replacement method.

Residues Improvement Workshop

A toolbox has been set to adjust the conformation of the residue according to the density map. A permanent menu appears when activating the TOOLS-BOX option of the BUILDING submenu. To find the best density fit manually, you can use the RESI ROT/TRANS option, which is an adaptation of the FRAG ROT/TRANS option. The whole residue is automatically selected for both translation or rotation, and there is no need to break some bonds to isolate the residue. The rotation center is the last picked atom on that residue. The ROTAMERS option proposes all the different rotamers on the side chain, as they have been defined by J.W. Ponder and F.M. Richards, and offers a way of getting the conformer that is best suited to your density. The RSR MAIN and RSR

SIDE options perform Real Space Refinement of the residue's main chain (CB included) and side chain, respectively. All these options work on the most recently picked residue.

An Alternative: The Bones Map

You may wish to get help from a bones map. To obtain the bones map, you need to run the following supplied program, **$TURBO_DIR/map/bones_turbo**, and then load this new bones map by using Textport's LOAD BONES order. When selecting a bones point, you will see a label window that appears in the bottom-left corner of the screen. In this label window, you may change the name (or label), the bones type, or the residue type of the last picked bones point by clicking with the right mouse button on the appropriate place. The bones type is one of the four following types: C ALPHA, MAIN, SIDE, UNDEFINED; choose one in the appearing pop-up menu. You will be able to change the residue type for those that have a C ALPHA bones type. These residue types include Short, Medium, Long, and Aromatic. When moving along the bones map, and with the help of the electron density map, you should spot all the bones points which seem to be at a C Alpha place and identify them properly. You can then have a sort of CA trace: A C Alpha bones trace, by selecting the BONES C ALPHA option of the BONES submenu.

Once you have located these CA points, you may want to have real residues. You can create residues of only one CA atom by activating the CREATE CA option of the BONES submenu. Select the first and last CA bones points of a connected set. The new residues will come at the exact location of the identified C Alpha bones point and will also have their residue type. Above all, however, you can create TPPRs by activating the CREATE TPPR option of the BONES submenu. These TPPRs will be created via the BUILD MODE 1 (both rotation and translation performed), where the C Alpha bones point will be the starting point (first location of the CA atom). When you have the TPPRs, you can enter the alignment screen and thus identify the fragment that you have built.

Index

101