XEASY - MANUAL

ETH Automated Spectroscopy for X Window Systems

User Manual

Tai-he Xia and Christian Bartels

General information:

Table of Contents

Changes since Release 1.2

Miscellaneous

- Availability
- Upgrades
- System requirements

SPSCAN manual for CSIC group members

recommended for peak picking and integration of 3D spectra, ECOSY evaluation, reduced-dimensionality spectra.

Please report bugs to Ralf W. Glaser, xeasy@mol.biol.ethz.ch.

Last modified: March 16th, 1998

Looking

Table of Contents

[Acknowledgements](#)

2

3

4

5

Known Bugs

6

Acknowledgments

1. An Overview

Today computer-aided spectrum analysis has become routine in the 3-dimensional structure determination of macromolecules by NMR. In our laboratory a package called EASY3D (C. Bartels, K. Wüthrich, unpublished results) an extension of the program EASY (C. Eccles, P. Güntert, M. Billeter and K. Wüthrich, J. Biomol. NMR 1, 111-130, 1991) has been developed for this purpose. EASY3D is an interactive computer program based on the SUNVIEW graphic user interface supported by the Solaris operating system version 1.1. Unfortunately, the new generation of the operating system Solaris 2 released by SUN will not continue to support the SUNVIEW window system. Therefor a new program based on a new window system is required.

The X window system is a portable software standard developed at MIT in the United States of America. Applications which use X windows to operate a workstation display can be run easily on a variety of workstations from a variety of computer vendors. In this sense, a program written to run on a workstation will be largely machine-independent if the X library is employed. In recent years, the X window system has become increasingly more popular.

XEASY is written de novo, based on the X Window System to satisfy the requirement for NMR spectrum analysis. Only the routines for integration, automatic anti-phase peak-picking and contouring are translated directly from EASY3D. One aim when rewriting EASY3D was to add functionality by making the routines more flexible yet remaining compatible to the EASY program. The program has been tested on Silicon Graphics and SUN workstations. To compile and link the program, the OSF/Motif widget set is needed.

XEASY does not require any particular assignment strategy but offers a set of flexible commands, data structures and display routines. Possible applications include assigning and integrating cross peaks in a homonuclear NOESY spectrum, identifying spin systems in an HCCH TOCSY, connecting back bone fragments using a HNCA experiment, or integrating cross peaks of a series of spectra to measure amide proton exchange rates. Since the interplay of the different commands and options is crucial, the chapter The XEASY Model on page 17, provides general explanations of how the different commands may be used for the different tasks. Examples of how the different commands can be integrated for use with specific assignment tasks are given in the chapter Applications on page 109.

A number of small accessory programs not integrated into XEASY, but designed to be used in conjunction with XEASY, are described in section 15.0 on page 103. They include programs for automatic picking of in-phase peaks and for peak integration (N. Schäfer, diploma thesis ETHZ).

Section 16.0 on page 121 gives guidelines for the installation of XEASY and the accessory programs

---

Looking

---

2. Using XEASY

The XEASY program is designed to assist the spectroscopist in making decisions. Thus the interaction of the program with the spectroscopist is an important design consideration of the program, This chapter explains the main parts of the user interface.

2.1 Starting XEASY

Log into a Unix workstation running X Windows and issue the command

```
xeasy [-l] [-ini param_file] [-m macro_file]
```

to run the program. The command-line flags given in brackets are optional.

Option
-l
> This option indicates that the output binary file called "xeasy.default" from the previous session will be loaded. The contents will be restored. While the program is running, all data will be saved into the file "xeasy.default" in your working directory every five minutes. This is also done when the program is terminated normally. If there is a crash, the data from five minutes ago can be retrieved with this option.

-ini param_file
> The parameter file param_file (absolute or relative to the default spectra_path) is loaded.

-m macro_file
> The macro file macro_file (absolute or relative to the default xem_path) is executed before the program accepts interactive commands.

In addition, XEASY accepts all of the standard X command line options. Some of them are described here.

-fn font_name or -font font_name (e.g., -fn 9x15)
> This option specifies the font resource. The default font_name used in XEASY is 8x13. On many workstations, the available fonts can be found in directories such as /usr/lib/X11/fonts and viewed with the tool xfontsel

-fg color or -foreground color (e.g., -fg Yellow)
> This option specifies the foreground color resource. The default is "Black".

-bg color or -background color (e.g., -bg Red)
> This option specifies the background color resource.

-display display (e.g., -display troll:0)
> This option specifies the X server to contact. If XEASY will be run on machine A, and the display will be on machine B, two steps are needed to start the program. (1) Log into the machine B, issue the command "xhost A"; (2) log into the machine A and issue the command "xeasy -display B:0".

If the option "-l" is not used, XEASY will create the main window, and popup the file window for input (refer for example to "New spectrum [ns]" command). XEASY can be terminated by selecting the command "Exit" from the menu or by typing the two-letter code "ex".

Improper keyboard setup may result in warning messages when starting XEASY on some SUN workstations. This can be avoided by defining the proper libraries. In our case it could be omitted by issuing the following command before running XEASY:

```
setenv      LD_LIBRARY_PATH      /usr/lib/X11
```

2.2 The mouse

8

The mouse has three buttons. The left mouse button is generally designed to select a region or to select a point in a window. The middle button is designed to start an action after a selection with the left button. Holding down the Shift key while pressing the middle mouse button will abort the current command. In addition, pressing only the middle mouse button refreshes the display (screen). The right mouse button is designed to popup the menu. In the text window (which will be described in the section "Windows") the popup menus on the command bar should be popped up by pressing the left mouse button, which is the only exception in XEASY where not the right mouse button is used for the menus.

2.3 Selecting a region

To select a region, press and hold down the left mouse button while moving the mouse. A dragged rectangle will be drawn as the mouse moves. Release the left mouse button to end the selection and **move the mouse away from the release point.**

2.4 Selecting a point

Position the cursor, press the left mouse button and release it.

2.5 Two-letter codes

Two-letter codes are designed for speeding up operations. For each command there is a corresponding unique two-letter code. Every key press is recorded in the memory, after two strikes the program will try to execute the command. The "RETURN" key is designed for cleaning up any wrong typing. When pressing the "." key the action of the last key is repeated. A special command is the "Apply to all ... [aa]" command. Using this command together with another will `synthesize' a new command, which is not listed in the menu. For example, "Remove lines [rl]" command removes lines one by one. After one selection with the left mouse button, the selected line will be removed from the screen. But if the user types "aa" before using the command, all lines are removed at once without confirmation. Therefore, the user is expected to use "aa" carefully. The "aa" command can only be applied to a limited subset of the XEASY commands.

2.6 Directories

For all releases up to 1.2.11 the user must have a directory "~/easy" and "~/easy/data", and some environment variables must be defined (see XEASY Installation).

In release 1.3 and later the program uses X Window user-defined resources to define the default behaviour. The relevant resources are listed and commented in XEasy. More details about default file paths can be found in the chapter about user-defined resouces.

From Release 1.3.4 on, the files "xeasy.default" and "xeasy.text" are stored in the startup directory. "xeasy -l" searches for "xeasy.default" only in the current directory.

2.7 Windows

There are 9 windows in XEASY. They are

1. main window;
2. default window;
3. text window;
4. color window;
5. overview window;
6. peak data window;
7. peak editing window;
8. assignment window;

9

9. slice window.

Each window can be iconized, moved, resized, and destroyed with the exception of the peak data window, which is a dialog-type window. If the main window is destroyed then the program is terminated.

Main window
> This window consists of two subwindows. The small one on top displays the coordinates of the mouse and the functions of the mouse buttons. The big one at the bottom is used for displaying spectra. This window is created when XEASY is started. By moving the handle on the border line between them vertically, the relative size of the two subwindows in the y-direction can be adjusted.

Default window
> This window is used for setting up the controlling parameters for peak display, color of contour lines etc. It is an independent window, that is, it can be resized, iconized, destroyed and recreated. The command for creating or removing this window is "Default window [dw]". the "dw" in the bracket indicates the equivalent action by typing the two-letter code from the keyboard. This is also true for other commands which will be described hereafter.

Text window
> This window is used to display the log file which records the individual commands, some statistical results and error messages. In addition, it can be used as a normal editor. The contents of the text file is stored in the file "xeasy.text" in the easy directory. The command for creating or destroying this window is "Text window [tw]" .

Color window
> This window is used for adjusting the colors in the displayed spectra. There are two slide bars in the panel. The upper one controls the change of the contrast and the lower one adjusts the threshold for displaying the intensities. The command to create or remove the color window is "Color window [cw]" .

Overview window
> This window is used to monitor the display in the main window. The whole spectrum is displayed, and the zoomed regions from the main window are indicated by rectangles. The command to create or destroy this window is "Overview window [ow]".

Peak data window
> This window is used for setting up the information displayed with the peaks, such as, peak number, volume, assignment. The command to create this window is "Peak data window [pw]". This window is a dialogue-type window, that is, it blocks the functions in the main window. After the "OK" button on the panel is pressed, the functions in the main window become again active, and the peak data window is removed.

Peak editing window
> This window is used for editing the peaks and their assignments. As soon as the contents in the editable fields are changed, the lists will be updated accordingly. Illegal modifications will be ignored. The command to create or destroy this window is "Peak editing [ew]" .

Assignment window
> This window is used for working with possible multiple assignments of a selected peak. The command to create or destroy this window is "Assignment [aw]".

Slice window
> This window is used for displaying 1D cross sections from the spectrum. The command to create or destroy it is "Slice window [sw]".

In addition, there are two more dialogue-type windows, they are the file popup window and the user's response window. The former is used to define the names of input and output files, the latter is used for getting various user's responses.

---

next: 3. The XEASY Model / contents

Looking

10

3. The XEASY Model

To enable computer supported spectra interpretation, an operational model is needed which defines the terms on which the computer interacts with the spectroscopist. This Model has to include the spectra which are the raw data and abstractions such as peaks, assignments, line shapes, geometries and strips which are derived from the spectra and constitute the results of the program. These different elements are detailed in this chapter.

3.1 Spectrum

A spectrum is a 2D, 3D or 4D dimensional box of intensities containing the frequency domain data of an NMR experiment. Spectra are loaded using the command "New spectrum [ns]" .

The intensity information for each data point is either encoded in 8 bits or in 16 bits. The format with 8 bits uses a logarithmic representation of the data with 1 byte per real data point. For a given data point s_k the program first determines the integer l that minimizes the expression

$$\left| |s_k| - \sqrt{2}^{\,l} \right|$$

3.1.A

(i. e. $l = \sqrt{2} \log |s_k|$) and then stores in one byte

$$
\begin{aligned}
e_k &= \min(l+1, 47), && \text{if } s_k \geq 0; \\
&= 95 - \min(l, 47), && \text{if } s_k < 0.
\end{aligned}
$$

3.1.B

This format can represent numbers approximately in the range $-\sqrt{2}^{\,47} \leq s_k \leq \sqrt{2}^{\,46}$; i. e. $-1.2 \cdot 10^{7} \leq s_k \leq 8.4 \cdot 10^{6}$ with a relative error of less than 20%. The format with 16 bits uses a 16 bit floating point format with the "exponent" e_k given by Eq. 3.1.B in the lower valued byte and the mantissa

$$a_k = 721 \frac{|s_k|}{\sqrt{2}^{\,l}} - 615, \qquad \text{if } l \neq 0$$

3.1.C

(if) in the higher valued byte (Eccles et al., 1991). This format can represent numbers in the same range as the 8 bit format but with a relative error of less than 1%.

The intensity data is stored in a spectrum data file with the extension ".3D.8" or ".3D.16" depending on the used accuracy. The information about the spectrum is kept in the parameter file with the extension ".3D.param". The example below

```
Version ....................... 1
Number of dimensions .......... 2
16 or 8 bit file type ......... 16
Spectrometer frequency in w1 .. 60.811001
Spectrometer frequency in w2 .. 600.138000
Spectral sweep width in w1 .... 34.840000
Spectral sweep width in w2 .... 6.719000
Maximum chemical shift in w1 .. 135.384872
Maximum chemical shift in w2 .. 11.538800
```

11

```
Size of spectrum in w1 ........ 128
Size of spectrum in w2 ....... 512
Submatrix size in w1 .......... 256
Submatrix size in w2 .......... 256
Permutation for w1 ............ 2
Permutation for w2 ............ 1
Folding in w1 ................. RSH
Folding in w2 ................. RSH
Type of spectrum ............. C
```

is from a 2D spectrum. The intensities are stored in the 16 bit format. Spectrometer frequencies and spectral sweep widths for both dimensions are given in ppm. The maximum chemical shifts give the ppm frequency of the lower left corner of the spectrum. The total size of the spectrum is given in data points. The submatrix size gives the size of the blocks used to store the intensities in order to allow fast access to small parts of the spectrum. The permutation defines the sequential order in which the intensities are stored in the data file. The folding has to be set to either RSH, for Ruben States Haberkorn (Marion, D., Ikura, M., Tschudin, R. & Bax. A. (1989) J. Magn. Reson. 85, 393-399), or TPPI, for time proportional phase increment (Marion, D. and Wüthrich, K. (1983) Biochem. Biophys. Res. Comm. 113, 967-974). The folding type is used to fold the peaks into the spectrum when loading a peak list. The type of spectrum is not used by the program.

To set the calibration of a spectrum the command "Calibration [ca]" is used. Since the peak positions are stored in ppm it is not recommended to work with uncalibrated spectra. The way the parameter file is interpreted is influenced by the value of the resource "XEasy*traditional_calibration".

A spectrum can be turned around without changing the data file by editing the parameter file. Using the following parameter file defines the data file from the above example to be from a three dimensional spectrum extending only one data point in the third dimension. An arbitrary calibration is used in w1:

```
Version ...................... 1
Number of dimensions .......... 3
16 or 8 bit file type ......... 16
Spectrometer frequency in w1 .. 500.000000
Spectrometer frequency in w2 .. 60.811001
Spectrometer frequency in w3 .. 600.138000
Spectral sweep width in w1 .... 10.000000
Spectral sweep width in w2 .... 34.840000
Spectral sweep width in w3 .... 6.719000
Maximum chemical shift in w1 .. 500.000000
Maximum chemical shift in w2 .. 135.384872
Maximum chemical shift in w3 .. 11.538800
Size of spectrum in w1 ........ 1
Size of spectrum in w2 ........ 128
Size of spectrum in w3 ........ 512
Submatrix size in w1 .......... 1
Submatrix size in w2 .......... 256
Submatrix size in w3 .......... 256
Permutation for w1 ............ 3
Permutation for w2 ............ 2
Permutation for w3 ............ 1
Folding in w1 ................. RSH
Folding in w2 ................. RSH
Folding in w3 ................. RSH
Type of spectrum ............. C
```

The different possibilities to generate spectra files in the XEASY format are given below. A description of conversion programs between different formats is given under external programs .

- directly from PROSA: using the spectrum processing program PROSA (P. Güntert, V. Dötsch, G. Wider and K. Wüthrich, J. Biomol. NMR, 2 (1992) 619-629) data suitable for XEASY may be written out with the command write easy8 [filename] or write easy16 [filename]
- converting a Bruker smx file into a 2D spectrum using the filegen2d program.
- converting a series of Bruker smx files into one 3D spectrum using the filegen3d program.
- copying Bruker smx files from the X32 to the workstation and converting them into 2D or 3D spectra. This can be done with the shell scripts cpx32 and cpx32_16bit.
- converting a spectrum from the old EASY format into the new XEASY format with the program filecon2d.

In addition to the file format for XEASY, two outdated spectral formats from the EASY and EASY3D

12

programs exist. The format from the EASY program can no longer be read by the XEASY program. These files have to be converted by the filecon2d program. Files used in EASY3D can still be read by the XEASY program. Since these parameter files contain binary information they may not be edited.

3.2 Zoom Regions

Many routines of the XEASY program display selected regions of a spectrum. One dimensional cross sections of the spectrum are displayed as plots of intensity vs. data points in the slice window (see section 12.8 on page 95). Two dimensional regions of the spectrum may be shown in the main window as contour plots or as intensity plots. The intensity plot maps the intensities onto a color scale. The contour plot draws lines of equal intensity. Following commands may be used to change between the different display modes:

- "Restore intensity[ri]"
- "Int. and contour [ic]"
- "Contour plot [cp]"

3.2.1 Selecting Zoom Regions

XEASY provides routines to define and select regions from the spectrum for displaying. The interactive routines for zooming using no information from peaks, assignments or strip sequences are described below. In addition, there exist a number of commands which select zoom regions based on picked peaks or on strip sequences. These are described in the corresponding chapters about peak lists (page 23) and strip sequences (page 36).

To select an arbitrary region in the spectrum and also to restore the full spectrum the command

"Permutation [pm]"

can be used. This command as well as the "New spectrum [ns]" command define the view onto the spectrum, that is which dimension will be displayed horizontally (in x direction) and which vertically (in y direction). Many routines within XEASY refer to this view.

The following commands allow the definition of regions to be displayed by manually selecting their boundaries within the currently displayed zoom region:

- "Zoom one region [zo]".
- "Zoom multiple [zm]".

In addition the two letter code "mo" may be used to select a region in the overview window. To modify the size or position of the displayed regions the following commands are available:

- "Zoom transpose [zt]".
- "Enlarge vertically [ev]".
- "Enlarge horizontally [eh]".
- "Goto left [gl]".
- "Goto right [gr]".
- "Go up [gu]".
- "Go down [gd]".
- "Forward plane [fp]".
- "Backward plane [bp]".
- "View area [va]".
- "Goto plane [gp]".

The zoom commands to move or resize regions take as a parameter the zoom factor which can be defined using the command "Zoom factor [zf]".

13

3.2.2 Comparing Zoom Regions

To compare different spectra, it is possible to display the contour plot of one spectrum on top of the intensity plot of an other spectrum. The command

"Alternative spectrum [as]".

may be used to define additional spectra for which the commands

"Replace contour [rc]".
"Replace spectrum [rs]".

select the ones to display as contour and intensity plots. The spectra currently used for the intensity plot and the contour plot are indicated in upper left corner of the main window. The upper name gives the spectrum used for intensity plots the lower the one for the contour plot. The routines in XEASY which access the intensities of a spectrum use always the spectrum of the intensity plot. To compare different regions the command

"Zoom alignment [za]".

is used to get several aligned regions. The large cursor

"Draw cursor [dc]".

allows the precise comparison of the positions of peaks within these aligned regions. If displaying the intensity plots of different spectra together it might be necessary to adjust the scale of color vs. intensities of one or more of them. This is possible with the command

"Scale display [sd]".

3.2.3 Zoom Region History

A history of the displayed zoom regions is kept and is used to call back previous displays. The command

"Restore zoom [rz]".

brings back the last stored display. Commands that can easily be reversed (e.g. the commands to move around a zoom region) are not stored in the zoom region history and can not be restored with the [rz] command. The two-letter codes "zb", "zf" and "zo" can be used in the overview window to select any zoom from the zoom history.

Zooms may also be written to a file and recovered again. The commands are

"Load zoom [lz]".
"Write zoom [wz]".

These files have the extension .zoom and contain the information about the zoom regions. The data in the zoomed regions are not stored, but rather the ppm coordinates of the regions in the spectrum are saved. Whenever such a file is loaded, the display will be accordingly updated.

3.3 Phase Correction

The program XEASY can be used to determine phase correction parameters interactively on the basis of 1D cross-sections that are displayed in the slice window (see "Slice window [sw]").

With the conventions used by the program PROSA (P. Güntert, V. Dötsch, G. Wider and K. Wüthrich, J.

$$\hat{s}_0, \ldots, \hat{s}_{n-1}$$

14

Biomol. NMR, 2 (1992) 619-629), the phase-corrected spectrum　　　　　　is related to the original spectrum $s_0, \cdots, s_{n-1}$ by

$$\hat{s}_k = s_k e^{-i\left(\phi_0 + \phi_1 \frac{k}{n+1}\right)}$$

                                                3.1.B

for the data points 0, ..., n-1 in every row along the dimension of interest. is the constant, the linear (or first order) phase correction parameter. The program XEASY allows real-time change of the phase correction parameters and and display of the corresponding, phase-corrected 1D cross-sections, and thus provides a convenient environment for the accurate interactive selection of phase correction parameters. The phase correction parameters found with the help of XEASY can subsequently be used in data processing programs such as PROSA to perform the phase correction on the complete multidimensional spectrum.

To use the interactive phase correction routine in XEASY, first two separate spectrum files containing the real and imaginary parts of the complex spectrum are prepared (e.g., with the program PROSA), then the real part of the spectrum is read into XEASY with the "New spectrum [ns]" command and the imaginary part of the spectrum is defined as alternative spectrum withe the "Alternative spectrum [as]" command. Next the user selects and displays suitable rows (or columns) in the slice window (see "Slice window [sw]") and clicks the Phase button in the slice window. Phase correction will be performed for the spectrum in the dimension of the current slice (i.e., the slice for which the Current button on top of the slice window is activated). The alternative spectrum window pops up and the user selects the file containing the imaginary part of the spectrum. The constant and linear phase correction parameters can now be adjusted using the keyboard according to table 3.3.

The linear phase correction parameter is changed such that the phase of the current slice at the cursor position in the slice window remains constant. The actual values of the phase correction parameters are monitored in the lower left corner of the slice window.

The user deactivates the phase correction mode by clicking the Phase button again. The 1D cross-sections of the original spectrum (i.e., the real part loaded with the [ns] command) are displayed again.

3.4 Assignments

Assignments represent the main result of the work with the program XEASY. Three lists, namely the peak list, the atom list and the fragment list contain the information about the assignments. The peak list contains the coordinates of the picked peaks used to assign the spectrum. The atom list contains the names and frequencies of possible resonances. They define the possible assignments for each dimension of a peak. For homonuclear, single quantum, proton spectra the atom list contains the names and frequencies of all protons in the molecule. In this case it corresponds to the proton list used in the programs EASY and EASY3D. The fragment list contains the names of the fragments which are used at the different stages of the assignment process. In early stages of the assignment, spin systems independent of the primary sequence of the molecule are used. Then increasingly more of them become mapped onto the primary sequence of the protein until finally the residues of the molecule under investigation are used. In this final stage the fragment list corresponds to the sequence list of the programs EASY and EASY3D.

In addition to these three lists, XEASY uses a library file defining the atoms and pseudo atoms for each fragment type. The following chapters provide first, a detailed description of the three lists and the library file, and then a description of how they can be used in conjunction in order to proceed through the different stages of the assignment process.

3.4.1 Peak List

A peak list contains entries for peaks picked in the spectrum. The following paragraphs present in detail the peak list file format, peak picking, assigning peaks, editing peaks, displaying relevant information

15

from a peak list, folding, and how to treat different dimensionality of peak lists and spectra.

A peak list contains an entry for each peak. The fields describing a peak are listed in Table 3.4.1.A. In a 2D spectrum the fields for the w3 and w4 dimension and in a 3D spectrum those for w4 are not used. The peak numbers in a peaklist must be unique but not necessarily continuous.

Table 3.4.1.A Peak Fields

```
Dim.     Fields              Description

         peak number         unique number identifying the peak
         colour              colour in the range [1,6] used for displaying the peaks
         volume              volume of the peak
         volume error        volume error in percent
         integration method
                             method used for the integration: d, r, e, m, a, -
         comment             user defined comment
         possible ass.       data structure containing possible assignments

         shift               folded w1 position in ppm
w1       fold                number of times peak is folded in w1
         atom number         assignment in w1: reference into the atom list

         shift               folded w2 position in ppm
w2       fold                number of times peak is folded in w2
         atom number         assignment in w2: reference into the atom list

         shift               folded w3 position in ppm
w3       fold                number of times peak is folded in w3
         atom number         assignment in w3: reference into the atom list

         shift               folded w4 position in ppm
w4       fold                number of times peak is folded in w4
         atom number         assignment in w4: reference into the atom list
```

The peak list is stored in a peak list file with extension .peaks. It can be read by the program ASNO (P. Güntert et al., 1993) which uses a peak list, an atom list and selected structure coordinate files to generate possible assignments for NOESY cross peaks which can be loaded into XEASY. The program CALIBA (P. Güntert et al., J. Mol. Biol. (1991) 217, 517-530) translates the peak lists containing integrated peaks into distance constraints which can be used by the program DIANA (P. Güntert et al., J. Mol. Biol. (1991) 217, 517-530). An example of the first few lines of a two dimensional peak list is given below:

```
# Number of dimensions 2
  11  7.289  10.169 1 ?        2.048e+03  0.00e+00 -   0  126  128  0
      #  first peak
  12  7.119   9.413 1 ?        1.280e+02  0.00e+00 -   0  517  506  0
   3  7.106   7.497 1 ?        4.096e+03  0.00e+00 -   0  129  130  0
   4  7.228   7.411 1 ?        4.096e+03  0.00e+00 -   0  131  127  0
   5  7.106   7.411 1 ?        5.120e+02  0.00e+00 -   0  327  328  0
   6  6.838   7.094 1 ?        8.192e+03  0.00e+00 -   0  489  488  0
```

The number on the first line after the hash "#" indicates the dimensionality of the peak list. Subsequent lines, starting with a number, contain the fields for one peak. Additional lines starting with a hash "#" are comments for the peak on the line above. The first field for each peak is the peak number, followed by: the unfolded chemical shift coordinates in ppm in w1 and w2 (more numbers are listed for higher dimensional spectra), the color code (a number from 1 to 6), the user defined type of the spectrum where the peak is observed, the peak volume, the uncertainty of the volume in percent, the integration method ("d" for Denk integration, "r" for rectangular integration, "e" for elliptical integration, "m" for maximum integration, "a" for automatic integration, "-" for not integrated), an unused number, the assignments in w1 and w2 is given by the two following atom numbers (more numbers are listed for higher dimensional spectra), the last number is not used. The commands to load or write peak lists are "Load peaklist [lp]" and "Write peaklist [wp]".

Loading a peak list with different dimensionality than the spectrum is possible. If the peak list has more dimensions than the spectrum the additional dimensions are just ignored when working with the peak list. When loading a peak list with less dimensions than the spectrum, it has to be specified whether the needed additional dimensions are copied from the available dimensions or set to a fixed ppm value. For example, the 2D peak list from a [1H,15N]-HMQC spectrum can be loaded onto a 3D 15N-correlated [1H,1H]-NOESY by copying the 1H chemical shift of the 2D list to both 1H dimensions in the 3D

16

spectrum. The resulting peaks are positioned on the diagonal of the 1H planes in the 3D spectrum.

Peaks may be picked either automatically or manually. For automatic peak picking of anti phase peaks in 2D spectra the command "Anti-phase peaks [an]" is used. For automatic picking of inphase peaks the external program pick described on page 105 or the command "In-phase peak picking [in]" can be used.

Issuing the command "Peak picking [pp]" enters the manual peak picking mode. By selecting a position on the screen a peak placed at the corresponding position of the spectrum is added to the peak list. In 3D and 4D spectra the added peak is picked at the position of the displayed 2D region.

The peak list file contains always the unfolded chemical shifts. When loading a peak list onto a spectrum all peaks are folded into the spectrum. The folding information is stored in the field fold. When writing out the peak list, this field is used to back calculate the unfolded peak positions. The information of the unfolded chemical shifts of a peak can in this manner be transferred from a spectrum with a large sweep width (i.e. 2D HMQC experiment) to spectra with smaller sweep widths (i.e. 3D 15N correlated [1H,1H] spectra). When picking additional peaks, one wants to retain the folding information from peaks already present in the spectrum. This is possible with the command "Copy and move peaks [cm]", which allows to copy the folding information and the assignment from an existing peak to new peaks. Other methods to set the folding of a peak are to enter the unfolded chemical shift into the w field of the peak editing window or to use the "Set selected peak entries [pe]", command which can be used to set or reset any field in the peak list.

Once peaks are picked, they may be assigned to atoms present in the atom list. This is done either for a single peak using the command "Assign peak [ap]" or for all selected peaks within a region using the command "Assign peaks in one region [ar]".

For example, all peaks in the fingerprint region of a homonuclear proton COSY spectrum can be assigned to the HN and Ha atoms of a set of spin systems, or all peaks in a [1H,15N]-HMQC spectrum may be assigned to the N and HN atoms of back bone fragments.

When working with a peak list it is crucial to display information relevant for the current assignment task. First, often only a subset of peaks is relevant (e.g. only the peaks assigned to a certain fragment). The commands "Select peak [sp]" and "Display all peaks [da]" provide different criteria for selecting peaks to display. In 3D and 4D spectra, usually peaks of interest are only those close to the displayed 2D zoom region. The option "Maximal distance in planes" in the default window controls the number of adjacent planes from which to display peaks.

Second, peaks with different properties (i.e. assigned peaks versus unassigned peaks) can be displayed with different colors or peak shapes. The options: "Peak color determined by", "Peak cross displayed is", "Lineshape displayed for" and "Coloring interval" in the default window, control the peak colors.

Third, peaks may also be labelled with their assignment or the volume. The command "Peak data window [pw]" pops up the peak data window in which the peak labels may be defined.

Fourth, peaks may be used to zoom interesting regions out of a spectrum. In the simplest case of the command "View peak [vp]" a region is zoomed that contains a specified peak. The command "Zoom peak [zp]" displays a small region around each of the peaks selected with the "Select peak [sp]" command. The command "View reference peak [vr]" displays for each peak with the same assignment as the selected one, a slice in the slice window. An other powerful method to select relevant regions out of a spectrum is the strip list. It is described in the section 6.0 on page 56.

It is possible to work with several peak lists together. When a peak list is loaded it is kept in memory until it is removed from memory with the

"Erase peak list [ep]"

command. To select a different peak list for displaying with a spectrum the command

17

"Exchange peak list [xc]"

is used. This command does not change the dimensionality or the folding of the loaded peak list. To adapt the folding of the peaks to a certain spectrum, to change the dimensionality or to permute the dimensions the command "Adapt peaks to spectrum [ad]" is used.

Several commands exist to edit the peak list entries. they are:

- "Delete peaks one by one [dp]"
- "Remove peaks in one region [rp]"
- "Move peaks one by one [mp]"
- "Fetch peak into plane [pl]"
- "Move reference peak [mr]"
- "Move peak list [ml]"
- "Editing peaks [ed]"
- "Set selected peak entries [pe]"

Worthwhile mentioning is the "Move reference peak [mr]" command. It allows one to move together all peaks with the same assignment. This helps when adjusting a peak list picked and assigned in one spectrum to a related, but slightly different spectrum. Two examples are pH titrations and the use of a TOCSY peak list to identify the intra-residual peaks in a NOESY spectrum.

To check or to list the assignments the commands "List peak entries [le]" and "Report ass. stat.[ra]" are used. They check for duplicated assignments, i.e. two peaks that have the same assignment, and for large chemical shift errors, i.e. peaks at different frequencies which are assigned to the same atom. Both informations are useful to identify wrong assignments.

3.4.2 Atom List

The atom list contains the names and frequencies of resonances. They define the possible assignments for each dimension of a peak.

Table 3.4.2.A - Atom Fields

```
Fields            Description

atom number       unique number identifying the atom
shift             mean chemical shift in ppm
shift error       deviation of the assigned peaks from the mean value
name              atom name
fragment number   number of the fragment to which the atom belongs
lineshapes        data structure containing the reference lineshapes
```

The fields listed in Table 3.4.2.A constitute an atom entry. They can be modified in the peak editing window. The atom numbers, used to reference the atoms, must be unique but not necessarily continuous. The number -9999 is reserved to denote invalid entries. The average chemical shift and the shift error can be calculated from the assigned peaks. The command is "Average chem. shift [ac]".

If the chemical shift is not defined it is set to the value 999.000.

A new atom list is generated each time when a fragment list is loaded. For each fragment the corresponding atoms are looked up in the fragment library file and added to the list. New atom entries are added to the atom list if a non existing atom is used with the "Assign peak [ap]" command, if the fragment type is changed in the peak editing window, or with the "Add new fragment [af]" command.

The atom list file has the extension ".prot" originating from the old EASY format. The following line is taken from such a file:

```
32   4.370  0.004   HA  2
```

The first number is the atom number, followed by its mean chemical shift and the deviation from the

mean value. The atom name and the fragment number follow. The commands to read or write an atom list are "Load atoms (chem. shift) [lc]" and "Write atoms (chem. shift) [wc]".

3.4.3 Fragment List

The fragment list contains the fragments currently used for the assignment. Depending on the stage of the assignment, the fragments are either spin systems or residues of the molecule. In the latter case, the fragment list corresponds to the sequence list used in the programs EASY and EASY3D.

Table 3.4.3.A Fragment Fields

```
Fields                  Description

fragment number    unique number identifying the fragment
fragment type      name of the fragment
mapping number     used to map spin system fragments to a residue fragments
comment            used to store possible spin system types or possible sequential
neighbours
```

The fields listed in Table 3.4.3.A constitute a fragment. The fragment numbers, which are used in the atom list to reference fragments, must be unique but not necessarily continuous. They are referenced in the atom list. The number -9999 is reserved to denote invalid entries. The fragment type must be defined in the fragment library. When starting a resonance assignment it is usually set to "SS" meaning that there is no information about the specific spin system type. In later stages it may be changed, for example to "ASP", using the peak editing window. The command "Add new fragment [af]" will add additional fragments to the list.

The mapping number defines a mapping from one set of fragments (e.g., spin systems) to another set (e.g., amino acid residues). When working with spin system fragments the mapping number is the residue number of the spin system. For example: a fragment with number 207, type GLY and mapping number 55 describes spin system 207 which corresponds probably to GLY 55. The mapping number -1 is reserved to denote invalid mappings. The comment can, for example, be used to store the fragment numbers of the sequentially neighboring spin system. When using the automatic sequential assignment routine "Sequential assignment [op]" the fragment comment can be used to indicate for each spin system the probable amino acid types.

The fragment list is stored in a file with extension ".seq". If the first line starts with a hash "#", it is treated as a comment. The subsequent lines list the fragment types with one line per fragment. For example, the following file defines a tripeptide:

```
# tripeptide
ASP 0
GLY 5 203
LYS 7 209 "-1: 203"
```

The number, if given, after the fragment name denotes the fragment number. If there is no number given for a fragment, its number will be assigned according to the last fragment number plus one. If no number is given for the first fragment, it will be set to one. The optional second number is the mapping number. An entry "-1" for the mapping number indicates that it is not defined. The optional text within the quotes is the comment. If the comment is present the mapping number must also be listed.

The command to read or write a sequence file is "Load sequence [ls]" and "Write sequence [ws]".

3.4.4 Fragment Library

The fragment library defines the different fragment types. All the atoms constituting a given fragment type are listed. The format is the one of the program DIANA (P. Güntert et al., J. Mol. Biol. (1991) 217, 517-530). Amino acid fragments, nucleic acid fragments, a general spin system fragment and a general amino acid backbone fragment are currently defined. The extension is ".lib". The default library, defined by the environmental variable XEASY_LIB is loaded when starting up XEASY. To load a different library the command "Load library [ll]" is used.

3.4.5 From Spin System Assignments to Residue Assignments

A reasonable setup for starting out with resonance assignments is to define a set of general spin systems in the fragment list and generate the corresponding atom list. The first few lines of such a fragment list are given below:

```
SS 201
SS 202
SS 203
SS 204
SS 205
SS 206
SS 207
SS 208
SS 209
...
```

Some lines extracted from the corresponding atom list, which is generated by loading the fragment list into XEASY are given as well:

```
 1 999.000 0.000 N     201
 2 999.000 0.000 HN    201
 3 999.000 0.000 CA    201
 4 999.000 0.000 HA    201
 5 999.000 0.000 QA    201
 6 999.000 0.000 CB    201
 7 999.000 0.000 HB2   201
 8 999.000 0.000 HB3   201
 9 999.000 0.000 QB    201
...
34 999.000 0.000 N     202
35 999.000 0.000 HN    202
36 999.000 0.000 CA    202
37 999.000 0.000 HA    202
38 999.000 0.000 QA    202
39 999.000 0.000 CB    202
40 999.000 0.000 HB2   202
41 999.000 0.000 HB3   202
42 999.000 0.000 QB    202
...
```

With these two lists, peaks may be grouped by assigning them to the same fragment and further classified by assigning them to certain atoms within the fragments. For example when working with heteronuclear spectra, each peak in a [1H,15N] HMQC experiment can be assigned to the N and HN atoms of different fragments. Similarly when working with homonuclear spectra, each COSY cross peak in the fingerprint region may be assigned to the Ha and HN atoms of a different fragment. Peaks in the COSY or TOCSY spectrum lying on the same amide frequency as the HN Ha peak can then be assigned to the HN Hb2, HN Hg2, ... of the same fragment.

After the resonance assignment has been finished, the fragments of choice are the amino acid residues and the atom list has entries for the atoms in the molecule. The first few lines of the fragment list for the interleukin receptor antagonist are given below:

```
ARG 1
PRO 2
SER 3
GLY 4
ARG 5
LYS 6
```

Some lines extracted from the corresponding atom list with undefined chemical shifts are given below:

```
 1   999.000   0.000   C     1
 2   999.000   0.000   CA    1
 3   999.000   0.000   CB    1
 4   999.000   0.000   CD    1
 5   999.000   0.000   CG    1
 6   999.000   0.000   CZ    1
 7   999.000   0.000   HA    1
 8   999.000   0.000   HB2   1
 9   999.000   0.000   HB3   1
10   999.000   0.000   HD2   1
11   999.000   0.000   HD3   1
```

20

```
12     999.000   0.000    HE        1
13     999.000   0.000    HG2       1
14     999.000   0.000    HG3       1
15     999.000   0.000    HH1       1
16     999.000   0.000    HH21      1
17     999.000   0.000    HH22      1
18     999.000   0.000    HN        1
19     999.000   0.000    N         1
20     999.000   0.000    NE        1
21     999.000   0.000    NH1       1
22     999.000   0.000    NH2       1
23     999.000   0.000    QB        1
24     999.000   0.000    QD        1
25     999.000   0.000    QG        1
26     999.000   0.000    QH2       1
27     999.000   0.000    C         2
28     999.000   0.000    CA        2
29     999.000   0.000    CB        2
30     999.000   0.000    CD        2
31     999.000   0.000    CG        2
32     999.000   0.000    HA        2
33     999.000   0.000    HB2       2
34     999.000   0.000    HB3       2
35     999.000   0.000    HD2       2
36     999.000   0.000    HD3       2
37     999.000   0.000    HG2       2
38     999.000   0.000    HG3       2
39     999.000   0.000    N         2
40     999.000   0.000    QB        2
41     999.000   0.000    QD        2
42     999.000   0.000    QG        2
...
```

These two lists can be used to make assignments of peaks to atoms in the molecule under investigation. For example, with these lists NOESY cross peaks can be assigned and integrated in order to extract distance constraints.

Starting from peaks assigned to spin system atoms, a method is needed to proceed to peaks assigned to atoms of the residues in the molecule. Changing assignments of single peaks is too time consuming. Changing the assignments for all peaks assigned to a given residue together is preferable. This is achieved by leaving the peak list and the atom numbers untouched and changing only the fragment and atom list.

The fragment type field can be changed from a general spin system (SS) into a glycine (GLY). Thereby it is ensured that not only the fragment type in the fragment list is updated but also that all the atoms of the new fragment type will be present in the atom list. In the above example, changing from a spin system fragment to a glycine fragment the Ha1, Ha2 and Qa atoms are inserted at the appropriate position in the atom list. For changing the fragment number (i.e. proceeding from spin system numbers to residue numbers) the mapping number is used. The mapping number is set in the peak editing window. In the case of a spin system fragment it is equal to the residue number of the spin system. The command "Switch fragment/map numbers [sn]" exchanges the fragment number and the mapping number. In order to keep the fragment and the atom list consistent the fragment number is changed in both lists together.

It is important to ensure that a certain fragment number is not used at same time for a spin system fragment and a residue fragment. The first possibility to avoid this, as illustrated in the above example, is to use different sets of numbers for the spin systems and the residues. The second possibility is to first set the mapping of all spin system fragments and only then exchange the fragment and the mapping numbers.

3.4.6 Changing Assignments

Different cases can be distinguished: either a peak is missassigned, a resonance is missassigned or the fragment number or fragment type is missassigned. Each of these cases is discussed separately.

To correct a missassigned peaks the atom list and the fragment list are not changed and only the atom number fields in the peak list are changed in order to reflect the new assignment. This is done with the command "Assign peak [ap]" In contrast, to change a resonance assignment, the peak list and the fragment list should remain unchanged and only the atom list is adapted. This can be achieved by editing the atom fields in the peak editing window. If changing an atom name or fragment number the corresponding fields in the atom list are updated.

To change the fragment type or the fragment number the peak list remains unchanged while the atom list and fragment list are adapted. This is done by editing the fragment and the mapping field in the peak editing window and using the "Switch fragment/map numbers [sn]" command as described in the chapter From Spin System Assignments to Residue Assignments on page 31.

3.5 Possible Resiude Types

Based on observed chemical shifts of a given fragment it is possible to identify likely residue types. Especially when 13C chemical shifts are known discrimination between different amino acid types is readily achieved, e.g., in more than 50% of the case the correct amino acid can be identified. The "Residue Type window [rw]" allows to match a set of identified frquencies to the frequencies expected for different residue types. The expected frequencies and their standard deviation are stored in the fragment library. If for a certain atom no expected frequency is specified in the fragment library the atom will not be considered.

3.6 Possible Assignments

Once the resonance assignments are nearly completed the known frequencies can be calculated using the "Average chem. shift [ac]" command and subsequently can be used to generate possible assignments for peaks. For each dimension of a peak, the atoms resonating at about the same frequency can be identified and their atom numbers stored in a data structure. The command "Possible assign. [pa]" asks for a maximal allowed deviation between the frequencies of the picked peak and the resonances of the atoms. It then generates all the assignment possibilities for all the dimensions of the peaks. In addition to specify the possible assignments for each dimension separately one often wants to include or exclude combinations of assignments. Since a 3D or 4D list of assignment possibilities would be too big, XEASY treats only two dimensional arrays of assignment possibilities. In a 3D or 4D spectrum only combinations of assignments in the two dimensions displayed horizontally and vertically are treated. In each dimension up to 10 possible assignments can be stored, that is, in the displayed plane a maximum of 100 possible combinations. Each time when there are more than ten assignment possibilities in a dimension, a warning message will be prompted and the maximal allowed deviation between the peak and the atoms will be lowered for this peak. To view and edit the assignment possibilities the assignment window is available. It can be popped up by the command "Editing assign. [ea]" In the lower left side of the panel, the horizontal text lines list the possible assignments in the vertical dimension, and the vertical text lines in the upper right part list the possible assignments in the horizontal dimension. The assignments may be toggled using the buttons in the window. To select a single assignment possibility, the corresponding button has to be pushed while the shift key is held down. Once a unique assignment is present in the assignment window the peak list is updated automatically. Another possibility to update the peak list according to the assignment list is provided by the command "Update assign. [ua]" It can also be used to check the two lists for contradictory assignments and to update the assignment list according to the peak list.

Two possibilities exists to reduce the number of proposed assignments. The command "Reduce to intrares. ass. [ia]" toggles all inter-residual assignments off. The command "Reduce using ass. peaks [ru]" uses already assigned peaks to reduce the assignment possibilities of unassigned peaks. Its main application is to include the information of the third and fourth dimension in 3D and 4D spectra.

The assignment possibilities are stored in a file with extension ".assign". This file contains the possible assignments for peaks. The number in the first line of the file indicates the dimension of the peak (2 means two dimensions and 3 means three dimensions). Following for each peak are several lines containing the assignment information. A typical block of data for a 2D peak is given here

```
# 127
3 220 225 294
10 22 82 197 283 293 388 403 432 438 457
1050625 0 0 0
```

The peak number (127 in the example) is given after the hash "#". When loading the assignment list and no peak with the listed peak number is present in the loaded peak list, a warning message will be

prompted and the loading is aborted. Subsequent lines (in the example the second and third line) list the possible assignments for each dimension. On both lines the first number indicates the number of possible atom assignments in this dimension followed by the individual atom numbers. Atom numbers in the file that are missing in the loaded atom list will be ignored. The four numbers on the last line encode the allowed assignment combinations corresponding to the toggle button matrix in the assignment window. The commands to read or write an assignment file are "Load assignment [la]" and "Write assignment [wa]" In later stages of the NOESY cross peak assignment, the number of assignment possibilities can be reduced using preliminary structures. This can be done using the program ASNO (P. Güntert et al., 1993). The program reads in structure coordinates, a peak list and an atom list. It produces all assignment possibilities not contradicting the structures or the frequencies of the atoms. The resulting assignment file from ASNO can also be read with the "Load assignment [la]" command. The first few lines of such a file are given below.

```
Assignment file
    Corresponding peaklist: toc.peaks
    Corresponding coherencelist: tend.prot
    Number of dimensions: 2
    Uncertainties:    0.020   0.020


#     1
          126  128
#     2
           24   22
          173  171
          361  359
          253  251
          316  314
```

The first five lines give the peak list, atom list and chemical shift tolerances in ppm used to produce the assignment file as well as the dimensionality of the assignment list. The entries for each peak start with a hash "#" followed by the peak number. The following lines list the assignment possibilities. Each line defines one nD assignment possibility defined by atom numbers in all dimensions.

3.7 Strips

The concept of strips originates from work with 3D spectra but can also be used for 2D and 4D spectra. Strips are 1D cross sections extending in a second dimension. They lie at the position of an atom (in 2D spectra) or atom group (in 3D or 4D spectra) and contain all the peaks involving this atom group. For example in spectra involving 1H and 15N of the amide group, as the 15N-correlated [1H,1H]-NOESY, 15N-correlated [1H,1H]-TOCSY, HNCA or HN(CO)CA experiments, for each amide group a strip can be defined. In the case of 2D 1H spectra a strip is defined for each proton and in the case of 13C correlated spectra as the HCCH TOCSY a strip can be defined for each proton bound to a carbon. In order to judge the lineshapes of the cross peaks in two dimensions the 1D cross sections are extended into a second dimension. To use strips in 4D spectra the width of the second dimension should be set equal to the sweep width.

For 3D or 4D spectra, working with a set of strips instead of planes reduces the complexity of the assignment steps. Methods to selectively search for certain strips further simplify the assignment procedure. The following paragraphs introduce the strip data structure, methods to define strip lists, display and edit them as well as methods to selectively search for strips that have either a specified intensity pattern or lie at a selected position in the spectrum.

A strip in XEASY has similar fields as a peak, however a strip is not used to hold relevant assignment information. This has to reside in the peak lists. Assignments stored together with a strip are only used to remove duplicated entries, to sort the strip list, and to identify peaks belonging to a strip. The fields defining a strip are given in Table 3.7.0.A.

Table 3.7.0.A Strip Fields

Dim.	Fields	Description
	spectrum	pointer to the spectrum

23

```
          shift            folded w1 position in ppm
w1        fold             number of times strip is folded in w1
          atom number      assignment in w1: reference into the atom list

          shift            folded w2 position in ppm
w2        fold             number of times strip is folded in w2
          atom number      assignment in w2: reference into the atom list

          shift            folded w3 position in ppm
w3        fold             number of times strip is folded in w3
          atom number      assignment in w3: reference into the atom list

          shift            folded w4 position in ppm
w4        fold             number of times strip is folded in w4
          atom number      assignment in w4: reference into the atom list
```

As for the peak list the folded position of the strip is used within the program but the unfolded shifts are written out to the strip list file. The folding information is stored in the field fold. The atom numbers define the assignment of the strip. In contrast to peaks, strips can not be edited. They can only be defined from peaks or loaded from a file.

The strip list file has the extension .strips. An example of the first few lines out of a strip list file is given below:

```
15nnoe-scp  3     0 116.385    8.312    8.311 1854 1853 1853
15nnoe-scp  3     1 118.317    8.170    8.185 1887 1886 1886
15nnoe-scp  3     2 117.788    7.822    7.817 1920 1919 1919
15nnoe-scp  3     3 115.097    6.957    6.957 1953 1952 1952
15nnoe-scp  3     4 112.377    8.099    8.099 1986 1985 1985
15nnoe-scp  3     5 112.319    7.891    7.891 2019 2018 2018
15nnoe-scp  3     6 112.020    6.961    6.960 2052 2051 2051
15nnoe-scp  3     7 125.553    8.410    8.410 2085 2084 2084
15nnoe-scp  3     8 111.984    8.932    8.932 2118 2117 2117
15nnoe-scp  3     9 115.893    7.997    7.963 2151 2150 2150
15nnoe-scp  3    10 106.044    7.482    7.478    0    0    0
15nnoe-scp  3    11 117.601    7.746    7.733 2184 2183 2183
15nnoe-scp  3    12 121.138    7.431    7.431 2217 2216 2216
15nnoe-scp  3    13 119.677    8.409    8.409 2250 2249 2249
```

For each strip there is one line. First the name of the spectrum in which the strip was defined is indicated followed by the dimensionality of this spectrum. When loading a strip sequence, XEASY tries to find the indicated spectrum. If this is not present, a spectrum with the same dimensionality is taken - if possible the currently displayed one. The next number, the strip number, is currently not used. Then come the chemical shifts in all the dimensions and at the end the atom numbers specifying the assignment in each dimension. The commands to load and write strip lists are: "Load strip list [sl]" "Write strip list [ss]" A strip list can be build up from peaks in two different ways. First, all peaks selected with the "Select peak [sp]" command can be appended to the strip list using the command "Strip sequence [se]" New strips are generated at the positions of each of the selected peaks. The assignments of the new strips are taken from the peaks. The newly generated strips are appended to the current strip list. The assignment of the strips can be used to sort them or to delete strips that have the same assignment. Second, with the commands "Append strip [sa]" and "Insert before reference [ib]" a single peak can be selected to define a strip, which is either appended or inserted into the strip list.

Once a strip list is defined it can be displayed. In order to be able to work with strips efficiently only about ten strips should be displayed on the screen at once. The selection of strips to be displayed and the number of strips to display on one screen are specified with the command "Goto strips [gs]". It displays one screen full of strips from the current strip list. With this command the width of the strips in the horizontal dimension is also defined. The height of the displayed strips will be equal to the height of the zoomed region where the command was issued.

To move around in the strip list the commands "Forward strips [fs]" and "Backward strips [bs]" can be used. They move to the next or previous screen of strips. The command "Strip find [sf]" allows a fragment number to be specified. It then displays a screen of strips including the first strip assigned to this fragment. The whole strip list or single strips can be removed from memory using the commands "Erase strip list [es]" and "Remove one strip [ro]".

In many of the assignment tasks one starts with a strip and searches for one or several related strips. The

24

command "Hold strip [sh]" fixes one or several strips. These fixed strips are thereafter shown as reference always on the left side of the screen. They can only be removed with the "Release strip [sr]" command.

In addition to defining and displaying strips methods are necessary to search for relevant strips. Two such methods are implemented in XEASY. The first searches for strips at a selected position. Possible applications of this method include assigning NOESY cross peaks in heteronuclear correlated 3D NOESY spectra, spin systems assignment using the HCCH TOCSY, sequential assignment using the 15N correlated NOESY or the HNCA and HN(CO)CA experiments. For a more detailed description of how to work with such spectra, refer to the section 16.0 on page 109.

The second method searches for strips with a selected intensity pattern. The positions of the strip can in addition be used to narrow down the set of strips for which the intensity patterns are compared. The pattern are derived from observed intensities. They can be modified using peaks with the same assignment as the reference strip. The intensities at the positions of these peaks are multiplied by a user specified factor (refer to the detailed description below). Since the similarity of the strips are determined by calculating the correlation coefficient between strip intensity patterns the method is called the spectral correlation method (Bartels, C and Wüthrich, K. (1994) J. Biomol. NMR, ...). Its main application is the sequential assignment of 15N-correlated [1H,1H]-NOESY spectra: sequentially neighboring spin systems show similar cross peak patterns due to their spacial proximity. The frequencies of the sequential cross peaks can be identified using the peaks from the 15N-correlated [1H,1H]-TOCSY (see also section 16.0 on page 109).

In order to work with the method to search for strips at a selected position, reference strips have first to be defined. The command to define these reference strips is "Define reference strips [rd]". It copies the current strip list into the reference list. After this command has been issued the current strip list can be altered without affecting the search for close strips. The command "Compare close strips [pc]" can then be used to search for close strips. It allows to select a frequency in the vertical dimension of a zoomed region. It then displays a screen of strips. The first strip comes from the zoomed region at the position where the frequency was selected. The following strips are from the reference list. They are sorted by increasing distance from the selected frequency. To view the strips further away from the selected position the commands "Forward comparison [fc]" and "Backward comparison [bc]" are provided. Again the first displayed strip is the one where the frequency was selected.

The next few paragraphs describe the correlation method followed by instructions of how to apply the method. For a detailed discussion refer to Bartels, C and Wüthrich, K. (1994) J. Biomol. NMR, ...

For the given strip s all the remaining candidate strips k are sorted according to a distance measure d_k such that the strip which is searched gets a low rank. The candidate strips, starting with the lowest ranking ones, are then displayed in order to allow the final assignment to be made interactively. The distance measure d_k is an enhancement of the correlation function introduced for the sequential assignment using 3D 15N-correlated [1H,1H]-NOESY in Bartels and Wüthrich (1994). In contrast to the sequential assignment using 3D 15N-correlated [1H,1H]-NOESY in all other assignment tasks it is possible to identify a frequency f in the given strip s at which there must be a peak in the sought strip k (see Applications section). This is used in the definition of the distance measure:

$$ d_k = \begin{cases} 1 + \left| \mathbf{f} - \mathbf{p}_k \right|, & \text{if } \left| \mathbf{f} - \mathbf{p}_k \right| > \Delta. \\ 1 - \left| \dfrac{\vec{\mathbf{v}}_k \cdot \vec{\mathbf{v}}_e}{\left| \vec{\mathbf{v}}_k \right| \left| \vec{\mathbf{v}}_e \right|} \right|, & \text{if } \left| \mathbf{f} - \mathbf{p}_k \right| \leq \Delta; \end{cases} $$

[1] The condition $\left| \mathbf{f} - \mathbf{p}_k \right| > \Delta$ discriminates those strips that are unlikely to correspond to the sought strip, since p_k, their position in the vertical dimension, differs more than the user specified parameter Δ from the expected frequency f. Δ is usually set to a value larger than the expected error in the determination of the frequency f or the peak positions p_k. The other strips with $\left| \mathbf{f} - \mathbf{p}_k \right| \leq \Delta$ are possible candidates for the sought strip. They are further sorted according to the correlation function (Bartels and Wüthrich, 1994) which expresses the similarity

25

of the peak pattern $\vec{v}_k$ observed in strip k to the peak pattern $\vec{v}_e$ expected for the sought strip and derived from the peak pattern observed in the given strip s. The peak patterns v = (i1,i2, . . ., in) are n-dimensional vectors with n equal to the number of data points along the vertical dimension which are derived from the experimentally observed intensities iex according to

$$i_j = A_m \frac{i_j^{ex}}{|i_m^{ex}| + i_b} \qquad (p_l \leq j < p_h),$$

[2] Here i_m^{ex} is the experimental intensity at the local maximum m of the absolute intensities, pl and ph are the positions of the two adjacent local minima of the absolute intensities and ib is a parameter usually set to 15 times the standard deviation of the noise. The constant Am - usually set to 1 - gives a weight to every local maximum. For particular assignment tasks, as for example sequential assignment using 3D 15N-correlated [1H,1H]-NOESY spectra, it can be used to emphasize or suppress subsets of peaks (Bartels and Wüthrich, 1994).

Note that setting Δ in Eq. [1] to infinity, allows to handle cases where their is no peak in the sought strip whose frequency f can be identified in the given strip, e.g., sequential assignment using 3D 15N-correlated [1H,1H]-NOESY and setting Δ to 0.0 ppm causes the candidate strips to be sorted only by the deviation of the strip position pk from the identified frequency f.

To apply the spectral correlation method the noise level must first be defined. This is done using the contour plot command "Contour plot [cp]" Currently displayed peaks with the same assignment as the strip define the positions Am used to modify the peak patterns. The remaining parameters necessary to calculate the correlation coefficients are defined using the command "Define correlation [cd]" It copies the current strip list to the correlation strip list and calculates the correlation coefficients between pairs of strips. Again the current strip list can be modified afterwards without affecting the correlation method. To search for correlated strips the command "Compare correlated strips [cc]" is available. It asks for a reference strip and a frequency f to be selected and then displays a screen full of strips for comparison. The first strip displayed is the selected reference strip. The following ones come from the correlation strip list and are sorted by decreasing correlation to the pattern of the reference strip. To display the previous or next screen of correlated strips the commands "Forward comparison [fc]" and "Backward comparison [bc]" are used. Again the first displayed strip is the reference strip.

The command "Sequential assignment [op]" uses the spectral correlation coefficients for determining the sequential assignment of a protein. The method combines simulated annealing with an algorithm described by R. Bernstein, C. Cieslar, A. Ross, H. Oschkinat, J. Freund and T. A. Holak. in J. Biomol. NMR, 3 (1993) 245-251. Already known sequential assignments and assignments of spin systems to possible spin system types can be provided to the routine.

3.8 Integrals

The final step in spectra interpretation for the structure determination of proteins is the integration of the NOESY cross peaks. The peak integrals are in a first approximation proportional to rij-6 where rij denotes the inter-proton distance between protons i and j. Because of this inverse sixth power relationship the volume accuracies are not critical - an error of 100% in volume leads to only a 12% change in inter-proton distance.

The problem remaining when trying to integrate cross peaks is overlap. Proteins of even moderate size will have upwards of several thousand cross-peaks many of which will overlap with others. If one of the overlapping components is weak the evaluation of its volume can be difficult - even getting the order of magnitude correct may be impossible by conventional techniques. In hetero-nuclear correlated 3D spectra the NOESY cross peaks are split up in a third dimension. This significantly reduces the problem of overlapping peaks and therefore integration.

The approach used in XEASY for 2D spectra is line-shape integration. This method stems from the fact

26

that line-shapes of peaks with the same assignment have the same lineshape. Distortion may however arise from zero quantum effects. Cross peaks which overlap are integrated by taking the w1 and w2 line shapes for each peak and then deconvoluting the peak cluster with the lineshapes into volumes and volume errors. This technique was proposed by Denk, W. Baumann, R. & Wagner, G. (1986) J. Magn. Reson., 67, 386-390. The approach consists of three steps: first reference line-shapes along the x and y direction are determined for each resonance in the 2D spectrum. Next the peaks are grouped together into clusters of overlapping peaks. Here, two peaks are said to overlap if the rectangles defined by their line-shape extents intersect. Finally the volume of each peak in the cluster is determined by a linear least squares fit of the peak shapes constructed from the reference line-shapes to the experimental data points in the spectrum.

In mathematical terms the problem is one of adjusting the volumes Vp to minimize the following expression:

$$\sum_{(\omega_1,\omega_2)} \left[S(\omega_1,\omega_2) - \sum_{p=1}^{m} V_p L_1^p(\omega_1) L_2^p(\omega_2) \right]^2$$

3.8.A where S(w1,w2) is the spectral intensity at coordinate (w1,w2), Vp is the volume of peak p, m is the number of peaks in the cluster and Li is the reference line-shape for the wi resonance of peak p. This linear least squares problem is solved using standard methods (Press W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T. (1986) Numerical Recipes, the Art of Scientific Computing. Cambridge University Press). An uncertainty can be obtained for each peak volume by calculating the square-root of the above function over the peak region (not the total cluster). This can then be expressed as a percentage of the calculated volume.

For line shape integration, the command "Line-shape integration [li]" is used. Before it is applied, the lineshapes for each atom have to be defined and edited with the following commands

- "Select x reference line-shape [sx]"
- "Select y reference line-shape [sy]"
- "Select x & y ref. line-shape [sb]"
- "Find reference automatically [fr]"
- "Reset ref. line-shape list [re]"
- "Remove reference line-shape [rm]"

Lineshapes are saved into a file with extension ".ref". An example is given below:

```
# Number of dimensions 2
  1 -9999    0.000    0.000 -9999    0.000    0.000
  2   774    4.211    4.266   155    4.206    4.258
  3 -9999    0.000    0.000 -9999    0.000    0.000
  4 -9999    0.000    0.000 -9999    0.000    0.000
  5   155    3.057    3.167  1452    3.063    3.153
```

The first line gives the dimensionality of the reference lineshape list. Each of the following lines defines for one atom the position in all dimensions where the lineshapes can be read out from the spectrum. The first number on each line is the atom number. The rest of the line is subdivided into groups of three numbers for each dimension. The first of which is the number of the reference peak, the second and third number give the lower and upper bound of the lineshape in ppm. The entry for atom 2 is illustrated on the

27
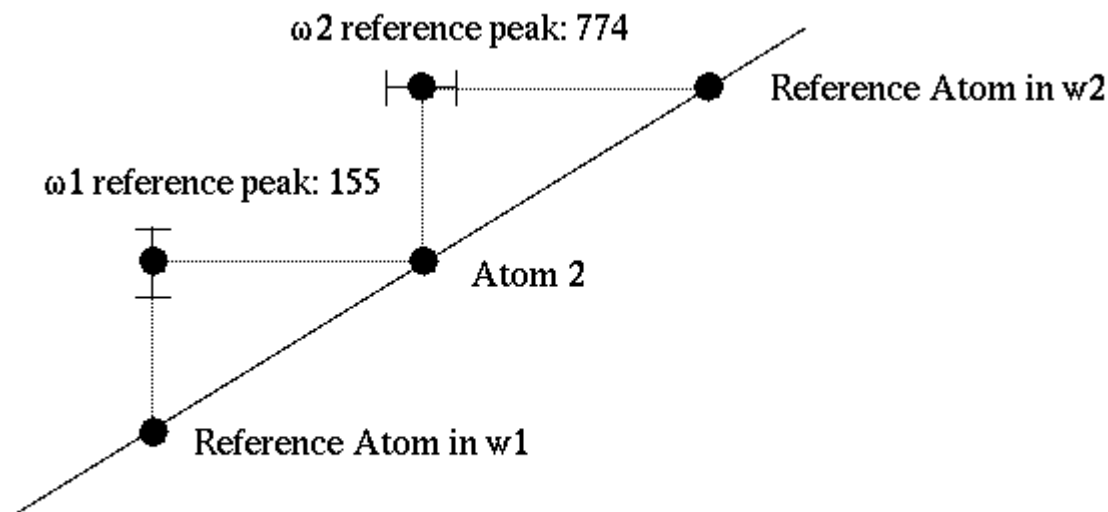
Figure 3.6a: line-shape definition for atom 2

figure below:

3.8.A The peak number "-9999" means that the w1 or w2 line-shape is not defined. The commands to load or save a line-shape list are "Load reference list [lr]" and "Write reference list [wr]".

The formats for the line shape lists of the programs EASY and EASY3D are different. Instead of specifying a reference peak, as in EASY and EASY3D, in XEASY a reference atom is specified for each dimension. Since this format is difficult to adapt to higher dimensional spectra it has been dropped. The old format can still be read by the XEASY program. However only the new format will be written out.

Two additional commands exist to check the lineshapes. The command "Check ref. line-shape list [cr]" looks which lineshapes are defined and whether the lineshapes position and the atom frequency are consistent. The command "View reference peak [vr]" displays for each peak with the same assignment as the selected one a 1D cross section in the slice window.

In 3D spectra, where overlap is a smaller problem, it suffices to partition the spectral data points into areas that belong to a given peak. Then the data points in these areas can be summed up to give the peak volume. The external program peakint (described in section 15.12 on page 106) is a routine to integrate peaks by partitioning. These two automated approaches have to be complemented by interactive methods where it is possible to specify the region to integrate in the spectrum. For interactive peak integration the commands "Interactive integration param. [ip]" "Interactive integration [ii]" can be used.

A different problem occurs when determining rate constants by evaluating a time series of spectra. It is crucial that the same regions out of each of the spectra at the different time points are integrated. This can be achieved by defining the regions to integrate by drawing rectangles (refer to section 3.9 on page 45) which can be saved and loaded again. The integration is performed with the command "Integrate rectangular regions [ir]"

Since NOESY spectral interpretation is an interactive process, where cross peak assignment, cross peak integration and structure calculation steps are repeated, efficient book keeping is crucial. For each peak, XEASY stores the integration method, the volume integral, and the error of the volume integral. The peak color can be used to reflect the state of the integration. In this way peaks with valid line shapes can be distinguished from peaks where lineshape information is missing or from peaks that were interactively integrated. Peaks can also be colored according to their integration method. To control the desired display mode the default window and the "Check ref. line-shape list [cr]" command are available.

In order to apply the lineshape integration selectively, peaks integrated interactively will not be changed by line shape integration. This allows the user to apply the lineshape integration routines to the whole spectrum without loosing the information of the interactively integrated peaks. In order to change the volume of an interactively integrated peak by line shape integration its integration method must be set back to "-" (for example by using the "Set selected peak entries [pe]" command).

28

3.9 Geometries

A basic feature necessary to work with spectra is to draw lines and rectangles (i.e. geometries) to mark features in the spectrum. Correspondingly a number of drawing commands exist in XEASY. They are:

- "Draw lines [dl]"
- "Draw ss lines [ds]"
- "Draw rectangles [dr]"
- "Draw transpose [dt]"
- "Draw freq. lines [df]", all in the Geometries Menu

The [ds] command draws lines from each selected peak to the diagonal. When applied to the TOCSY or COSY peaks of one fragment this corresponds to drawing spin system connectivities. The [df] command draws lines at selected frequencies. It can be used to identify a spin system using a TOCSY spectrum: clicking at all the peaks of a TOCSY tower will produce a grid identifying all the peaks belonging to one spin system.

Geometries can be removed with the commands "Remove lines [rl]" and "Remove rectangles [rr]" In order to remove all lines or rectangles at once these two commands can be combined with the "Apply to all ... [aa]" modifier command.

Geometries are saved in a file with the extension ".geom". The coordinates of the lines and rectangles are stored in ppm. A flag denotes whether the geometry is a line or a rectangle: "0" marks a line, "1" marks a rectangle. The commands to read or write a geometry file are "Load geometries [lg]" and "Write geometries [wg]".

---

next: 4. Spectrum Menu / contents

Looking

4. Spectrum Menu

New spectrum [ns]

Loads a spectrum. The program responds to this command by popping up a file window. After selecting a spectrum, the parameter file is read. In the second pop up window which appears, the view of the spectrum can be defined. The view indicates which dimension of the spectrum is displayed on the screen in the horizontal x direction and which in the vertical y direction. In addition the position and size of the first zoom region to display can be specified. The default settings for 2D spectra is to display the whole spectrum with w1 vertically and w2 horizontally. In 3D spectra by default the first w1 plane is displayed with w2 vertically and w3 horizontally.

The defined view is used by other routines in XEASY. It is changed with the Permutation [pm] command described on page 53.

Replace spectrum [rs]

Replaces the spectrum displayed as intensity plot. The new spectrum can be selected from all the spectra loaded with the [ns] or [as] commands. The intensity plot of the region where the command was invoked is updated. To change the intensity plot of all zoomed regions, the command can be combined with the [aa] modifier command. The spectra used for the intensity and contour plots are the upper and lower spectrum names, respectively, present in the upper left corners of each zoom region.

The zoom region history is not changed by this command.

Replace contour [rc]

This command works as the [rs] command described above. Instead of replacing the intensity plot, the spectrum displayed as contour plot is exchanged.

Alternative spectrum [as]

This command loads additional spectra and defines their view (i.e., which dimension is displayed in the horizontal x direction and which in the vertical y direction). It works as the New spectrum [ns] command. However, the new spectrum is not displayed in the main window and the zoom history remains unchanged.

Scale display [sd]

In different zoom regions intensity plots of different spectra can be displayed (see Zoom alignment [za], Zoom multiple [zm], Replace spectrum [rs]). If the intensity range of the different spectra differ, their color scale has to be matched. With the command [sd] a scaling constant for the intensity colors of one of the spectra is defined. A dialogue window is popped up where the scaling factor can be entered. The spectrum where the command was invoked is then displayed with the new scaling factor applied.

Calibration [ca]

Calibration is done in three steps: first a reference point in the spectrum has to be selected.

Second the ppm coordinate of the reference point together with the spectral sweep width, the spectrometer frequency (in MHz) and the spectrum type has to be entered. At this point the coordinate of the reference point may be adjusted. By default the spectral sweep width is entered in ppm. However Hz and the dwell

30

time in ms are also accepted. To indicate the desired unit the number has to be followed by either `Hz' or `us'.

Third, the maximal chemical shift in each dimension is calculated, displayed and can be corrected. The maximal chemical shift is the ppm frequency of the lower left corner of the spectrum.

After confirming the maximal chemical shift the parameter file of the spectrum is updated and the spectrum redisplayed with the new calibration.

XEASY supports two ways to calibrate spectra: If XEasy*traditional_calibration: true, the maximum chemical shift parameters are the ppm values of the lower left corner of the lower left spectral point. If XEasy*traditional_calibration: false, the maximum chemical shift parameters are the ppm values of the center of the lower left spectral point. In both cases the sweep width is n times the ppm difference between two data points (where n is the number of data points in the spectrum).

[next: 5. Zoom Spectrum Menu](#) / [contents](#)

Looking

31

5. Zoom Spectrum Menu

Zoom one region [zo]

This commands allows to select a zoom region in the main window. The user is expected to select the region in the main window. The selected zoom region is displayed in the main window.

The previously zoomed regions are saved in the zoom region history and can be called back using the "Restore zoom [rz]" command. To restore the full spectrum use the Permutation [pm] command.

Zoom multiple [zm]

The user can select several areas on the screen. Each of the selected areas is subsequently shown in a separate zoom region.

The previously zoomed regions are saved in the zoom region history and can be called back using the "Restore zoom [rz]" command. To restore the full spectrum use the Permutation [pm] command.

Zoom alignment [za]

The user can select several areas on the screen. If the selected areas overlap in the vertical dimension they are displayed in zoom regions arranged one beside the other. If they overlap only in the horizontal dimension they are arranged one above the other. If the regions do not overlap in any of the two dimensions a warning message is prompted and the command is aborted. In this case the Zoom multiple [zm] command can be used.

The sizes and positions of the regions are matched such that they show the same ppm range in the overlapping dimension. The command is useful to compare in detail different regions from the spectrum. To compare the peak positions in detail the cursor can be changed with the Draw cursor [dc] command, into a cross extending over the whole main window.

Dublicate zoom [zd]

The zoom region in which the cursor is when invoking the command is dublicated. When combined with the Apply to all ... [aa] modifier command all currently displayed region are dublicated. To restore a single zoom region the View area [va] command is available.

Zoom transpose [zt]

This command transposes the zoomed region onto the other side of the diagonal. The ppm coordinates are used to calculate the new position of the zoom region.

The command can be used in conjunction with the Draw transpose [dt] command to check assignments and peak picking at the symmetric region in the spectrum.

Zoom symmetric peaks [zs]

This command can be used to check symmetric peaks in 3D spectra. The user is expected to select a reference peak. The reference peak and its symmetric peak and the corresponding diagonal peaks are then displayed - each in a small region. The size of the region can be defined using the Zoom factor [zf] command.

32

The symmetric peak and the diagonal peaks are identified based on their assignments: the symmetric peak has the assignment in the x and y direction from the reference peak exchanged. For example, when displaying w3 in x direction and w2 in y direction, the reference peak N 22 / HN 22 / HN 21 defines its diagonal peak N 22 / HN 22 / HN 22, the symmetric peak N 21 / HN 21 / HN 22 and its diagonal peak N21 / HN 21 / HN 21. Either the symmetric peak or its diagonal peak must be picked and assigned. If this is not the case, a warning message is prompted and the command aborted.

To search for unassigned symmetric peaks the Compare close strips [pc] command can be used.

Zoom peak [zp]

Each visible peak is displayed in a separate region. The command works as if all the parts of the spectrum which do not contain a visible peak are cut out and the remaining regions of the spectrum are displayed. To select the peaks the command Select peak [sp] can be used to define the size of the displayed regions the Zoom factor [zf] command.

This command works only for 2D spectra.

Zoom geometries [zg]

The command is similar to the Zoom peak [zp] command. However it does not use peaks to define the regions to zoom but rather the geometries (i.e., drawn lines and rectangles, see Geometries Menu). Each drawn line in the spectrum defines a frequency in either x or y direction. Each point where two frequencies (one defined in the x direction the other defined in the y direction) cross each other a small region is zoomed. In the scheme of Fig. 5.7,
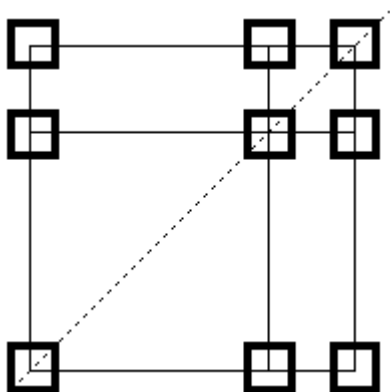
Fig.5.7

the diagonal of the spectrum is shown with a dotted line and the geometries are shown with a thin solid line. The regions that will be zoomed using the indicated geometries are indicated by bold squares. As suggested by the scheme, this command helps to assign spin systems using 2D TOCSY and COSY spectra: first the frequencies of the spin system are identified in a TOCSY tower (i.e., peaks located at the same frequency in w1) using the Draw freq. lines [df] or Draw transpose [dt] commands. Then all the regions containing the peaks of the spin system are expanded.

The command works only in 2D spectra.

View peak [vp]

This command is used to search a peak with given peak number. The peak number or the assignment (i.e., HA2 33) in the horizontal dimension must be specified in a dialogue window. If an assignment is specified the first peak with this assignment in the horizontal dimension is selected. When the command is invoked a second time and the same assignment is specified the next peak with this assignment is shown.

If the found peak lies in one of the currently displayed zoom regions it is marked with a square. If the

33

peak is not visible, the zoom region, where the command was invoked, is moved such that the desired peak lies in the middle of the new region.

Enlarge vertically [ev]

The zoomed region, where the command is invoked, is expanded in the vertical y direction. To apply the command to all regions simultaneously the Apply to all ... [aa] modifier command can be used. The size sn of the new region is determined by the old size so and the zoom factor f, which is defined with the Zoom factor [zf] command.

```
sn = so * (1 + f)                 Eq. 5.10
```

To shrink a region using this command the zoom factor f has to be set to a value between -1 and 0.

The zoom region history is not changed by this command.

Enlarge horizontally [eh]

The zoomed region where the command is invoked, is expanded in the horizontal x direction. To apply the command to all regions simultaneously the Apply to all ... [aa] modifier command can be used. The size sn of the new region is determined by equation 5.10 using the old size so and the zoom factor f, which is defined with the Zoom factor [zf] command. To shrink a region using this command the zoom factor f has to be set to a value between -1 and 0.

The zoom region history is not changed by this command.

Goto left [gl]

The zoomed region, where the command is invoked, is moved to the left. To apply the command to all regions simultaneously, the Apply to all ... [aa] modifier command can be used. The new position pn of the region is determined by equation 5.12 from the old position po, the size of the region s and the zoom factor f:

```
pn = po - f * s                 Eq. 5.12
```

The zoom factor is defined using the Zoom factor [zf] command.

The zoom region history is not changed by this command. To revert its action the Goto right [gr] command can be used.

Goto right [gr]

The selected region or regions are moved to the right. The command works as the Goto left [gl] command described above.

Go up [gu]

The selected region or regions are moved upward in the spectrum. The command works as the Goto left [gl] command described above.

Go down [gd]

The selected region or regions are moved downward in the spectrum. The command works as the Goto left [gl] command described above.

"Restore zoom [rz]"

34

Restores the previous zoom of the zoom region history. All the zooming commands that cannot simply be reversed, store the previous zoom to the zoom region history. To call back the last zoom, the [rz] command is used. To go more than one step back in the zoom history, the zb and zo two letter codes in the overview window (Overview window [ow]) are used to step backwards.

Forward plane [fp]

The next plane in a higher dimensional spectrum is zoomed. In the case of 4D spectra a pop up window appears where the dimension to increase must be specified.

The zoom region history is not changed.

Backward plane [bp]

The next plane in a higher dimensional spectrum is zoomed. In the case of 4D spectra a pop up window appears where the dimension to increase must be specified.

The zoom region history is not changed.

Permutation [pm]

The view of the spectrum is defined and a zoom region to display can be specified by entering its position and size in points. The command works as the New spectrum [ns] command but does not load a new spectrum. A pop up window appears where the desired view can be specified. The view defines which dimension of the spectrum is displayed on the screen in the horizontal x direction and which in the vertical y direction. In addition it is possible to define the position and size of the first zoom region to display. The default settings for 2D spectra is to display the whole spectrum with w1 vertically and w2 horizontally. In 3D spectra by default the first w1 plane is displayed with w2 vertically and w3 horizontally.

The defined view is used by other routines in XEASY.

Goto plane [gp]

The zoomed region is moved to the indicated plane. A pop up window appears, where the desired plane can be specified in data points. In 4D spectra the dimension in which to change the plane must be specified as well.

The zoom region, where the command was invoked, is then moved to the specified plane.

The command does not change the zoom region history.

Contour plot [cp]

Changes the drawing mode to contour lines only. The user is prompted to enter parameters defining the contour plot: intensity of the first contour level, factor between the intensities of two subsequent contour levels, number of contour levels, and whether to plot only positive, only negative or positive as well as negative contours.

After the parameters are entered, the zoom regions are redisplayed. Only contour lines will be shown. The spectrum used for the contour plot is the lower spectrum name indicated in the upper left corner of each of the zoom regions. It can be changed using the Replace contour [rc] command . To switch back to the intensity mode the Restore intensity[ri] is available. To see both contour and intensity plots, use the Int. and contour [ic] command.

Int. and contour [ic]

Changes the drawing mode to display both the intensity plot and contour lines. As with the Contour plot [cp] command the parameters for the contour plot must be specified. After the parameters are entered, contour lines and intensity plot are displayed. The spectra used for the intensity and contour plots are the upper and lower spectrum names, respectively, present in the upper left corner of each zoom region. They can be changed individually using the Replace contour [rc] and Replace spectrum [rs] commands. To display only the intensity plot or the contour plot use the commands Contour plot [cp] or Restore intensity[ri].

Restore intensity[ri]

Changes the drawing mode to intensity plot only. Only the intensity plot will be shown. The spectrum used for the intensity plot is the upper spectrum name indicated in the upper left corner of each region. It can be changed using the Replace spectrum [rs] commands. To switch to the contour plot mode or to view both contour and intensity plots the commands Int. and contour [ic] and Contour plot [cp] can be used.

Plotting file [pf]

Generates a plot file from the displayed zoom regions. As with the Contour plot [cp] command, the parameters for the contouring must be specified. Different file formats are available for output: postscript formats for A3 and A4 paper sizes and HP GL formats for A4, A3 (for model HP7550) and A0 (for model HP7596) paper sizes. The plot is scaled by default to fill the selected paper size. The following attributes can be changed. For each of the parameters the first entered value specifies the horizontal axis, the second entered value the vertical axis. If only one value is specified it applies to both axes. An asterix * denotes the default value supplied by the program.

Size in cm:
> size of the displayed zoom regions (without axis and labels) on the plot.

Offset in cm:
> position of the lower left corner of the plot on the paper.

Label spacing in ppm:
> distance between subsequent labelled tics in ppm. When plotting many small zoom regions (e.g., strips) on one plot, the center of each zoom region can be labelled by entering a `-' instead of a distance in ppm. For plotting strips, for example, `- *' is a good setting for the label spacing. It results in the default spacing on the vertical axis and on the horizontal axis each strip is marked with a single label at its center

Tic spacing in ppm:
> distance between subsequent small tics drawn in-between adjacent labels.

Draw grid:
> Whether to draw a grid. When the option is turned on, grid lines are plotted at the position of the labels.

Strip plotting mode:
> If you enter "y" ppm values of the indirect dimension(s) are plotted at the top of the strip. To plot strips it is advisable to enter
> Label spacing "- *"
> Tic spacing "2 *" (or some other large value)

After the parameters are entered a file window pops up where the name of the output file can be selected. A plot file with extension .plot is written. The plots will contain the contour plot of the regions, the displayed peaks with their labels and, if selected, a grid. Geometries are not plotted. The plot file can be printed or plotted on a suitable output device using the UNIX commands lpr (BSD) or lp (System V).

36

View area [va]

The command allows a zoom region to be selected, which will be expanded to fill the main window. The previous zoom regions are stored in the zoom region history.

Zoom factor [zf]

Defines the parameters used when shifting around a zoom region in the spectrum and when displaying strips (Strips Menu) or small regions for a subset of peaks ("Zoom peak [zp]" command).

Load zoom [lz]

Loads and displays zoom regions written out with the "Write zoom [wz]" command. The file name can be specified in a file window. Only the position and the size (given in ppm) in the displayed x and y direction are read. In 3D and 4D spectra the plane from the zoom region where the command was invoked is taken to display the loaded zoom regions.

Write zoom [wz]

Writes the zoom regions from the current zoom to a file. The file name can be specified in a file window. Only the size and position in ppm in the displayed x and y direction are saved to the file. In 3D and 4D spectra the position of the current plane is not stored.

---

next: 6. Strips Menu / contents

Looking

37

6. Strips Menu

Load strip list [sl]

A strip list file can be selected in the file window. The strips listed in the file are appended to the current strip list. They are assigned to the spectrum indicated in the strip list file. If this spectrum is not loaded, the program tries to find a spectrum with the same dimensionality. First the intensity spectrum from the zoom region where the command was issued is considered. If this does not fit, the remaining spectra are examined.

To load a completely new strip list the command Erase strip list [es] can be used to reset the current strip list prior to loading the new one.

Write strip list [ss]

The file name for saving the current strip list can be selected in the file window. The user specified extension is always removed and the extension .strip is appended. If no directory is specified the file is written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environment variable defining the groups directory and user_name is the users login name.

If a file with the same name exists already a `%' is added to the existing files name prior to writing out the strip list.

Strip sequence [se]

All selected peaks are appended to the current strip list. A pop up asks whether the produced strip list should be sorted and whether strips with same assignments should be deleted.

For each selected peak (see also the Select peak [sp] command) a new strip is added to the current strip list. The position, the folding, and the assignment of the peak are copied to the strip. The strip is assigned to the spectrum where the command was issued.

When Sort strips [y/n] is turned on, the whole strip list is sorted using the assignments. 2D strips are first sorted according to their assignments in the x direction then according to their assignment in the y direction. 3D strips are sorted first according to their assignment in the x direction and then according to their assignment in the dimension perpendicular to the current view. 4D strips are sorted in the two dimensions perpendicular to the current view. The ordering of the assignments is given by the position of the atoms in the atom list file. When working with a standard setup this corresponds to sorting according to residue number. However by reordering the atom list, the order of the strip list can be controlled. For example if HN of residue 33 comes before HN of residue 34 in the atom list file (which should normally be the case), then a strip assigned to HN 33 will be sorted in front of a strip assigned to HN 34.

When Sort and delete duplicated strips [y,n] is turned on the whole strip list is first sorted according to the above description. Then all strips having the same assignment are removed. For strips from 2D spectra only the assignments in the x direction are considered. In the case of 3D spectra, the assignments in the x direction and in the direction perpendicular to the view are compared. In 4D spectra, the assignments of the two directions perpendicular to the view are used.

Goto strips [gs]

A screen full of strips from the current strip list is displayed. The number of strips to display, the width and the starting position within the current strip list must be entered in a dialogue window. The height in

38

ppm is taken from the zoom region where the command is invoked. For each strip, a region at the position of the strip is zoomed from the spectrum to which the strip is assigned.

If the strips are assigned (i.e., generated from peaks that were assigned) the number of the fragment, assigned to the strip in the horizontal x direction, is displayed in the lower left corner of each strip. If the mapping number or fragment comment of the corresponding fragment is set it is also displayed.

If the number of strips to display on one screen is bigger than the number of strips in the current strip list the number of strips to display is reduced. The starting position within the current strip list is always adjusted in order to display the desired number of strips together.

Strip find [sf]

Moves the position of the displayed strips in order to show the first strip assigned to a specified fragment. The fragment number must be entered in a dialogue pop up. The current strip list is searched for the first strip assigned to the specified fragment. If it is found it is displayed as the first strip of the displayed series of strips. The starting position to display strips from the current strip list is adjusted in order to display the desired number of strips together. So there may be cases where the selected residue does not appear as the first strip. The strips are displayed as with the Goto strips [gs] command.

Forward strips [fs]

Displays the next screen full of strips from the current strip list. The position of the strips to display from the current strip list is incremented by the number of strips to display on one screen. The strips are displayed as with the Goto strips [gs] command.

Backward strips [bs]

Displays the previous screen full of strips from the current strip list. The position of the strips to display from the current strip list is decremented by the number of strips to display on one screen. The strips are displayed as in the case of the Goto strips [gs] command.

Erase strip list [es]

The current strip list is removed from memory. The command is important since the Load strip list [sl] and Strip sequence [se] command only append strips to the current strip list.

Append strip [sa]

Appends a single strip to the strip sequence. After a peak is selected, its position, folding and assignment are then copied to a new strip. This new strip is appended to the current strip list.

Insert before reference [ib]

Inserts a new strip before a selected reference strip in the current strip list. First the reference strip must be selected. This is done by clicking the mouse within the desired reference strip. Then the new strip to insert must be defined. This is done by selecting a peak at the position of the new strip. The position, folding and assignment of the selected peak are copied to the new strip and the new strip is inserted into the current strip list in front of the selected reference strip. The modified strip list is displayed: if the strips last displayed contain the inserted strip they are redisplayed. Otherwise the dialogue window from the Goto strips [gs] command is popped up. The default value for the starting position of the displayed strips within the current strip list is set in order to show the insertion position.

Remove one strip [ro]

Removes a single strip from the current strip list. The strip to remove must be selected by clicking the mouse within the strip. The program selects the strip closest to the selected position. If several strips occur at identical positions the one occurring first in the current strip list is selected. The selected strip is removed from the current strip sequence. The modified strip list is displayed. If the strips last displayed contain the position of the removed strip they are redisplayed. Otherwise the dialogue window from the Goto strips [gs] command is popped up. The default value for the starting position of the displayed strips within the current strip list is set in order to include the position of the removed strip.

Hold strip [sh]

Defines a strips to be displayed on the left side of the screen. After a peak is selected, its position, folding and assignment are copied to a new strip. This new strip is appended to the list of fixed strips. The list of fixed strips is displayed on the left of the screen whenever strips are displayed.

Release strip [sr]

Removes a single strip from the list of fixed strips. The strip to remove must be selected by clicking the mouse within the strip to remove. To remove all the strips from the list of fixed strips the command can be combined with the Apply to all ... [aa] command.

Define reference strips [rd]

Defines the reference strip list to be used for the Compare close strips [pc] command. The current strip list is copied to the reference strip list for the position comparison which is used for searching strips that lie close to a user defined position. Subsequently the current strip list can be changed without affecting the Compare close strips [pc] command. For an overview of the commands dealing with strips refer to Strips section in XEASY Model.

Define correlation [cd]

Defines the correlation strip list and calculates the correlation coefficients. The calculated correlation coefficients are used by the Compare correlated strips [cc] command.

The current strip list is copied to the correlation strip list. For each pair of strips the correlation coefficient (Eq. [1]), describing the spectral similarities between the strips is calculated. The following parameters can be specified:

Factor to enhance peaks:
   the factors Am (Eq. [2]) used to enhance the peaks in the expected peak patterns , derived from the intensities observed in the reference strips s. Cross peaks with the same assignment as the reference strip define the positions to be scaled. These cross peaks are sorted by decresing chemical shift in the vertivcally displayed dimension. The first of the specified factors applies to the position defined by the cross peak with highest chemical shift, the second factor for the position defined by the cross peak with second highest shift, etc. The last of the specified factors applies to the remaining cross peaks. The default setting (1.0 4.0) is suitable for sequential assignment using 15N-correlated [1H,1H]-NOESY spectrum together with the peak list from a 15N-correlated [1H,1H]-TOCSY spectrum: the diagonal N-HN-HN peak is not emphasized while the local maxima at the positions of the other peaks are all multiplied by 4.0. For all other assignment tasks it is recomended to set this factor to 1.0.
Scaling Factor for intensities:
   the noise level used for the contour plot multiplied by the specified factor defines the constant ib used in Eq. [2]. The default setting 5 works usually fine.
Number of diagonal points to suppress:
   may improve the performance of the correlation method in cases where there are many strips with diagonal peaks at the same position in the y direction. It is specified in data points and should be set to the peak width of the diagonal peaks in the y direction Usually this is not necessary and the

40

default value 0 can be used.

[1]inter or [2]inter spectral search:

Whether to calculate the correlation between strips defined in the same spectrum or only between strips from different spectra. In all the cases discussed in the Applications the default value 1 can be used.

Number of weighting regions:

weighting regions are only necessary if the spectrum contains strong artifacts (e.g. axial peaks). If the spectral region containing the artifact, does not contain any peaks it can be excluded from the calculation using a weighting factor of 0.0, if the region contains some relevant peaks its influence can still be reduced by choosing, for example, a weighting factor of 0.3.

Factor to reduce peaks:

the factors used to suppress peaks in the peak patterns of the candidate strips k. Local maxima at the positions of picked peaks are divided by the specified factors (the inverse of the specified factors corresponds to Am of Eq. [2]). Again the first of the specified factors applies to the position defined by the cross peak with highest chemical shift, the second factor for the position defined by the cross peak with second highest shift, etc. The last of the specified factors applies to the remaining cross peaks. The default setting (1.0 3.0) is suitable for sequential assignment using 15N-correlated [1H,1H]-NOESY spectrum together with the peak list from a 15N-correlated [1H,1H]-TOCSY spectrum: the diagonal N-HN-HN peak is not supressed while the local maxima at the positions of the other peaks are all divided by 3.0. For all other assignment tasks it is recomended to set this factor to 1.0.

Tolerance for identified frequency:

The tolerance D (Eq. [1]) in ppm units. The default value -1.0 specifies that the first part of Eq. [1] should not be used and the strips get only sorted according to the similarities of the peak patterns. This is appropriate for sequential assignment using 15N-correlated [1H,1H]-NOESY spectra. In other cases D should be set to the uncertainty with wich peak positions in the vertical dimension can be identified.

Compare close strips [pc]

Warning: pc does not work - as you might expect - by selecting strips with peaks near to the position where you clicked the mouse. It selects strips of which **the peak that defines the strip** is near to the selected position. Consequently, you have to "se" with Sort and delete duplicated = n, "rd", "es" and again "se" with Sort and delete duplicated = y. Now you will possibly get all the strips that you want.

Displays the strips from the reference strip list (see also the Define reference strips [rd] command) sorted by increasing distance from a selected position. The position is selected by clicking the mouse in a zoom region. If the intensity spectrum of this zoom region is a 2D or 3D spectrum, the ppm position in the y direction py is used. In this case the distance d of the selected position and the position of the strip in the y direction sy is calculated as

$$d = \left| p_y - s_y \right|$$

6.16.A

In the case of a 4D spectrum both the position in the x direction px and the position py in the y direction are used. Since the sweep widths wx and wy in the two dimensions might differ considerably, it is necessary to scale the distances in the two dimensions. The distance d to the position sx and sy of a given strip is thus calculated as

$$d = \sqrt{\left(\frac{p_x - s_x}{w_x}\right)^2 + \left(\frac{p_y - s_y}{w_y}\right)^2}$$

6.16.B

A reference strip to display together with the sorted strips from the reference list is defined. The reference strip is taken from the zoom regions at the position where the mouse was clicked. This allows close

41

comparison of the reference strip with the strips lying close to the selected position. For example, the slice window can be used to make a detailed comparison of lineshapes selected in different strips.

Displaying the comparison strip list (i.e., the reference strip together with the sorted strips from the reference strip list) works by displaying parts of the current strip list. The width of the displayed strips and the number of strips to display are set with the Goto strips [gs] command. The height in ppm is defined by the zoom region where the [pc] command is invoked. For each strip, a region at the position of the strip is zoomed from the spectrum to which the strip is assigned.

If the number of strips to display on one screen is bigger than the number of strips in the reference strip list plus one, the number of strips to display is reduced. The starting position within the reference strip list is always adjusted to display the desired number of strips.

To view additional strips from the comparison strip list the commands [fc] and [bc] can be used. They show the following or previous set of strips from the sorted reference strip list. The first strip displayed is always the reference strip.

The Forward comparison [fc] and Backward comparison [bc] commands are used together with both the Compare close strips [pc] and the Compare correlated strips [cc] commands to compare strips either by correlation or by position. After the [pc] command was used the [fc] and [bc] commands display the next or previous set of strips close to the selected position. After the [cc] was used they display the next or previous set of strips correlated to the selected reference strip.

Compare correlated strips [cc]

Displays strips from the correlation strip list (see also the [cd] command) sorted by decreasing distance as defined by Eq. [1] to a selected reference strip. The reference strip and the frequency f used in Eq. [1] aee selected by clicking the mouse at a position in the desired strip. The frequency in the vertical dimesnion of the selected position defines the frequency f. The strips from the correlation strip list are displayed together with the reference strip. They are sorted by decreasing spectral correlation coefficient to the reference strip.

Displaying the comparison strips (i.e. the reference strip together with the sorted strips from the correlation strip list) works as displaying parts of the current strip list. The width of the displayed strips and the number of strips to display are set with the Goto strips [gs] command. The height in ppm is defined by the zoom region where the [pc] command is invoked. For each strip, a region is zoomed from the spectrum to which the strip is assigned at the position of the strip.

If the number of strips to display on one screen is bigger than the number of strips in the correlation strip list plus one, the number of strips to display is reduced. The starting position within the correlation strip list is always adjusted to display the desired number of strips together.

To view additional strips from the comparison strip list the commands Forward comparison [fc] and Backward comparison [bc] can be used. They show the following or previous set of strips from the sorted correlation strip. The first strip displayed is always the reference strip.

The Forward comparison [fc] and Backward comparison [bc] commands are used together with both the Compare close strips [pc] and the Compare correlated strips [cc] commands to compare strips either by correlation or by position. After the [pc] command was used the [fc] and [bc] commands display the next or previous set of strips close to the selected position. After the [cc] was used they display the next or previous set of strips correlated to the selected reference strip.

Forward comparison [fc]

Views additional strips from the comparison strip list. Depending on whether the comparison strip list was defined using the Compare close strips [pc] command or the Compare correlated strips [cc] command, the displayed strips are either from the sorted reference strip list or the sorted correlation strip

42

list. The first strip displayed is always the reference strip.

Backward comparison [bc]

Views the previous set of strips from the sorted correlation strip list or the sorted reference strip list. Depending on whether the comparison strip list was defined using the Compare close strips [cc] command, or the Compare correlated strips [cc] command the displayed strips are from the sorted reference strip list or the sorted correlation strip list. The first strip displayed is always the reference strip.

Sequential assignment [op]

The command invokes an automatic assignment routine, originally proposed by R. Bernstein, C. Cieslar, A. Ross, H. Oschkinat, J. Freund and T. A. Holak. in J. Biomol. NMR, 3 (1993) 245-251. To improve the performance the proposed method has been combined with a simulated annealing scheme. Experience so far has shown that for bigger proteins where only limited information about the spin system types is present, the performance of the method is only sufficient to provide partial (less than 50%) assignment. The defined score, however, seems to be capable of distinguishing wrong assignments from the correct one. This indicates that the sampling property of the algorithm is the problem.

The routine tries to assign the strips form the correlation strip list to amino acid residues in the fragment list. The strips should be assigned to spin systems in the fragment list. This implies that the fragment list to be used should contain entries for both the amino acid residues and the spin systems. This is possible by using different sets of fragment numbers for the spin systems and the amino acid residues.

The used score includes contribution from the quality of the sequential connectivities and from the matching of the spin system type to the amino acid residue. The spectral correlation coefficients are used to judge the quality of the sequential connectivities. Possible amino acid types of a spin system can be specified by the user in the fragment comment field in the peak editing window (Editing peaks [ed] command). To specify that a given spin system is, for example, an AMX spin system (i.e., one of the amino acids Asp, Asn, Phe, Tyr, Trp or His) the comment field must be set to DNFYWH, or to specify that the spin system is most likely a glycine, it must be set to G.

Known sequential assignments of a spin system can be used by setting the mapping number of the spin system to the fragment number of the assigned amino acid residue. The assignment of the spin systems for which the mapping numbers are set, can be fixed and are then not changed by the optimization algorithm.

The parameters to be specified for running the routine are described below. In general the proposed default values can be used. By setting the starting temperature to zero the routine works as the one proposed by R. Bernstein et al.

First residue to assign:
> the fragment number of the first amino acid residue that should be assigned.

Last residue to assign:
> the fragment number of the last amino acid residue that should be assigned.

Number of iterations:
> number of annealing steps to perform

Number of fragments to reset:
> the algorithm works by reassigning stretches of amino acid residues. The size of these stretches is defined with this option.

Use mapping to [s]tart, [f]ix, [n]ot:
> indicates how the mapping numbers of the spin systems are used. [s] the mapping numbers are used to define assignment with which the algorithm starts. This can be used to locally optimize an

43

assignment already made. [f] the mapping numbers are used to fix the sequential assignment of the corresponding spin systems. For example setting the mapping number of spin system 236 to 22 assigns spin system 236 permanently to amino acid residue 22. [n] the mapping numbers are not used.

Continue last optimization [y/n]:
to implement involved simulated annealing schemes the routine can be called several times. If this option is set to y the routine continues with the previous run.

Copy solution to mapping number [y/n]:
the results of the optimization are listed in the text window. If desired they can also be reported in the mapping numbers of the spin system fragments.

Number of iterations to equilibrate:
number of iteration for each annealing step (i.e., number of equilibrating steps at each temperature).

Start temperature:
temperature used for the first annealing step.

Cooling factor:
cooling factor to multiply the temperature after each annealing step.

---

next: 7. Peak Analysis Menu / contents

Looking

44

7.0 Peak Analysis Menu

7.1 "Load peaklist [lp]"

The peak list to load can be selected in the file window. The peaks are loaded as a new list into memory. The new peak list is named after the peak list file. If a peak list with the same name is already present it is replaced. The list remains in memory until it is removed with the "Erase peak list [ep]" command or overwritten with a new peak list with the same name. The peaks are displayed together with the spectrum shown in the current zoom region. To associate the peak list to all the spectra currently loaded the modifier command "Apply to all ... [aa]" must be issued before using the [lp] command. Refer to the "Exchange peak list [xc]" command for an explanation of how to associate the different loaded peak lists to the different loaded spectra.

The peak list is adapted to the spectrum of the zoom region where the command was issued: all the peaks are folded into the spectrum, the dimensionality is adapted to the one of the spectrum. The peaks are folded depending on the setting of folding type in the spectral parameter file. Allowed settings are "TPPI" and "RSH". If **sw** denotes the sweep width in ppm of the spectrum, **p**max and **p**min the maximum and minimum chemical shift in ppm of the spectrum and **p**$unfold$ is the unfolded chemical shift of the peak then the folded chemical shift **p**RSH and the folding value **n**$fold$ are calculated in the case of RSH folding as

$$n_{fold} = (p_{unfold} - p_{min}) \operatorname{div} sw$$

7.1.A

$$p_{RSH} = p_{unfold} - n_{fold} \cdot sw$$

7.1.B

In the case of TPPI for uneven folding values nfold the shift calculated according to equation 7.1.B is reflected according to

$$p_{TPPI} = p_{max} + p_{min} - p_{RSH}$$

7.1.C

When writing out a folded peak list, the unfolded chemical shifts are re-calculated from the folding value nfold and the folded chemical shift. It is possible to load a peak list with different dimensionality than the spectrum. However it must be defined how the dimensions of the peak list file should be used for the adapted peak list. If the adapted peak list has still more dimensions than the corresponding spectra the surplus dimensions will be kept in memory and can be edited with for example the "Editing peaks [ed]" command, but they are not used for displaying the peaks.

A pop up window appears where each dimension of the adapted peak list must be defined. Three possibilities to define one dimension exist. First, entering a number sets the chemical shift in this dimension of all peaks to the entered value. Second, specifying a dimension from the loaded peak list copies, for each peak, the chemical shift and the assignment from the specified dimension of the peak list. To define the dimension a "w" must proceed the number. Third, when loading a peak list with more dimensions than the spectrum, it is possible to discard the surplus dimensions by entering a dash - for the surplus dimension(s). This option is only available for the surplus dimensions since the peaks have to be defined in all dimensions of the spectrum.

For example when loading a peak list from a 15N-correlated HMQC-spectrum onto a 3D 15N-correlated spectrum setting the field "For new w3 use [w? ,ppm or -]" to 4.65 puts all the peaks into one w1/w2 plane at a w3 chemical shift of 4.65 ppm. Setting the same field to w2 generates a list of diagonal peaks with equal w2 and w3 shifts. When loading a 3D peak list onto a 2D spectrum, it is possible to discard the third dimension by setting "For new w3 use [w? ,ppm or -] " to -. Once a peak list has been adapted to a

45

particular spectrum, it can also be displayed together with different spectra. The command "Exchange peak list [xc]" selects a peaklist to display together with a given spectrum. It does not adapt this peak list to the spectrum. To adapt (dimensionality and folding) a loaded peak list to another spectrum, the command "Adapt peaks to spectrum [ad]" must be used.

7.2 "Write peaklist [wp]"

The peak list displayed in the region where the command is issued is written to a file. The file name can be entered in the file window. Any extension is removed and the extension ".peaks" is added. If no directory is specified the file will be written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environmental variable defining the groups directory and user_name is the users login name.

The peaks are unfolded prior to writing them out. The dimensions are written out as they are displayed in the peak edit window.

If a file with the same name already exists the existing file is backed up by appending a `%' to its name prior to writing out the new peak list.

7.3 "Delete peaks one by one [dp]"

Single peaks can be removed by clicking the mouse close to the picked peak position. The peak is deleted from the peak list in memory. The peak numbers of all the remaining peaks are not changed. Only displayed peaks are considered.

7.4 "Remove peaks in one region [rp]"

All peaks within a selected region are deleted. The region is selected with the left mouse button. All peaks currently displayed in the selected region are deleted. The peak numbers of all the remaining peaks are not changed. Since only the displayed peaks are deleted, the "Select peak [sp]" command can be used to select a subset of peaks to delete. For example, all diagonal peaks or all sequential peaks can be selected and removed.

7.5 "Erase peak list [ep]"

Removes a peak list from the memory. The peak list to remove can be selected in a pop up window.

7.6 "Assign peaks in one region [ar]"

Assigns all displayed peaks within a selected region to different fragments. The first fragment number used and the atom names in all dimensions must be specified. The first displayed peak is assigned to the specified atoms and fragment. The second peak is assigned to the same atoms but the fragment number is incremented by one. For each subsequent peak, the fragment number is incremented by one. The corresponding atoms must be present in the atom list, otherwise an error message is prompted and the command is aborted. To generate the necessary atom entries, load a sequence of suitable fragments using the "Load sequence [ls]" command.

To restrict the command to only a subset of peaks, the "Select peak [sp]" command can be used. For example, all the diagonal peaks in a 3D 15N-correlated spectrum can be assigned to the N and HN of subsequent back bone fragments.

7.7 "Move peaks one by one [mp]"

This command enables the user to move single peaks. First, the peak to move has to be selected by clicking the left mouse button close to its picked position. The selected peak is emphasized by a circle

drawn around it. Then the new position can be defined by pressing the left mouse button with the cursor at the desired position. The peak is moved to this new position. The folding of the peak remains unchanged. To change the folding, use the peak editing window or the "Set selected peak entries [pe]" command.

7.8 "Fetch peak into plane [pl]"

Fetches single peaks into the currently displayed zoom region of a 3D or 4D spectrum. A displayed peak can be selected by clicking the mouse close to its picked position. The selected peak is moved in the dimension(s) perpendicular to the displayed x and y direction into the middle of the displayed plane. Its position in the horizontal x direction and the vertical y direction as well as its folding remain unchanged.

7.9 "Copy and move peaks [cm]"

This command allows the user to copy the assignment and folding information from already picked peaks to new peaks. First, the reference peak must be selected by clicking the left mouse button close to its picked position. The position of the new peak is then defined by clicking the mouse at the desired new position. In 2D and 3D spectra, the position of the new peak is only allowed to differ in one of the two displayed dimensions: the program selects the dimension where the new position shows the largest difference from the position of the reference peak. In 4D spectra both displayed dimensions are allowed to vary. The positions, foldings, and assignments in the dimensions which are kept constant, are copied from the reference peak to the new peak.

7.10 "Move reference peak [mr]"

Move a reference peak and all peaks with the same assignment. First, the reference peak to be moved must be selected by clicking the left mouse button close to its picked position. The selected reference peak is indicated by a circle drawn around it. Once this is done, the new position can be defined by pressing the left mouse button. The reference peak is then moved to the new position. If the peak is assigned, the new position and the folding define the chemical shifts of the corresponding atoms. All peaks assigned to these atoms are moved onto this shift. Since the positions of the picked peaks are adjusted in all dimensions, it is important that the spectrum is carefully calibrated. Folding is handled as follows: the shifts of the reference peak are unfolded to calculate the unfolded chemical shifts of the atoms, the shifts of the peaks with the same assignment are then set to these unfolded shifts and subsequently the peaks are again folded into the spectrum.

7.11 "Move peak list [ml]"

Moves the entire peak list. A selected reference peak is moved. All the other peaks from the peak list are moved by the same amount. The reference peak is selected by clicking the left mouse button close to its picked position. The new position is then selected by clicking the left mouse button at the desired position. The peak positions in the dimensions perpendicular to the displayed view and the folding values are kept constant.

7.12 "Exchange peak list [xc]"

Associates a particular peak list with a given spectrum. The peak list can be selected in a pop up window. The chosen peak list is then associated with the spectrum displayed in the current zoom region. To associate a peak list to all loaded spectra, the "Apply to all ... [aa]" modifier command can be used. Neither the folding nor the order of dimensions are changed.

7.13 "Adapt peaks to spectrum [ad]"

The folding and the order of dimensions are adapted to the spectrum displayed in the current zoom region. The command corresponds to writing out the peak list to a dummy file and reading it in again. First, the

47

peaks are unfolded. The dimensions can then be exchanged according to the descriptions given in the chapter about the "Load peaklist [lp]" command. Finally, the peaks are folded into the new spectrum.

7.14 "Peak picking [pp]"

This command is designed for manual peak-peaking. Clicking the left mouse button adds a new peak to the peak list shown in the corresponding zoom region. If no peak list is loaded a new peak list with the default name of the displayed spectrum is generated. Pressing the middle button terminates this command. The added peaks are unassigned and not folded. In the case of 3D and 4D spectra they are positioned in the middle of the displayed plane.

7.15 "Assign peak [ap]"

This command allows to assign a peak by specifying either the atom number or atom name and residue number. In a popup window, the assignment in all dimensions must be specified. If an unknown atom is entered the command aborts and the program asks the user whether the new atom should be added to the atom list. On confirmation, a new atom entry with the given atom name and fragment number is added to the atom list.

Alternative ways to add new atoms to the atom list are the "Add new fragment [af]" command, changing the fragment name in the peak edit window (refer to the "Editing peaks [ed]" command) and changing the atom name in the peak edit window.

7.16 "Select peak [sp]"

Instead of displaying all peaks of a peak list, this command enables the user to make a selection of peaks to be displayed according to a number of criteria. The criteria are entered in a dialogue window. With the option "Select (s), subtract (-) or add (+) peaks:" set to the default value `s' all peaks which match the specified criteria are displayed. If several criteria are specified only peaks that match all of them are displayed. If the option is set to `-' the specified peaks are removed from the set of displayed peaks. If the option is set to `+' the specified peaks are included into the set of displayed peaks.

The command can be combined with many other commands as the "Remove peaks in one region [rp]", "List peak entries [le]" or the "Strip sequence [se]". This type of command perform their action only on those peaks that are displayed (i.e., which were selected with the [sp] command).

The "Display all peaks [da]" command is a short cut to redisplay all peaks. The following paragraphs describe the different selection possibilities.

Fragment numbers (e.g. 12 13 or 44-74):
> allows the user to specify a range of fragment numbers or a set of fragment numbers. If a `-' occurs after the first number the specification defines the allowed range of fragment numbers, otherwise all the listed numbers define the set of allowed fragment numbers. Peaks having assignments in at least one dimension to one of the specified fragments are displayed.

Atom name or number (e.g. HA 22):
> allows to specify a single atom. Peaks that are assigned in at least one dimension to this atom, are displayed. The atom can either be specified by its name (e.g. HA 22) or by its number listed in the atom list file (e.g., 1301).

Peak numbers (e.g. 234 1001):
> allows to specify single peaks by their peak numbers. The selected peaks are displayed. To zoom into a region containing a certain peak the "View peak [vp]" command can be used.

Assignments (e.g. HA HN):
> allows the user to specify atom names of an assignment in all dimension without specifying the fragment numbers. The wild cards `*' and `?' are allowed. The `*' stands for any valid assignment. The `?' also allows unassigned peaks. Examples are given below:

48

HA HN: selects all peak assigned in w1 to an HA and in w2 to an HN.
- * HN: selects all peaks with valid assignment in w1 that are assigned in w2 to an HN
- ? HN: selects all peaks assigned in w2 to an HN
- * *: selects all assigned peaks

Diagonal peaks (tolerance in ppm, e.g., 0.02):
> selects all peaks lying on the diagonal of the current view. The tolerance that has to be specified in ppm, gives the maximal allowed difference between the shift in the x direction and in the y direction. If the difference is bigger the peak is not selected.

Sequential range (intra and seq. = 0 1):
> selects peaks whose assignments lie in a specified sequential range. The lower and upper bound of this range are specified. If only one number is specified, it is taken as the upper and lower bound. For each peak, the maximal difference in the fragment numbers occurring in the assignment is calculated. If this difference lies within the specified range, the peak is selected. Examples are given below:

- 0: selects all intra-residual peaks
- 1: selects all sequential peaks
- 1 4: selects all sequential and medium range peaks

Asymmetric peaks (tolerance in ppm, e.g 0.02):
> selects asymmetric peaks. Only peaks that have no symmetric counterpart are selected. For a specified tolerance Dp a peak at the ppm position px in the x direction and py in the y direction is considered symmetric to a peak at position sx and sy if

$$\left| p_x - s_y \right| < \Delta p \;\wedge\; \left| p_y - s_x \right| < \Delta p$$

7.16.A

7.17 "Display all peaks [da]"

Redisplays all peaks. This command is used after a subset of peaks was selected with the "Select peak [sp]" command.

7.18 "List peak entries [le]"

The information stored for each peak is written out. An example is given below. The fields are those from the peak list described in section 3.4.1 on page 23. The listed assignments are generated by looking up the atoms in the atom list using the atom number stored in the peak list and by looking up the fragments in the fragment list using the fragment numbers stored in the atom list. On each line the following entries are listed: first is the peak number followed by the peak color, then the folded chemical shifts in all dimensions, subsequently the assignments. Each line ends with the integration information including the volume, the volume error and the integration method.

```
 No.  Color    w1      w2      w3      ass. in w1      ass. in w2      ass. in w3
Volume      Vol. Err  Method

  97     1    113.009   7.957   7.957    N   51 THR     HN   51 THR     HN   51 THR
0.000e+00   0.00e+00     -
 402     1    113.009   5.699   7.957    N   51 THR     HA   51 THR     HN   51 THR
0.000e+00   0.00e+00     -
 403     1    113.009   3.826   7.957    N   51 THR     HB   51 THR     HN   51 THR
0.000e+00   0.00e+00     -
 404     1    113.009   1.024   7.957    N   51 THR     QG2  51 THR     HN   51 THR
0.000e+00   0.00e+00     -
 663     1    125.271   5.699   9.133    N   41 VAL     HA   51 THR     HN   41 VAL
0.000e+00   0.00e+00     -
 694     1    113.009   4.452   7.957    N   51 THR     HA   50 ALA     HN   51 THR
0.000e+00   0.00e+00     -
 695     1    113.009   4.711   7.957    N   51 THR       inv. ass.     HN   51 THR
0.000e+00   0.00e+00     -
 696     1    113.009   0.455   7.957    N   51 THR     QB   50 ALA     HN   51 THR
0.000e+00   0.00e+00     -
 699     1    125.160   1.024   8.999    N   52 TRP     QG2  51 THR     HN   52 TRP
```

49

```
0.000e+00  0.00e+00       -
 700    1   125.160    3.826    8.999    N  52 TRP     HB  51 THR     HN  52 TRP
0.000e+00  0.00e+00       -
 703    1   125.160    5.699    8.999    N  52 TRP     HA  51 THR     HN  52 TRP
0.000e+00  0.00e+00       -
```

7.19 "In-phase peak picking [in]"

This command invokes the routines written by N. Schäfer (diploma thesis, ETH Zürich) for picking in-phase peaks. First, local maxima that lie above the specified noise level are identified. The number of local maxima can be reduced on the basis of criteria designed to distinguish noise and artifacts from peaks. The position of the remaining maxima can be adjusted by moving them to the maximum of a second order polynomial fitted to the data points adjacent to the identified maximum. The following parameters are used in the routine:

Base level:
> intensity level above which local maxima are considered as possible peaks.

Pick only:
> whether to pick only positive peaks (+), only negative peaks (-) or both (b).

Mini. linewidth at 1/2 height in w ... :
> criterion to eliminate local maxima that are too narrow. All local maxima with a full width at half-height smaller than the specified value (in data points) are discarded.

Max. delta fit:
> the integer coordinates of the local maxima are adjusted by at most the specified value towards the maximum of the fitted second order polynomial.

Dim_1, Dim_2, max_distortion:
> criterion to eliminate local maxima with asymmetric peak shapes. The curvatures, i.e., the second derivativs of the fitted second order polynomial, in the two specified dimensions Dim_1 and Dim_2 are compared. If the ratio of the two curvatures lies below max_distortion the local maxima is discarded.

Wiggle checking (0/1) in w ...:
> criterion to eliminate local maxima likely to be wiggles. For each dimension it must be specified whether local maxima likely to be wiggles should be discarded (1) or not (0). Local maxima are identified as wiggles if they are part of a pattern of local maxima consisting of a small maximum followed by a larger maximum of opposite sign and again a smaller maximum with the same sign as the first one.

Hyperplanes in w ...:
> criterion to eliminate local maxima in noise bands. For each dimension of the spectrum a threshold has to be entered. The number of local maxima found at each position in each dimension of the spectrum is counted. If for a given position this count is larger than the specified threshold all local maxima at this position are eliminated.

Delete hyperplane with one peak:
> criterion to remove local maxima at positions where no other local maximum is found.

Remove diagonal peak with offset:
> criterion to remove local maxima close to the diagonal. All local maxima whose position in the different dimensions differ by less than the specified value are removed.

7.20 "Anti-phase peaks [an]"

This routine automatically picks peaks with a COSY type symmetry in the selected region. The routine works by looking for peaks which have symmetry of the form: The criterion for the selection of a peak is based on two parameters; it's height and symmetry.

Four parameters need to be entered before this routine can be used:

1) A base level - this corresponds approximately to the noise level in the spectrum although its worth trying a number a factor of 2 above or below to see if performance is improved.

2), 3) Peaks are found by calculating the symmetry about each point in the spectrum. The region considered in the symmetry calculation is set by the next two parameters which define the range of square sizes to search. The meaning of these numbers is indicated in figure 7.19b.

Typically a range of window sizes of from 2 to 3 points is adequate. The larger the range the longer the pick will take. Window sizes of 1 data point are not recommended as too many noise peaks will be found.

4) The final parameter selects the symmetry sign. The peaks in figure 7.19a have, from left to right, sign -1 and +1.

Where possible the routine will centre peaks correctly. However, in the case of distorted peaks, this may not be possible and these peaks are given a different color and their positions should be corrected manually.

Complex multiplets will be picked more than once as they contain a number of simple symmetry centres. They will therefore need to be corrected manually.

7.21 "Editing peaks [ed]"

Allows the user to edit the information associated with a selected peak. After selecting a peak the peak editing window pops up. The displayed fields can be edited. The corresponding lists are updated according to the performed changes. The upper panel of the peak editing window contains the information directly stored with each peak. The lower panel gives the information referenced by the atom numbers and stored in the atom and fragment lists. Changes in the upper panel affect only the peak list while changes in the lower panel affect the atom and fragment lists. Following is a description of the various fields and the changes occurring in the different lists when the fields are edited:

Peak nr.:
> gives the unique peak number of the peak. By setting the field to the number of a valid peak the information of the corresponding peak is displayed. If the entered peak number does not exist a warning message pops up.

w1, w2, w3, w4:
> give the folded position of the peak in ppm. Entering a new value will shift the peak to the indicated position. If this position lies outside the spectrum the peak is folded into the spectrum and the folding value is set accordingly. Thus it is possible to set the folding value of a peak by entering the unfolded shift into these fields.

Atom 1, Atom 2, Atom 3, Atom 4:
> atom numbers for each dimension. They give the assignment of the peak. The value 0 indicates that the peak is not assigned in the corresponding dimension. The atom number is used to look up the information about the atom from the atom list. This information is then displayed in the lower panel of the peak editing window.

Color:
> gives the user defined color value of the peak. Allowed values are 1 to 6. Peaks with different color values are displayed with different colors. The colors associated with each number can be defined in the default window (refer to section 12.1 on page 91).

Volume:
> the volume of the peak. It can be set either manually by entering a value or calculated from the spectrum using one of the integration methods.

Method:
> one letter giving the integration method. The program uses the following convention but does not enforce it:

```
`-' no integration performed
`d' denk integration
`r' rectangular integration
`e' elliptical integration
`m' maximum integration
`a' automatic peak integration performed by the external program peakint
```

Comment:
> a user defined comment that is stored and can be displayed with each peak.

51

In the following four lines of the peak edit window the information for each of the assigned atoms is given. The fields are:

Shift:
>   gives the atoms unfolded chemical shift in ppm. Changing this value will only affect the value stored in the atom list. It can also be calculated from the assigned peaks using the "Average chem. shift [ac]" command.

Error:
>   gives the deviation from the mean of the chemical shifts of all the peaks assigned to this atom. An exceptional large value indicates that a peak is missassigned. The value is set by the "Average chem. shift [ac]" command. Changing it will only affect the value stored in the atom list.

Atom Name:
>   gives the name of the atom as listed in the atom list. Editing this field will update the atom name in the atom list. This causes the assignment of all the peaks assigned to this atom to change. Care must be taken that the atom list does not degenerate (i.e., that there are not two entries for an HB2 produced and no entry for an HB3).

Fragment:
>   gives the fragment type as stored in the fragment list. The fragment type is looked up in the fragment list using the indicated fragment number. Editing this field causes two changes: first, the name in the fragment list is changed. Second all atoms belonging to the new fragment type and not already present in the atom list, are loaded from the fragment library and added at the appropriate position into the atom list. These two changes ensure that the assignments of all peaks assigned to this fragment are correctly updated and that all the necessary atoms are present for further assignments using the "Assign peak [ap]" command. Entering a name of a fragment type not present in the fragment library will cause an error message to pop up. "XEasy.new_atom_numbers: false" must be set to support this mechanism.

Fragment nr.:
>   gives the fragment number of the atom as stored in the atom list. This fragment number is used to reference the fragment in the fragment list. Changing it affects only the number stored in the atom list. This causes the assignments of all the peaks assigned to the atom to change. Care must be taken that the atom list does not degenerate (i.e., that there are not two entries for an HB2 of fragment 23 produced and no entry for the HB2 of fragment 77).

Mapping:
>   gives the mapping number of the fragment. This number is stored in the fragment list. When working with spin system fragments the mapping number indicates the amino acid residue number to which the spin system corresponds. The mapping number can be exchanged with the fragment number using the "Switch fragment/map numbers [sn]" command. This can be used to switch from assignments to spin system fragments to assignments and to amino acid residue fragments.

Frag. Com.:
>   a user defined comment. It can be used to hold possible spin system types when working on the spin system assignment or possible sequential neighbors, when working on the sequential assignment of a protein.

7.22 "Set selected peak entries [pe]"

Allows the user to change the information of all peaks lying in a selected region. A dialogue pop up window appears where the different fields of the peak can be specified. After leaving the pop up window the region containing the peaks can be selected using the left mouse button. When combining the command with the "Apply to all ... [aa]" modifier command it applies to the whole spectrum and it is not necessary to select a region.

The fields of all the peaks within the selected region are set according to the specification made in the pop up window. Leaving a field in the pop up window empty indicates that the corresponding peak field should not be changed. Using this command it is possible to set the folding value as defined above.

---

next: 8. Assignment Menu / contents

Looking

previous: 7. Peak Analysis Menu / contents

8.0 Assignment Menu

In the section on the Peak Analysis Menu, a command to make assignments was described where one could specify the atom names in each dimension. Here we present another method to make assignments based on the peak list and the atom list. By comparing the peak coordinates with known atomic chemical shifts, possible assignments can be produced for the peaks within the selected region. The user can make choices from the proposed assignments using the "Assignment window".

8.1 "Load assignment [la]"

The assignment list (refer to section 3.6 on page 34) to be loaded can be selected in the file window. The possible assignments are loaded onto the peaks present in the current zoom region. If no peak is loaded corresponding to the one in the assignment file a warning message is given. If for a loaded peak no assignment entry is present in the assignment file, the assignment possibilities for this peak are not altered. When there are more possible assignments listed than the program can handle, a message is displayed in the text window. As mentioned in section 3.6 on page 34, two different file formats exists. Both formats can be read, but only the XEASY format is written out.

8.2 "Write assignment [wa]"

The assignment possibilities associated with the peaks displayed in the current zoom region are written out to a file. The file name can be entered in the file window. The last extension is removed and the extension ".assign" is added. If no directory is specified, the file will be written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environment variable defining the groups directory and user_name is the users login name. Otherwise the specified directory is used. If a file with the same name exists already it is backed up by appending a `%' to its name prior to writing out the new assignment list.

8.3 "Possible assign. [pa]"

This command is used to find possible assignments for peaks within a selected region by comparing the coordinates of the peaks with the chemical shifts of the atoms. The chemical shifts of the atoms can either be defined by loading an atom list with defined chemical shifts ("Load atoms (chem. shift) [lc]" command) or using the "Average chem. shift [ac]" command. In each dimension a maximum of 10 assignments will be proposed. If more than 10 possibilities are found, the allowed differences between the peak coordinates and the chemical shift of the atoms is tightened. The corresponding peaks are reported in the text window.

This command treats only assignments in the current x and y direction. When the view of the spectrum is changed the command must be repeated in order to get valid assignment possibilities.

8.4 "Reduce to intrares. ass. [ia]"

In some spectra (e.g., [1H,13C]-COSY) the assignment possibilities produced with the "Possible assign. [pa]" command can be reduced by allowing only intraresidual assignments. Using this command after the "Possible assign. [pa]" command toggles all the interresidual assignment possibilities off. This may produce new unique assignments for a number of peaks. To transfer these assignments into the peak list the "Update assign. [ua]" command is available.

8.5 "Reduce using ass. peaks [ru]"

This command allows to restrict the assignment possibilities of an unassigned peak to those assignments

present in peaks at the same position in the specified dimension. In particular it can be used to reduce the assignment possibilities in heteronuclear correlated 3D spectra which were produced by the "Possible assign. [pa]" command or the ASNO program. In both cases the heteronuclear shift is not used to reduce the assignment possibilities in the correlated proton dimension.

To reduce the assignment possibilities of a given peak the program searches for all assigned peaks that lie close (within the specified tolerances) to the given peak. The assignment of this set of peaks are considered to be valid assignments for the given peak, all assignments not occurring in this set of peaks are considered to be invalid. For example for a heteronuclear correlated 3D spectrum, if only the tolerances in the heteronuclear dimension and the correlated proton dimension are set, the program will search for all assigned peaks lying in strips within the specified tolerances and restrict the assignment possibilities to those of these strips.

8.6 "Editing assign. [ea]"

The assignment possibilities (see also Possible Assignments) for a selected peak can be viewed and edited. The peak is selected by clicking the left mouse button close to its picked position. The possible assignments for this peak in the x and y direction will be listed in the "Assignment window".

Clicking the left mouse button in the assignment window toggles the corresponding assignment possibility on and off respectively. Holding down the Shift key when pressing the left mouse button turns the corresponding assignment possibility on and all the other assignment possibilities off. This option is useful to select a unique assignment out of the numerous possibilities. If only a single assignment possibility is activated in the matrix, a unique assignment has been made and is then copied to the peak list.

8.7 "Report ass. stat.[ra]"

Checks the peak list and the chemical shift list in order to find assignment errors.

The number of assigned and unassigned peaks are counted and reported in the text window.

Duplicated peak assignments are reported. A peak assignment is considered duplicated if there exist two peaks that have the same assignment in all dimensions. Peaks with one or more unassigned dimensions are not considered. Atoms that have a large chemical shift error are listed. The chemical shift error can be calculated with the "Average chem. shift [ac]" command. A large chemical shift error indicates that there are peaks assigned to this atom which lie at different positions in the spectrum. This can result from incorrect assignments or from improperly picked peak positions.

8.8 "Update assign. [ua]"

Compares the cross peak assignments from the peak list with those from the assignment list. The command is useful when working with an assignment list generated by another program (e.g., ASNO). Different possibilities exist to update one of the two lists:

Compare:
> neither of the two lists is changed. The number of unassigned peaks that have an unambiguous, unique assignment in the assignment list is counted and reported. All assigned peaks are checked. If their assignment does not appear in the list of assignment possibilities from the assignment list, the peak is reported as contradictory in the text window.

Update peak assignments:
> all unassigned peaks for which there is a unique assignment in the assignment list are updated. Peaks assigned in at least one dimension are not changed by this option. The peak numbers of the newly assigned peaks are reported in the text window.

Overwrite peak assignments:
> all peaks for which there is a unique assignment in the assignment list are updated. With this

55

option, assignments present in the peak list are overwritten. The number of peaks for which the assignment has changed is reported in the text window. The peak numbers of the peaks which changed are also reported.

Update assignment list:

for each assigned peak, the possibilities in the assignment list are updated. If the atoms referenced by the peak assignment are not present in the assignment list, nothing is changed. Otherwise, all assignment possibilities, except the one specified by the peak assignment are turned off.

Looking

Integration Menu

In this section several methods for integration will be described. They range from the least sophisticate technique, which simply takes the intensities at the peak positions, to the most rigorous techniques, such as line-shape integration.

Load reference list [lr]

The list of reference line-shapes (refer to Integrals section in XEASY Model) to be loaded can be selected in a file window. The line-shapes are loaded onto the loaded atoms. If an atom for which a line shape is specified, is not present, a warning message is given and the command is aborted. In XEASY, reference line shapes are defined using reference peaks. In the old format from the EASY and EASY3D programs, reference lineshapes are defined by reference atoms. It is possible to read the old format. The program tries to identify the corresponding reference peak for both reference atoms. All cases where no reference peak can be identified are reported in the text window. For these atoms the reference line-shape remains undefined.

Write reference list [wr]

The line-shapes are written out to a file. The file name can be entered in a file window. The last extension is removed and extension ".ref" is added. If no directory is specified the file will be written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environment variable defining the groups directory and user_name is the users login name. Otherwise the list is written into the specified directory.

If a file with the same name already exists, a backup copy is made by appending a `%' to its name prior to writing out the new reference line-shape list.

Select y reference line-shape [sy]

Defines the extend of the reference line shape of an atom in the vertically displayed y direction. A rectangle can be drawn around an isolated peak which is assigned to the atom in the y direction. The size of the rectangle in the y direction will be taken as the line-shape extent.

Select x reference line-shape [sx]

As the "Select y reference line-shape [sy]" command but defines the line shape extend in the x direction.

Select x & y ref. line-shape [sb]

Both the horizontal and vertical dimensions of the rectangle are used to define the line-shape extent.

The defined line-shape extents may be viewed by setting the peak display mode entry in the defaults window to line-shape. To see the line-shape for a peak which is not a reference peak you need to use the option line-shape displayed for all peaks.

When selecting line-shape extents, you should try and go as close as possible to the noise level, otherwise base-plane offsets will be present in the integral.

Find reference automatically [fr]

An alternative to the manual line-shape selection described above, is this routine which finds line-shape

57

extents automatically. To use the function you must define a lower limit for the line-shape. Setting it too low causes the routine not to work properly. Usually just above noise level works out. The user should also define a region, in ppm, about the diagonal in which not to choose reference peaks, since the routine uses the most intense peaks as line-shape sources and those around the diagonal often suffer from severe baseline distortion. The program will also avoid peaks which do not have symmetrical line-shapes. Selecting the region in which line-shape extents are to be found starts the routine. At the end of the procedure the percentage of peaks which can now be integrated is reported (60-70% is typical). The remaining line-shapes must be chosen manually as isolated reference peaks could not be found.

Good line-shape selection is critical if the integration routine is to function reliably. Following the automatic search line-shapes should be carefully checked to make sure there are no obvious errors. If they are not good they should be corrected manually.

Reset ref. line-shape list [re]

This command removes all reference line-shape extends from memory.

Remove reference line-shape [rm]

This command enables the user to remove the line-shape extends for a selected peak. The peak and the dimension in which to reset the line-shape extend must be selected. If the peak is unassigned the command is aborted. The line-shape(s) of the atom(s) assigned to the peak in the selected dimension(s) is reset. The volumes of the other peaks that depend on the deleted line-shape(s) are reset. Only the volumes of peaks that are line shape integrated (integration method set to `d') and assigned to the atom whose line shape extend has been reset are set to zero. Peaks integrated by other methods or peaks not assigned to one of the atoms are not changed.

Check ref. line-shape list [cr]

Checks the reference line-shape list for consistency. Invalid line-shapes are either corrected or if this is not possible reset. The volume of peaks integrated by line-shape integration or non integrated peaks with invalid line-shapes are reset if the option "Reset invalid line shape integrated volumes" is selected. Atoms with valid chemical shifts but no defined line-shapes are reported. The different reported lists are described below: First, a list of all peaks used as reference peaks but not present in the peak list is reported.

```
>   Reference peak #1449 does not exist (atom HA 54 in w1)
>   Reference peak #1449 does not exist (atom HB2 54 in w2)
>   Reference peak #1644 does not exist (atom HB3 54 in w2)
>   Reference peak #1566 does not exist (atom HN 55 in w1)
>   Reference peak #1566 does not exist (atom HA 55 in w2)
```

Second, a list of all reference peaks that lie outside of the line shape extent is given. This may happen if there are missassignments, if the peaks have been moved, or the chemical shifts were not updated using the Average chem. shift [ac]" command described on page 85. The line-shape defined by the peaks appearing in this list are invalid and are reset.

Third, the assignment of the reference peak is checked against the atom for which the line shape is defined. All inconsistencies are listed and the corresponding line-shapes are reset. It is assumed that although the assignment of a peak may have changed, it still remains separated and defines a line-shape. So if the new atom assigned to the peak has no valid line-shape, the peak is used to define the line-shape for this atom.

Fourth, all atoms with a valid chemical shift but no line shape defined are listed.

```
>   Atoms with valid chemical shift and no line-shape
      #5   :    QB    1 in w2
      #11  :    CZ    1 in w1 w2
      #12  :    CG    1 in w1
      #14  :    HN    2 in w1
```

58

```
        #15 :    HA   2 in w1 w2
        #16 :   HB2   2 in w2
```

Fifth, all peaks that were line shape integrated or are not integrated (integration method set to `-' or `d'), yet have invalid line-shapes are listed. If the option "Reset invalid line shape integrated volumes" is selected the volumes of these peaks are reset.

```
>   Peaks with invalid line-shapes
     1      3      6      7      8
     9     10     12     14     15
    16     17     20     21     22
    23     24     25     26     27
    30     31     32     33     36
    37     39     42     43     44
```

Line-shape integration [li]

This uses the method described in section 3.8 on page 42 to integrate both overlapping non-overlapping peaks. Select a region containing those peaks to be integrated. The routine first identifies peaks which overlap and then applies line-shape integration (also known as Denk integration) to this cluster.

Integration will fail for a peak with an undefined line-shape, for unassigned peaks, if the line-shape is less than 3 data points wide or greater than 200 points (unlikely), if the line-shape extent falls outside the selected region or the line-shape extent and the atom position are inconsistent. If desired, the results of the integration are listed in the text window and displayed on the screen. The peak volume entry will be updated with the calculated integral only for those peaks with the integration method set to `d' or `-', i.e., those peaks that are not manually integrated. Note that occasionally you may get a negative integral for a small peak which is overlapped by much larger peaks. This arises due to small differences between the selected and the actual line-shape for the peaks involved and is largely unavoidable. One solution is to try changing the source of the offending line-shape(s).

If your spectrum has a base-plane offset you can account for it by answering `1' to the question apply base-plane correction when integrating? and entering the offset in the popup window which appears at the start of this command. Note that this is really only reliable if the base-plane offset is consistent over the entire spectrum. Normally its use is only recommended if applied to individual clusters where you have specific integration problems.

Apart from the volume of the peak the program also calculates a volume error. This is the root mean square difference between the simulated cluster obtained by combining line-shapes and volumes, and the spectral data. The calculation is limited to the region defined by the peaks line-shape extents. The resultant error is expressed as a percentage of the volume for that peak.

An example of output from the program is given below.

```
>    Peak 974 misses line-shapes in w1 or w2
>    Peak 983 is not unambiguously assigned
>    Peak 1126 is not unambiguously assigned
...
>  Cluster 1:
>    Peak  114  :   HG3  40 /  HD2  40: volume = 8.52e+03 : error = 2.11%
>  Cluster 2:
>    Peak  115  :   HG2  40 /  HD2  40: volume = 1.15e+04 : error = 1.71%
>  Cluster 3:
>    Peak  116  :   HB3  53 /   HA  53: volume = 5.09e+04 : error = 0.59%
>    Peak  123  :   HB2  40 /  HD3  40: volume = 6.25e+03 : error = 1.61%
>    Peak  795  :   HB2  40 /  HD2  40: volume = 5.22e+02 : error =29.60%
```

Interactive integration param. [ip]

This command allows the user to the set the parameters used by the interactive integration functions which are described in the following section.

Five parameters can be set, the integration mode (i.e., elliptical, rectangular, automatic or maximum) the base contour level below which data points will not be considered in the integration, a base plane offset

59

correction which will be subtracted from each data point, whether the data points integrated should be marked and how many planes should be integrated in 3D spectra.

Interactive integration [ii]

With exception of the maximum integration mode and the automatic mode, you can integrate peaks by selecting the region of interest about each peak. The type of integration carried out depends on the parameters set above. The default settings are rectangular, a base level of -9999 (i.e., all points are integrated) and zero base-plane offset. The results of the integration will be reported in the text window and drawn on the screen. Note that this method of integration is preferred over line-shape integration in those cases where the peaks are isolated, as it gives more accurate results, or where peaks are unassigned or unpicked.

The rectangular mode integrates all data points above the base contour level that lie within the selected region. If a single peak lies within the selected region its volume entry is updated and the integration method set to `r'. The elliptical mode works as the rectangular mode with the exception that only data points lying in an elliptical region are integrated. The size of the axis of the ellipse is defined by the extends of the selected region. The integration method entry of the integrated peaks is set to `e'.

The automatic mode checks each peak within the selected region to see if it is isolated. If it is then a rectangle is defined about the peak at the base level set in "Interactive integration param. [ip]". All points inside this rectangle above the base level will contribute to the final integral. This mode is only useful if you have little or no overlap in your spectrum (3D data sets for example). The integration method entry of the integrated peaks is set to `r'.

If maximum integration is selected the volume of all peaks in the selected region is set to the spectral intensity at the position of the peak. The integration method entry of the integrated peaks is set to `m'.

Integrate rectangular regions [ir]

Sometimes one wants to integrate the same regions in a number of spectra. This command allows you to do this. First define the regions to be integrated by drawing rectangles about them (use the "Draw rectangles [dr]" command). Save the rectangle list when finished. Define the integration parameters with the "Interactive integration param. [ip]" command and then select this option. Each region defined by a rectangle will be integrated and the results listed in the text window. If a single peak lies within the integrated rectangle its volume field is updated with the calculated integral. The integration method entry of the integrated peaks is set to `r'.

The procedure can then be repeated for all remaining spectra.

Integrate strips (spscan) [is]

Integrate peaks by fitting them to a gaussian or lorenzian lineshape. The command saves the current state of the peaklist to a file "*peaklist*_1.peaks" and writes a macro "xeasy_temp.spm". This macro is used by spscan to integrate the peaks with the parameters in "xeasy_int.scpa" and write the result to "*peaklist*_2.peaks". This list is read back to spscan.
Although the use of these commands is quite convenient I recommend to do these steps "by hand", and adjust the integration parameters accordingly. (see spscan: how to integrate) The default parameters in "xeasy_int.scpa" or "xeasy_pick.scpa" may not always give optimal results.
When you adapt the parameters in "xeasy_int.scpa" or "xeasy_pick.scpa" to your spectrum (with spscan: "integrate"/"parameters") consider the following:

- load the spectrum in orientation: xeasy x=1, indirect=2, y(strips)=3.
- if the distance between two planes is larger than the natural linewidth, set the linewidth Lw+min = Lw_e = Lw_max = 0.6 ... 0.8 x ppm distance between planes.

60

The file "xeasy_int.scpa" has to be in the directory defined by Spscan*scpa_path: or Spscan*path:, otherwise the command will fail. Processing of 3D spectra is time-consuming and requires a computer with large memory.

Peak-pick strips (spscan) [ps]

Same as above, but the existing peaks are only used to define strips along which all peaks are picked that meet the conditions defined in "xeasy_pick.scpa". For 2D spectra both commands are identical - spscan integrates only 2D peaks that are already picked.

View reference peak [vr]

After this command, the user is expected to select a peak, and to define the dimension to be taken. The program goes through the peak list to find all the peaks with the same assignment in the defined dimension as the selected peak. At these positions, slices are taken and displayed in the slice window. The slices are labeled with the assignment of the peaks.

/ contents

Looking

10.0 Atom List Menu

In the old EASY and EASY3D programs, the atom list is called the proton list. For heteronuclear spectra atoms other than protons are added.

Load library [ll]

Loads the fragment library (see also Fragment Library on page 30). The atoms occurring in the different type of fragments are defined in this library. The name of the library file can be selected in the appearing file window. By default the program reads the library defined by the environment variable XEASY_LIB.

Load sequence [ls]

The list of fragments (see also Fragment List on page 29) which can correspond to the amino acid sequence, is loaded. A new atom list corresponding to the read in fragments are generated: for each fragment in the list, the corresponding atoms are read from the fragment library and appended to the atom list. If an atom list other than the automatically generated default atom list shall be used, it must be loaded after the fragment list with the "Load atoms (chem. shift) [lc]" command.

Load atoms (chem. shift) [lc]

The list of atoms (see also Atom List on page 28), which can correspond to the proton list used in the old EASY and EASY3D programs, is loaded. The existing atom list in memory is replaced by the new one. Care must be taken that the fragment list is loaded ("Load sequence [ls]") before the atom list. Otherwise the loaded atoms are overwritten.

Average chem. shift [ac]

Calculates the chemical shift of the atoms from the peaks assigned to them. The average and the deviation of the unfolded positions of the peaks assigned to a particular atom is calculated and stored in the corresponding entries of the atom. If an atom is not used for any peak assignment, the average can not be calculated and it is set to 999.000 (meaning undefined). If an atom is used only in the assignment of a single peak the deviation of the chemical shift is not defined and is set to 0.0.

Usually only those peaks from the peak list displayed in the region where the command was invoked are considered. To calculate the average chemical shift of peaks from several peak lists, the "Apply to all ... [aa]" modifier command can be used. In this case the average is calculated using the peaks of all loaded peak lists.

Change new chem. shifts [nc]

This command works very similar to "ac", but atoms for which no chemical shift is found in the peaklist(s) are left with their previous shifts and errors. If shifts can be extracted from the peaklist(s), only these shifts are considered, and the previous value is not used for calculating the average.

Switch fragment/map numbers [sn]

The fragment numbers and mapping number are consistently exchanged in the atom list and the fragment list. First the atom list is updated. For each fragment with defined mapping number (see also Fragment Library on page 30) all the atoms are updated by replacing the fragment number by the mapping number. Second, the fragment list is updated by exchanging the fragment number and the mapping number. This

procedure ensures that the atoms before and after the exchange of the numbers still belong to the same fragment.

The command is useful to switch from assignments to spin system fragments to assignments to amino acid fragments. Care must be taken not to use the same fragment number for different fragments (see section [From Spin System Assignments to Residue Assignments](#)). The command can only be used correctly if "XEasy.new_atom_numbers: false".

Add new fragment [af]

Adds a new fragment to the fragment list and the corresponding atoms to the atom list. The fragment type and the fragment number of the new fragment must be entered. If the fragment number is already present in the fragment list, the fragment type is changed and the atoms of the new fragment type me added to the atom list. No additional fragment is appended in this case.

If the fragment type is not present in the fragment library (see section [Fragment Library](#)), the command is aborted.

The modified atom and fragment list are sorted by increasing fragment number. The order of the atoms within one fragment is not changed.

Write sequence [ws]

All the fragments from the fragment list are written out to a file. The file name can be entered in the file window. The last extension is removed and extension ".seq" is added. If no directory is specified the file will be written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environment variable defining the groups directory and user_name is the users login name. Otherwise, the list is written into the specified directory. If no mapping number is specified, only the fragment type and the fragment number are written out. Otherwise the fragment type and the fragment number are followed by the mapping number.

If a file with the same name already exists, it is backed up by appending a `%' to its name prior to writing out the new fragment list.

Write atoms (chem. shift) [wc]

All the atoms from the atom list are written out to a file. The file name can be entered in the file window. The last extension is removed and extension ".seq" is added. If no directory is specified the file will be written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environment variable defining the groups directory and user_name is the users login name. Otherwise, the list is written into the specified directory.

If a file with the same name already exists, it is backed up by appending a `%' to its name prior to writing out the new atom list.

[next: 11. Multiple Quantum Menu](#) / [contents](#)

Looking

11. Multiple Quantum Menu

The commands in this menu allow to treat reduced dimensionality experiments as for example the one published in Szyperski, T., Pellecchia, M. and Wüthrich, K. (1994) J. Magn. Reson. B 105, 188-191. In this sort of spectra peak pairs are observable whose positions correspond to the single and double-quantum coherences of two nuclei. If d1 and d2 denote the positions of the two peaks, the chemical shifts, n1 and n2, of the involved nuclei can be calculated according to

$$\nu_1 = \frac{\delta_1 + \delta_2}{2}$$

11.0.A

$$\nu_4 = \delta_0 \pm \frac{f}{2}(\delta_1 - \delta_2)$$

and   11.0.B

In these equations d0 is the offset and f is 1/kappa * sf_c/sf_s, where sf_c is the frequency for the nucleus that defines the center of the doublet and sf_s is the frequency for the nucleus that defines the splitting (Szyperski, T., Pellecchia, M. and Wüthrich, K. (1994) J. Magn. Reson. B 105, 188-191). In the subsequent explanation we will refer to n1 as to the mean frequency and to n2 as to the difference frequency of a given peak pair. Since it is a priori not known which of the two peaks of each pair is the single-quantum peak and which is the double quantum peak, the difference frequency is not uniquely defined and special care has to be taken to determine the correct shift.

"Pick multi quantum peaks [pi]"

The command allows to identify and assign pairs of peaks whose positions in one dimension are d1 and d2. As result of the command entries for the zero- and double-quantum coherences are added to the atom list (which in this context should rather be called coherence list), the two peaks are picked and they are assigned to the new entries in the atom list. To use the command one of the peaks should be picked already. Its assignments and folding values are used as default for the second peak of the pair. After invoking the command the picked peak must be selected by clicking the mouse close to it. Subsequently the position of the second peak of the pair is defined by clicking the mouse at its position. In the appearing dialogue box the assignment for the peak pair can then be specified in a similar manner as with the "Assign peak [ap]" command. For the assignment in the dimension with the zero- and double-quantum frequencies two atoms must be defined, whose chemical shifts correspond to the mean and to the difference frequencies respectively.

Identified and assigned peak pairs can be used together with the "Generate average peaks [ga]", "Generate difference peaks [di]" and "Generate [diff,avg.]-COSY [co]" commands to generate peak lists with single quantum peaks of the atoms involved in the zero and double-quantum coherences.

After having used this command the atom list has changed and should be saved ("Write atoms (chem. shift) [wc]")!

"Pick and average peak pair [av]"

The command allows to identify pairs of peaks whose position differ only in one dimension. The command generates an additional entry in the peak list at the averaged position of the two peaks. To use the command one of the peaks of the pair must have already been picked. The assignment and the folding values of this peak are copied to the new peaks in a similar manner as with the "Copy and move peaks [cm]" command. After invoking the command the picked peak must be selected by clicking the mouse close to it. Subsequently the position of the second peak of the pair is identified by clicking the mouse at

64

its position. The program then adds two entries to the peak list, one for the second peak of the pair and one at the averaged position. To allow subsequent selection ("Select peak [sp]") of all the averaged peaks their color is set to two.

"Generate average peaks [ga]"

A new peak list is generated with one entry for each identified and assigned peak pair (see "Pick multi quantum peaks [pi]"). As a parameter for the command the name of the new peak list must be specified. The peaks in the new list are positioned in the multiple quantum dimension at the average frequency n1 (equation 11.0.A) and they are assigned in this dimension to the atom with chemical shift corresponding to this frequency. In all the other dimensions the positions and assignments of the peak pair is copied to the new peak.

"Generate difference peaks [di]"

A new peak list is generated with one entry for each identified and assigned peak pair (see "Pick multi quantum peaks [pi]"). The peaks in the new list are positioned in the multiple quantum dimension at the difference frequency n2 (equation 11.0.B) and they are assigned in this dimension to the atom with chemical shift corresponding to this frequency. In all the other dimensions the positions and assignments of the peak pair is copied to the new peak. As parameters the name of the new peak list and the parameters f (Scaling factor) and d0 (Carrier position) must be specified.

In order to resolve the ambiguity of the difference frequency a reference spectrum can be specified which must have peaks at the mean and the difference frequencies of each peak pair. Beside the name of the reference spectrum, the dimensions which correspond to the mean and average frequencies and allowed tolerances between the calculated mean and difference frequencies and peaks in the reference spectrum must be specified. If a reference spectrum is specified new peak entries are only generated for peak pairs for which a reference peak is picked in the reference spectrum.

"Generate [diff,avg.]-COSY [co]"

A new 2D peak list is generated with one entry for each identified and assigned peak pair (see "Pick multi quantum peaks [pi]"). The peaks in the new list are positioned at the difference and the average frequencies n1 and n2 (equation 11.0.A and equation 11.0.B). Due to the ambiguity of the difference frequency for each peak pair two peaks are generated. As parameters the name of the new peak list and the parameters f (Scaling factor) and d0 (Carrier position) must be specified.

"Generate 4D peaklist [wm]"

A new peak list is generated as described above, which contains both the "average" and the "difference" entries. In addition, a range of chemical shifts for the "difference" entry can be given: If no reference peaklist is given or if the position cannot be resolved from the reference peaklist, then only the position(s) in this range is(are) written out.

After you have done that you have to "xc" to the list that you have created, and "wp" it to disc. When you repeat the command with the same name for the peaklist, you add peaks rather than creating a new list. "ep" the list you have yust created before repeating with different parameters.

Symmetrization commands

The following commands symmetrize the displayed contour plot along a horizontal axis. Intensity plots are not affected by the command. Symmetrization is done according to the following rules: 1) The axes of symmetry is the center of a data point.
2) Points for which the symmetric point has opposite sign, or is outside the spectrum, are set to zero.
3) All other points are set to the minimum of the intensities of the two symmetric points, or if both are

XEASY manual: Chapter 11

negative, to the maximum.

"Strip symmetrization [sm]"

Strips are displayed as symmetrized in y-dimension to the position of the peak that was used to construct the peaklist (independent of the peaklist that is currently loaded). The command has no immediate effect, it is realized with the next command that plots strips (e.g. gs, fs, bs). In the meantime, somewhat misleading, "working ..." status is indicated.

"Reset strip symmetrization [ms]"

Resets the above.

"Zoom region symmetrization [zn]"

When you click into a spectrum, the intensity plot is immediately symmetrized to the y-position of the cursor.

"Reset zoom symmetrization [nz]"

Resets the above.

/ contents

Looking

66

12.0 Windows Menu

12.1 "Default window [dw]"

Displays the default window. Using the command a second time will remove the default window from the screen. The first five options listed in the default window define the colors used for the different items. Possible colors are White, Magenta, Red, Yellow, Green, Cyan and Blue. Those items whose color can be specified are:

Contour +
> positive contour lines

Contour -
> negative contour lines

Grid
> grid drawn in the spectrum

Lines
> user defined lines

Rectangles
> user defined rectangles

The following section of the window contains a maximum of six buttons where the peak colors can be defined. Possible peak colors are White, Magenta, Red, Yellow, Green, Cyan and Blue. If the option "Peak cross displayed is" is set to "Various" the given choice is between different peak shapes. The possibilities are a square, a circle, a diamond and a plus sign. The buttons in this section of the default window depend on the criterium used to assign a color to a peak:

Peak color determined by:
> This option defines the criterium which is applied to assign colors to the peaks. The different criteria and the classes of peaks they define are given below:

> Color type:
> > the peak color is determined by its color field in the peak list.
> > Type 1, Type 2, Type 3, Type 4, Type 5, Type 6

> Assignment:
> > the peak color is determined according to the assignment state of the peak.

> x & y assigned
> > peak is assigned in the dimensions displayed horizontally and vertically

> y assigned
> > peak is only assigned in the dimension displayed vertically

> x assigned
> > peak is only assigned in the dimension displayed horizontally

> Multiple
> > peak has multiple assignments (i.e., atom number -1)

> Unassigned
> > peak is unassigned (i.e., atom number 0)

Spin sys:
> not used.

line-shape:
> the peak color is determined according to whether or not the line-shapes are defined. When this option is enabled, peaks are no longer displayed as small crosses, but also the extends of the line-shapes are indicated.

67

line-shape undefined
> one or more line-shapes of the peak are undefined

line-shape defined
> both the line-shapes in x and y direction are defined

Manual integration
> the peak is integrated with an other method than line shape integration (i.e., the integration method field is different from `d' and `-')

Integ. mode:
> the peak color is determined according to the integration method field of the peak.

Denk
> line-shape integrated peak, the integration method field is set to `d'

Elliptical
> the integration method field is set to `e'

Rectangular
> the integration method field is set to `r'

Invalid
> peak is not integrated, the integration method field is set to `-'

Others
> the integration method field is different from `d', `e', `r' and `-'

Plane:
> this criterium can be used in 3D and 4D spectra. The peak color depends on the distance between the peak and the displayed plane. The type 1 peaks are closest to the plane, the type 6 peak are furthest away from the displayed plane. The size of the interval for each of the types is defined with the option "Coloring interval". If it is set to 1 the different peak types have the following meaning:

Type 1
> within the displayed plane

Type 2
> in the plane above or below the displayed one

Type 3
> two planes above or below the displayed one

Type 4
> three planes above or below the displayed one

Type 4
> four planes above or below the displayed one

Type 5
> more than four planes above or below the displayed one

Fold:
> the peak color is determined by its folding. The following possibilities are distinguished:

Unfolded
Folded in w1
Folded in w2
Folded in w3
Folded in w4
Multiple folding
> the peak is folded in several dimensions

Peak cross displayed is:
> Normal displays all peaks as a small plus (+) signs with the option of using different colors, diagonal displays an `x' while various allows the selection of four different geometric symbols (square, circle, diamond and plus sign) all displayed in the same color. The latter option is useful for plotting.

line-shape displayed for:

68

determines which peaks will have line-shapes displayed when the line-shape option above is selected. Reference implies that only peaks which have been used to determine line-shapes will have the line-shape extents displayed, all means that line-shapes for all peaks will be displayed.

Peak color for peak pick:
determines the peak color attributed to new peaks that are appended to the peak list.

Maxima search for peak pick:
not used.

Residue code display mode:
allows either the conventional 1 or 3 letter residue codes to be displayed next to the cross-peaks when the assignment is displayed. A compact format displaying only the assignments in the two displayed dimensions is also available.

Maximal distance in planes:
this option is used in 3D and 4D spectra. It specifies how far the displayed peaks and the plane in a zoom region may lie apart from each other. If the distance of the peak to the plane is bigger than the specified limit, it is not displayed.

Coloring interval:
when the option Peak color determined by is set to Plane this option determines how fast the peak color changes with increasing distance from the displayed plane.

12.2 "Text window [tw]"

Pops up the text window. Using the command a second time removes the window from the screen. The window is used to display the log file which records the individual commands, some statistical results and error messages. In addition, it can be used as a normal editor. The contents of the text file is stored in the file "xeasy.text" in the easy directory.

The text window can be in one of two modes. In the Logfile mode it records the messages from the XEASY program. In the Editing mode no messages from the program are printed and the window can be used to edit arbitrary text files. The menus in the text window are activated with the left mouse button. In the File menu the command New allows to load a new text file for editing. The file name can be selected in a file window. The corresponding file is loaded and the text window switches automatically into the Editing mode.

The Save command allows the contents of the text window to be saved in a file. The name of the output file can be defined in a file window. No default extension is appended to the given name.

The Logfile and the Editing commands allow the user to switch between the two modes of the text window.

In the Edit menu the usual commands for cutting, copying and pasting are available. They function only within the text window. To get information out of the text window you have first to save its contents.

The commands in the Search menu allow the user to search and replace text. The text must be defined in the two fields Search Pattern and Replace Pattern.

12.3 "Color window [cw]"

This window is used for adjusting the color representation of intensity in the displayed spectra. There are two slide bars in the panel. The upper one controls the contrast and the lower one adjusts the threshold for displaying the intensities.

12.4 "Overview window [ow]"

This window is used to give an overview of the entire spectrum when working with zoom regions in the main window. The whole spectrum is shown and the zoomed regions in the main window are indicated by rectangles. There are several commands which can be activated by placing the mouse within this window and typing one of the following two-letter codes:

"mo"

> move one zoom region. A new zoom region can be selected in the overview window. If only one zoom region is currently displayed this one is updated to show the new zoom region. If several zoom regions are displayed, the one in which the cursor was located when typing the two letter code is updated.

"zb"

> zoom backward. This command will move one step back in the zoom history without changing the zooms displayed in the main window.

"zf"

> zoom forward. This command will move one step forward in the zoom history without changing the zooms displayed in the main window.

"zo"

> zoom. This command updates the zoom regions displayed in the main window to those currently shown in the overview window. It is normally used after the "zb" or "zf" commands.

12.5 "Peak data window [pw]"

This window is used for setting up the information displayed with the peaks, such as, peak number, volume and assignment. The following options are available:

```
Peak number      peak number field from the peak list
Coordinates      unfolded ppm coordinates of the peak
Color   peak color field from the peak list
Peak volume      peak volume field from the peak list
Volume error     volume error field from the peak list
Integ. method    integration method field from the peak list
Comment comment  field from the peak list
Atom numbers     assignment from the peak list
Atom names       names of the atoms with the above listed atom numbers
Assignment       depending on the option Residue code display mode in the default window
                 the assignment defined by the atom numbers is given in different
formats.
```

Either the atom name, the fragment number and the long (i.e., 3 letter) fragment type name are given or the atom name, the fragment number and the short (i.e., 1 letter) fragment type name. When compact is selected, only the fragment numbers and the atom names in the two displayed dimensions are given.

```
Mapping number   the mapping number of the assigned fragments.
Fragment number  the fragment number (which may correspond to the residue number)
                 of the assigned fragments.
```

12.6 "Peak editing [ew]"

This window is used for editing the peaks and their assignments. It is described in detail in the chapter about the "Editing peaks [ed]" command. The [ew] command can be used to remove the editing window from the screen or to fetch it from the background.

12.7 "Assignment [aw]"

This window is used for choosing the possible assignments for a selected peak. It is described in detail in the chapter about the "Editing assign. [ea]" command. The [aw] command can be used to remove the assignment window from the screen or to fetch it from the background.

12.8 "Slice window [sw]"

This window is used for displaying 1D cross sections (i.e., slices) out of the spectrum. Slices can either be selected in a zoom region or by using the "View reference peak [vr]" command. To select a cross section

in one of the zoom regions the dimension in which to take the cross section must be specified by selecting one of the toggles listed under Omega. Then the slice which should be replaced is defined by clicking a toggle listed under Current. Finally the position of the cross section can be defined by clicking the left mouse button at the appropriate position in the main window. The size and position for the new slice in the selected dimension is determined as follows: if no other slice from the same spectrum in the same dimension is displayed, the size and position of the zoom region where the new slice was selected are taken. If other slices from the same spectrum in the same dimension are displayed, the size and position of these slices are taken. To enforce the size and position to be taken from the zoom region, the SHIFT key must be hold down when selecting the slice in the main window.

A description of the function of the different buttons follows.

Current
>   Defines the slice to which the other commands apply. Allows to redefine the slice by clicking the mouse within one of the zoom regions.

Display
>   Defines the slices to be displayed in the slice window.

Omega
>   When selecting a new slice in one of the zoom regions, these buttons define the dimension from which the slice is taken.

Draw mode
>   Defines how the slices are displayed. `D' displays for each data point one dot. `L' draws the same dots and connects them by lines. `V' is not implemented.

Increment
>   Increments the position of the current slice (defined with the Current buttons) by one. For a slice selected in the displayed x direction the position in the y direction is incremented. For a slice selected in the y direction the position in the x direction is incremented. Slices perpendicular to the current view are not incremented.

Decrement
>   As Increment, but decrements the position.

Data
>   When the cursor is in the slice window, its position and the data intensity at the current position is indicated. The position of the cursor within the slice is shown in the main window.

Separate
>   Separates the slices in the vertical direction so they are individually visible.

Center
>   Centers all slices for line-shape comparison.

Scale
>   Scales all slices so they have the same height at the selected point and fit into the slice window. The point can be selected by clicking the left mouse button within the slice window. The scaling command is aborted by pressing the middle mouse button.

Reset
>   Resets the scaling applied with the Scale command. Afterwards the relative intensities of the slices correspond to the ones in the spectrum.

Add
>   Adds all the displayed slices and displays the sum in slice one. The scaling applied with the Scale or Reset commands is used.

Subtract
>   Subtracts the second slice from the first one. The scaling applied with the Scale or Reset commands is used.

Xmagnify
>   Allows to zoom into a part of a slice. The left and right border of the region to zoom must be selected by clicking the left mouse button in the slice window.

Xreset
>   Enlarges the displayed part of the slice. All data points are shown if the slice lies perpendicular to the displayed dimensions. Otherwise the smallest part including all the zoomed regions is displayed.

Yexpand

The vertical scale of all the slices is increased.

Yreduce

The vertical scale of all the slices is reduced.

Save

The first slice is saved to a text file which can be read by the external program infit. A file pop up appears where the file name can be defined.

Save All

The currently displayed slices are written to a file *.sla as ASCII data. They can be printed with ghostscript using the shellscript pslices, or imported by any grafics program.

Phase

Allows the interactive determination of phase correction parameters (see Phase Correction).

12.9 "Residue Type window [rw]"

This window is used to identify possible residue types based on sets of frequencies. The set of observed frequencies must be entered into the frequency fields. In order to distinguish different types of nuclei, e.g., 1H versus 13C or aromatic versus amide protons, frequencies may be grouped. The set of atoms included in each group is defined in the atom types field by their atom types - as defined in the fragment library. The frequencies can either be typed in, taken over from the frequencies stored in the atom list, or defined by clicking to selected peaks. After pressing the possible types button the possible types field contains the list of all possible amino acid types. The amino acid types that match best the specified frequencies are listed first, followed by those that are less likely. Finally using the statistics button likely residue types for all fragments present in the fragment list can be detemined.

A description of the function of the different buttons and fields follows.

Fragment

Field to enter the number of the fragment for which possible types should be determined.

Possible Types

Field containing the possible types found by the program. The most likely types are listed first.

Atom Types

Fields containing the types of the atoms that could correspond to one of the frequencies specified in the adjacent Frequencies field.

Frequencies

Field containing observed frequencies.

Select Peaks

After pushing this button peaks can be selected with the mouse. The frequencies of the selected peaks are appended to the frequencies listed in the Frequencies fields. For each selection only one frequency is added. To add all the frequencies of a peak the peak must be selected several times. The frequencies are added to the group of frequencies whose expected range matches best the spectral sweep width of the dimension from which the frequency was taken. Possibly occurring errors can be corrected by editing the Frequencies fields.

Use Atom List

The frequencies defined for the fragment specified in the Fragment field are copied from the atom list to the Frequencies fields.

Possible Types

A list of possible types is generated based on the frequencies specified in the Frequencies fields.

Statistics

For all fragments defined in the fragment list observed frequencies are defined from the frequencies in the atom list and possible residue types are determined. The results are written out to the Text Window.

The algorithm does not use assignment information within the fragment.

Looking

13.0 Geometries Menu

In the XEASY program lines and rectangles are called geometrical objects. They can be used for marking points, regions and frequencies in the two dimensions of the current view. Their coordinates (in ppm) can be saved and loaded with flags distinguishing the lines from the rectangles. The different routines to draw and remove geometrical objects are described below.

13.1 "Draw lines [dl]"

An open polygon can be drawn by clicking the left mouse button. The first click defines the starting point of the first line. Subsequent clicks define the end position of the previous and the starting position of the next line. The drawing is terminated by clicking the middle mouse button. Only horizontal and vertical lines can be drawn. The polygon is two dimensional and is projected on any plane in 3D or 4D spectra.

13.2 "Draw ss lines [ds]"

Draws lines extending from all the displayed peaks to the spectral diagonal in the two displayed dimensions. For example when using a COSY peak list and selecting a particular fragment with the "Select peak [sp]" command, the command draws the lines defining the corresponding spin system.

13.3 "Draw rectangles [dr]"

Allows the user to draw rectangles. The way of drawing a rectangle is exactly the same as selecting a region on the screen. The drawn rectangle may be used for rectangular integration (refer to "Integrate rectangular regions [ir]").

13.4 "Draw transpose [dt]"

Allows the user to draw transposed rectangles. The user is expected to select a point in a zoom region by clicking the left mouse button. A rectangle is drawn extending from the selected point to the spectral diagonal and then to the symmetric point on the other side of the diagonal.

13.5 "Draw freq. lines [df]"

Draws four lines marking the frequencies of the selected position. The position is selected by clicking the left mouse button, it defines two frequencies: one in the x direction and one in the y direction. At both of these frequencies in both of the displayed dimensions, lines are drawn extending over the whole spectrum. The command is especially useful in combination with the "Zoom geometries [zg]" command.

13.6 "Remove lines [rl]"

With this command lines can be removed. The user is expected to click the left mouse button close to the line to remove. To delete all lines at once, the command must be combined with the modifier command "Apply to all ... [aa]".

13.7 "Remove rectangles [rr]"

With this command, rectangles can be removed. The user is expected to click the left mouse button close to one side of the rectangle to remove. To delete all rectangles at once the command has to be combined with the modifier command "Apply to all ... [aa]".

13.8 "Load geometries [lg]"

The file to be loaded can be selected in the file window. The loaded geometries are displayed in the current zoom regions.

13.9 "Write geometries [wg]"

The geometries are written out to a file. The file name can be entered in a file window. The last extension is removed and extension ".geom" is added. If no directory is specified the file will be written into the directory XEASY_GROUP/user_name/easy, where XEASY_GROUP is the environment variable defining the groups directory and user_name is the users login name.

If a file with the same name already exists it is backed up by appending a `%' to its name prior to writing out the new list of geometries.

13.10 "Draw grid [dg]"

Grid lines are drawn on the screen. Issuing the command a second time removes the grid.

13.11 "Draw connectivities [cs]"

Draw lines emphasize connectivities between peaks. Lines are drawn between all displayed, assigned peaks that have the same assignment and are displayed at the same position on the screen. Two peaks are assumed to have the same position on the screen, if their screen coordinates differ by less than 20 points.

Looking

---

14. Miscellaneous Menu

Utility commands, not classified in other menus, are described in this section.

14.1 "Display unit(c/f/n) [du]"

Toggles the unit used to display the mouse position from ppm to Hz to data points.

14.2 "Draw cursor [dc]"

It enables or disables the drawing of a cursor extending over the entire main window. This is useful to compare the peak positions in the spectrum.

14.3 "Up and down mode [ud]"

Switches the display mode between horizontal arrangement and vertical arrangement of the zoom regions. Invoking the command once will increase the number of regions to display above each other by one.

14.4 "Measure noise [mn]"

After this command, the user can select a region. The average and the standard deviation of the intensities within the selected region is calculated and reported in the text window.

14.8 "Apply to all ... [aa]"

The command is used to modify the action of other commands. For example, using the "Goto right [gr]" command alone moves the zoom region where the cursor was when the command was invoked. Combining it with the [aa] modifier command will move all the displayed zoom regions. Or the "Average chem. shift [ac]" command alone averages the chemical shifts of the peaks assigned to the same atom. Using the command alone considers only peaks from the peak list displayed in the region where the command was invoked. Combining it with the [aa] command will average the shifts of the peaks from all peak lists loaded into memory.
The [aa] modifier can be used with the following commands:

- Replace spectrum [rs]
- Replace contour [rc]
- Enlarge vertically [ev]
- Enlarge horizontally [eh]
- Goto left [gl]
- Goto right [gr]
- Forward plane [fp]
- Backward plane [bp]
- Exchange peak list [xc]
- Select peak [sp]
- Display all peaks [da]
- Average chem. shift [ac]
- Draw ss lines [ds]
- Remove lines [rl]
- Remove rectangles [rr]

write xeasy.default [wd]

Dump the current state to the file "xeasy.default". This file is read from the startup directory when xeasy is started with the "-l" option. To save this state, mv it to another name, or to another directory. An automatic scheme dumps the state from time to time and overwrites "xeasy.default".

start spscan [xs]

The same as if you would type "spscan" on the shell, load the current spectrum and peaklist, and select a macro.
XEASY does not wait for SPSCAN to exit. see "is"/"ps"

Exit [ex]

This command terminates XEASY, and saves the current session into the binary file xeasy.default, which can be loaded by starting XEASY with the on line command option "-l".

next: Macro commands / contents

Looking

Macro Commands

Load Macro [lm]

Starts running an interactively selected *.xem macro file.
The command language is described below.

Go Back to Macro [gm]

When a macro has been interrupted by an "im" command, "gm" continues execution of the current macro. If no macro is open, the command is ignored. Dont use this command within a macro.

Interactive mode [im]

Suspends execution of a macro and allows to continue working in the normal interactive mode. The suspended macro is continued with "gm".

"im" can also be used outside a macro, if a macro fails du to an error, and leves the program in non-interactive mode.

Non-interactive mode [nm]

For testing purposes only. Switches to non-interactive mode without executing a macro.

Exit Macro [em]

Exit the current macro. If several levels of macros are open, the command must be repeated several times. If no macro is open, the command is ignored.

Command language

Files with macro commands should end with .xem. Macro files are read line by line, and interpreted as c-style strings. The first string on each line is regardes as a command, all following strings are regarded as arguments to this command. Arguments must not contain spaces. There are no characters with a special meaning for the macro interpreter.

All interactive commands are recognized, but some of them are handeled in a slightly different way. In addition there are some other commands which all start with an underscore (_). If a command is unknown, it is ignored. Thus lines that start with a # or $ can be used for comments.

Xeasy commands that do not require any further interaction by the user work exactly in the same way as

if they were entered from the keyboard.

Commands that normally lead to interactive selection of a filename (i.e. all load and save commands) require a filename as the first argument. Filenames can be absolute or relative to the default path valid for this type of file.

```
Example: ns /files/ralf/spectrum
         lp p1tocsy.peaks
```

Commands that normally cause a popup menu to appear, need additional arguments. For these commands arguments 2, 3, ... correspond to the string that is expected as an input to the 1st, 2nd, ... entry field. If the popup is used to toggle one or several fields, the arguments are "T" for on and "F" for off. Please consider that ususally only one argument can be true, and the others must be false. These commands must be used with great care, because the order, and the number of arguments to these commands can depend on the history of the program. If the command doesn't need a filename, the first argument is ignored. If a command causes two consecutive windows to appear, the arguments 2, 3, ... are used a second time, which is usually not intended. If not more than one argument is given, the program switches temporarily to interactive mode.

Commands that require mouse input, **are not adapted** to non-interactive use. The behaviour of these commands is undefined. Usually they will be ignored, because they are simply overwritten by the next command, as no mouse action is recorded.

Special cases:

Loading spectra
> Four arguments can be given to display the dimensions in x, y, or as planes ("-"). If no other argument but the filename is given, the spectrum is loaded in "standard orientation", which corresponds to " `- x y -` ". It is not possible to change offset and size of the displayed spectrum.

pm
> No filename, otherwise same as above.
> Example: `pm = - y x -`

Support for macros has been started in version 1.3 and is not yet complete.

Underscore commands

Be aware that these commands are not checked for unsuitable arguments. Proper use is impossible without some idea about the internal structure of xeasy.

_call

_call is used to call xeasy routines directly, which are normally not called by a two-letter command. As the routines may require different arguments, each subcommand is treated individually in void call_direct() (macro.c). The following subcommands are currently defined:

_call draw
> redraws the display. It is necessary to call this command after loading a peaklist, for example.
> Noninteractive commands do not automatically force a refresh of the screen.

_call test
> undefined behaviour. used for development only

_set

_set is used to modify global variables of xeasy directly. The routine is void set_global() in macro.c. The following subcommands are currently defined:

78

_set ptrhis 0..9

   set gPtrhis. With this command you can move backward and foreward in the zoom history.

_set parallel n

   set gHistory[gPtrhis], the number of spectra displayed in parallel

---

[next: User-Defined Resources](#) / [contents](#)

Last modified: Jan 30, 1996
Ralf W. Glaser, xeasy@mol.biol.ethz.ch

Looking

Setting User-Defined Resources

General X-Window defaults

The class name of xeasy is XEasy. The program recognizes some standard X_Window resources, but there is no systematic guide, which resources take effect in which windows.

Project-Specific Settings

In addition to the locations defined by the X-Window standard, a copy of XEasy in the startup directory is read, and meged in with the highest priority.

This allows you to overwrite the system-wide and the user-defined defaults with project-specific defaults, e.g. default paths for spectra, peaklists ...

Program-Specific Variables

*dont_use_default_colormap: false
>   If set to true, xeasy always creates a own colormap, even if there are enough cells free on default. Then all instances show the colors of the instance that has focus (on machines that support only one colormap)

*basecolors: 128
*keepcolors: 10
>   When xeasy opens a new colormap and if it is compiled with SINGLE_INSTALLED_COLORMAP=1 it copies basecolors read_only colorcells from the default colormap, then it allocates the read/write cells and immediately frees the read_only colorcells except keepcolors colorcells. You may want to adapt these values to minimize colormap interferences between programs.

*beep: off
>   for your nerves

*traditional_calibration: on
>   Xeasy has been interpreting the "Maximum chemical shift" information in parameter files as ppm of the very edge of the spectrum, i.e. the left side of the left pixel. *traditional_calibration: on supports this mode. "off" interprets the data as ppm in the center of the first pixel, what makes calibration independent of the spectral size in the frequency domain. Don't swith within a project, but I recommend to start new projects with "off". If you exchange data with other programs, check which mode is used there.

Modify XEasy according to your needs and copy it to your home directory. Alternatively, the resources can be specified in any other way (see any book about the X Window system to find out how else to specify application resources).

Default File Paths

Before release 1.3 the default path to search for spectra was $XEASY_GROUP$USER/easy/data/, and for all other files $XEASY_GROUP$USER/easy/.

From release 1.3 on, the directories ~/easy and easy/data are no longer necessary, default directories for spectra and data files can be specified anywhere on the system. You can specify different directories for all files with a different extension, according to the following rules:

- The default path for spectra (*.3D.16 / *.3D.8) and parameter files (*.3D.param) is the value of

80

"XEasy*spectra_path".

- The default path for all other files (*.*ext*) is the value of "XEasy*ext_path".
- All XEasy*ext_path are of class "path", so if XEasy*ext_path is not found, it defaults to the value of XEasy*path.

---

next: External Programs / contents

Last modified: November 28, 1995
Ralf W. Glaser, xeasy@mol.biol.ethz.ch

Looking

81

15. External programs

filegen2d

The program converts a two dimensional smx Bruker file (Aspect or X32) into a two dimensional XEASY file. The smx file must first be moved into your ~easy/data subdirectory. Th extension smx must be added to the file name on the workstation or smx+ if the file has been transformed on an Aspect computer and size extended to 32 bits on a X32 computer.

The spectrum can then be converted by typing the command

```
cd ~easy/data
filegen2d
```

Filegen will ask for the full file name, the sizes of the real spectrum (e.g., 1024 by 2048), where the data was transformed (X32 or Aspect), the submatrix size used (width and height), if transformed on an X32, and whether you want a 16 bit or an 8 bit file to be generated.

Since the submatrix size for files generated on the X32 can vary, you should first run the xau program foreasy on the X32 to obtain the necessary information.

To prevent data overflow you may need to scale the X32 data down. This can be done within filegen. The scale factor should normally be 8 as this will change the dynamic range from 32 to 24 bits. This is necessary as the 8 and the 16 bit files were designed for Aspect data.

filegen3d

This program converts a series of two dimensional Bruker files (AMX or X32) into a three dimensional XEASY file. The use of the program is analogous to the program filegen2d. The names of the Bruker files must be numbered continuously in front of the extension (i.e., for an experiment with name myexp the transformed 2D planes could be called `myexp1.smx, myexp2.smx, myexp3.smx, ...`).

cpx32

This program copies a two or three dimensional spectra from an Bruker X32 computer to the SUN workstation and converts it into the XEASY 8 bit format. The data on the X32 must be in the directory `/disc/data/user/nmr/filename/n/pdata/1` on the X32 computer (called station in the example below), where disc is the name of the hard disc, user is the name of the user, filename is the name of the spectra and n ranges from 1 to the specified number of planes nr. The names have to be given as parameters to the program:

`cpx32 station disc user filename nr` where nr stands for the number of planes. if nr is equal to one a two dimensional XEASY file is generated, if it is larger than one, a three dimensional XEASY file is generated. To convert a 3D spectrum that is in the 3rr format on the X32 computer, the nr parameter must be omitted:

`cpx32 station disc user filename`

cpx32_16bit

The program is used as the cpx32 program, but generates 16 bit XEASY format.

filecon2d

This program converts a two dimensional spectrum in the old EASY format into the new XEASY format. If a 16 bit file is available either a 8 or a 16 bit file may be generated. Otherwise only an 8 bit file is generated.

varian_xeasy2d

Program to convert varian 2D spectra into XAESY format. Oleg Berezhnoy / Christian Bartels.

varian_xeasy3d

Program to convert varian 3D spectra into XAESY format. Oleg Berezhnoy / Christian Bartels. The program is said to contain bugs.

convert3da

An old version of varian_xeasy3d, producing an outdated format, which can still be read by XEASY, however. Works reliable, in contrast to the program above.

converter

This program allows to transform varian and felix 2D and 3D file formats into XEASY file formats.

uxnmr_to_xeasy

uxnmr_to_xeasy converts 2D and 3D real spectra from Bruker UXNMR to XEASY format (Version 1/16 bit). Written by Ralf W. Glaser. It is used as

```
uxnmr_to_xeasy   <uxnmr 2rr/3rrr-file> <xeasy-filename>
```

where the first argument is the Bruker file, either in the current directory or with a full path. uxnmr_to_xeasy tries to extract the relevant information from procs (proc2s, proc3s) or proc files in the same directory as the data file.Normally uxnmr_to_xeasy works fully automatic, but if it cannot find the data it askes for interactive input. The program does not collect information from several text files: If it finds one of the procs or proc files, it assumes that it can find all the parameters there. The second argument is the name of the xeasy file. Two files are produced: filename.3D.16 and filename.3D.param. filename.3D.param is a textfile, which can be edited, if some of the extracted information is wrong or incomplete. uxnmr_to_xeasy does not change the (rather big) submatrix structures of the Bruker file.

Test releases:
0.2 (17.4.96)
- With option "-s" the program asks for the scaling factor (-s must be placed after the filenames)
- It is no longer necessary to have write permission on the UXNMR file
0.3 (22.7.96)
extracts ##$BYTORDP and does byte swapping if necessary.
0.4 (19.9.96)
extracts ##$BYTORDP from procs (Bruker feature, that data in proc and procs may be different)
0.5 (30.1.97)
float->xeasy16 from spscan - xeasy "-0" feature fixed: no more "spikes" in the spectra.

[pipe2xeasy](#) NMR Pipe to xeasy converter

flatt3d

The program is used for base line correction. It is an implementation of the flatt program (P. Güntert et al., J. Magn. Reson. 96,403-407, 1992) for higher dimensional spectra. The baseline correction is applied

83

to the original file. The user is responsible to make a backup of the original file as it will be overwritten. Flatt3d needs an additional parameter -r n m, where n defines the dimension in which the baseline correction is applied and m is used to define the diagonal for splitting the spectrum. For example

```
flatt3d -r 3 2 -d test
```

would perform a baseline correction on the spectrum `test' along w3 splitting the spectrum at the w2/w3 diagonal.

project

The program projects multidimensional spectra along one dimension. The dimension for the projection can be defined by the -d option. The norm used for the projection is defined by the -n option. If -n option is set to 1, the arithmetic mean of all the points in the projection dimension is calculated. If it is set to 2, the geometric mean is calculated. If it is set to a negative value, the maximal intensity along the projection dimension is determined.

Example:

```
project -d 3 -n -1 test test_w3_projection
```

Calculates the skyline projection along w3 of the spectrum called test, and stores it under the name test_w3_projection.

pick

The routine (N. Schäfer, diploma thesis, ETH) picks inphase peaks in multi-dimensional NMR spectra. First it locates all the local maxima that lie above the specified noise level. The number of local maxima can be reduced by a number of criteria designed to distinguish noise and artifacts from peaks. The remaining maxima can be positioned at the exact maximum by fitting a polynomial through the peak intensities.

Parameters for calling the program (n is the dimension of the spectrum) are:

```
pick  <specfile>         <peakfile>                filenames
      <start_point_1>...<start_point_n>            n integers
      <end_point_1>  ...<end_point_n>              n integers
      <noise_level>                                integer
      [-f] [-d] [-w] [-l] [-h] [-o] [-s] [-r]      options
```

If all end_point's = 0, then the whole spectrum will be picked. This is not advisable if you have a strong residual water signal or noise on the borders of the spectrum which can be easily excluded.

```
Options:    Arguments:                           Meaning:
  -f       <max_delta_fit>                        Integer peak coordinates are
            positive float                        adjusted by at most
                                                  points towards the maximum of the
                                                  fitted 2nd order polynomial.

  -d       <dim1> <dim2> <max_distortion>         Peaks are eliminated whose
            1..n   1..n     0.0 .. 1.0            fit-function curvature-distortion
                                                  (i.e.the ratio between the
                                                  curvature in dim1 and the
                                                  curvature in dim2) is below
                                                  max_distortion.
                                                  (Nice peaks have a distortion
                                                   ratio of 1.0).

  -w       <wiggle_dim_1>..<wiggle_dim_n>         If wiggle_dim_i = 1 then this
              0 or 1 each                          dimension i will be searched for
                                                  peaks which are surrounded by two
                                                  peak-candidates whose spectral
                                                  intensity is of the middle one's
                                                  reverse sign.
                                                  These two peak-candidates will be
                                                  called wiggles and are eliminated.
```

84

```
-l      <linewidth_1>...<linewidth_n>     A peak whose spectral line in
                n integers                dimension i is not over half of
                                          the peak's intensity for at least
                                          linewidth_i points is eliminated.

-h      <min_noise_1>...<min_noise_n>     Noise hyperplanes with at least
                n integers                min_noise_i local maxima are
                                          completely deleted.
                                          (`Hyperplane' = (n-1)-dim linear
                                           space, e.g. in 2D the waterline.)

-o                                        Delete hyperplanes with just one
                                          peak.

-s      + or -                            Pick only positive (+) or negative (-)
                                          peaks. Be careful with folded spectra.

-r      <max_diagonal_delta>              Remove peaks which are at most
        positive float in ppm             max_diagonal_delta ppm off diagonal.
                                          w_n-1 and w_n are proton dimensions.
```

Example:

```
pick    bpti    bpti.peaks    0 50 300    64 300 400    800    -f 0.4
                              -d 1 2 0.3    -w 0 1 0    -l 3 3 4
                              -h  100000 100 300
```

The above example will pick the peaks in a 3D spectrum called bpti.3D.* starting from (0,50,300), ending at (64,300,400), fitting peak-coordinates, excluding peaks which are distorted below 0.3 between w1 and w2, eliminating wiggles in w2, ignoring peaks with small line widths and deleting planes whose weighted average peak count is above 100000 in w1, 100 in w2 and 300 in w3. The resulting peaks are written to the file bpti.peaks.

This routine can be called from XEASY with In-phase peak picking [in].

peakint

The program (N. Schäfer, diploma thesis, ETH) integrates peaks in multidimensional spectra. The program tries to partition the points between the picked peaks. The following rule is used: all points around a peak are added up excluding those points that are:

- peaks them selves,
- under the noiselevel,
- bigger than their neighboring points.

Neighboring points are those that lie between the point and the picked peak.The last rule can be overcome by allowing points to be bigger than their neighbors by a factor which decreases from an initial 1.0 + to 1.0 at the specified distance from the peak. The distance for each dimension is specified in the parameter list.

A volume error is calculated. It is defined as the averaged border integral of the integration region. That is the sum over all border points divided by the number of border points.

The parameters for calling the program (n is the dimension of the spectrum) are:

```
peakint  <specfile>          <peakfile>               file names
         <noise_level>                                integer
Options
-t       <dist_1>...<dist_n>  <max_tolerance>         Tolerate slight intensity
              integer              float > 0.0         increase away from peak
                                                       to account for doublet
                                                       peaks and poorly
                                                       centered peaks.
```

Example:

```
peakint bpti bpti.peaks 800 -t 1 2 2 0.8
```

85

The given example will read in the whole 3D spectrum file bpti.3D.*, integrate all the peaks in the file bpti.peaks with an initial tolerance of 80%. The integrated peaks are written back into the file bpti.peaks.

infit

The program infit (T. Szyperski, P. Güntert, G. Otting, and K. Wüthrich, J. Magn. Reson. 99, 552-560 (1992)) is used for the determination of scalar coupling constants by inverse Fourier Transformation of in-phase multiplets. The input for infit is prepared using the "Add" and "Save" options of the "Slice window [sw]".

pslices

This shellscript can be used to print *.sla files created with the "Save All" command in slice window to the screen or to a postscript device using `gnuplot`. It was tested with gnuplot 3.5.

Usage:

```
sh
pslices filename
```

(Enter *filename* without the ".sla" extension.) The plot should appear on the screen for a few seconds. The following options are printed to stdout:

```
print immediately ------- p
```
With "p enter" the postscript file goes to the default printer (lp)
```
print and clean --------- c
```
As p, but the *filename*.sla file is deleted.
```
produce postscript-file - f
```
Writes *filename*.ps
```
do nothing ------------- n *default*
```
Leave without printing or writing a file.

Looking

---

16. Application Examples

The following chapters present some commonly used applications of the XEASY program in analyzing spectra to be used in the determination of the 3D structure of biomacromolecules. The proposed methods represent only one of many ways of applying the functions of XEASY to a particular task. The aim is to provide a starting point for a novice user by giving some examples. Two letter codes of the commands are used in the examples below.

16.1 Spin System Assignment using [1H,1H] COSY, TOCSY

The general idea is first to identify the frequencies belonging to one spin system in the TOCSY spectrum. These frequencies are then used to display all regions containing the TOCSY peaks of this spin system. To determine the spin system type the TOCSY spectrum is displayed together with the COSY spectrum. Having the whole spin system displayed, the spectroscopist can determine the spin system type, pick the peaks and assign them accordingly.

To start assigning the peaks, generic spin systems are used that can later be changed into any specific amino acid.

`[ns]`,`[as]`
Load the COSY and TOCSY spectrum.
`[ic]`,`[rs]`
Display the COSY as intensity plot and the TOCSY as contour plot.
`[ls]`
Load a fragment list defining the generic spin systems (i.e., ss.seq).
`[pp]`
Pick the COSY peaks in the fingerprint region.
`[ar]`
Assign each of the picked peaks to the Ha and HN atoms of different spin systems.
`[zo]`
Zoom into the TOCSY tower of spin system 201 (i.e., the region containing all the peaks that can be assigned to the amide proton of spin system 201 in the horizontal x direction).
`[df]`
Mark all the frequencies of the spin system.
`[zg]`
Display each peak of the spin system in a separate zoom region.
Decide on the probable spin system type (e.g., GLU).
`[ed]`
Edit the Ha HN peak of spin system 201.
Change the fragment type from the genuine spin system SS into a GLU by editing the fragment field in the peak editing window).
`[cm] or [pp]`
Pick the remaining peaks of spin system 201.
`[ap]`
Assign the peaks.
`[ws]`,`[wc]`,`[wp]`
Save the peaks, atoms and fragments.
Continue with the next spin system.

16.2 Spin System Assignment using 3D HCCH TOCSY

The general idea is first to define strips for each 13C bound 1H in the 3D HCCH-TOCSY using the 2D [13C,1H]-HMQC spectrum. Then all strips are collected that belong to the same spin system. Starting with a strip corresponding to the Ha and Ca atoms of a spin system, one searches for strips at the different proton frequencies of this spin system. Those strips which show a similar cross peak pattern are likely to belong to the same spin system. Strips identified in this way are appended to a new strip list. This new strip list can then be displayed and the peaks can be picked and assigned as in the case of the 2D

87

homonuclear TOCSY described in the previous section.

[ns]
> Load the 2D [13C,1H]-HMQC spectrum. It is advantageous if this spectrum has a large sweep width and does not contain any folded peaks.

[pp]
> Pick the peaks.

[ls]
> Load a fragment list defining the generic spin systems (i.e., ss.seq).

[ar]
> Assign the peaks that could correspond to the Ha and Ca atoms to different spin systems.

[wp]
> Write out the peak list.

[ns]
> Load the 3D HCCH-TOCSY.

[lp]
> Load the saved peak list, use w2(1H) of the loaded peak list for w2(1H) and w3(1H) of the spectrum. This generates a list of peaks lying on the diagonal of the proton planes.

[se]
> Define a strip for each peak.

[gs][fs][bs]
> Display the strips.

[mp]
> Adjust the peak positions in the w2(1H) and w3 (1H) dimensions, search for additional overlapping strips that were missed in the 2D spectrum.

[es][se]
> Redefine the strips at the modified peak positions.

[pm]
> Turn the spectrum around: display planes with w1(13C) horizontally and w2(1H) vertically.

[gs][fs][bs]
> Redisplay the strips.

[mp]
> Adjust the peak position in w1(13C), search for additional overlapping strips that were missed in the 2D spectrum.

[pm]
> Turn the spectrum to view again planes with both proton dimensions.

[es][se]
> Redefine the strips at the modified peak position.

[sf]
> Start with the strip assigned to Ha Ca of spin system 201.

[rd],[pc]
> Look for the strips at one of the proton frequencies of the spin system. These proton frequencies are defined by the cross peaks in the Ha Ca strip.

[fc],[bc]
> Visually inspect the displayed strips to find the one with the same pattern as the Ha Ca strip.

[es],[sa]
> Define a new strip list containing the strips of spin system 201.

[pc],[sa]
> Look for other strips belonging to spin system 201 and add them to the current strip list.

[sh][sr]
> As an alternative the found strips can be kept displayed until the whole spin system is assembled.

[gs]
> Display the strips belonging to spin system 201.

[pp][cm][ed][ap]
> Pick the peaks of the spin system and assign them as described in the section about Spin System Assignment using [1H,1H] COSY, TOCSY on page 109.

[ws]
> Save the strip list of spin system 201.

[es],[se],[sf]
> Continue with the next spin system.

16.3 Sequential Assignment using [1H,1H] NOESY

The general idea is first to build a stretch of sequentially neighboring spin systems by searching for the sequential neighbors of spin systems. This stretch of spin systems is then mapped onto the amino acid

sequence and the fragment numbers are changed from spin system numbers into residue numbers.

To find the sequential neighbor of a spin system, lines are drawn at the positions of the HN, Ha and Hb's frequencies of a given spin system. The sequential neighbor of the given spin system should have sequential peaks at some of the positions marked with the lines (e.g., Ha(i)/HN(i+1), HN(i)/HN(i+1), Hb2(i)/HN(i+1), ...). Possible candidates can be zoomed and compared in detail. Once a stretch of spin systems is found and mapped onto the amino acid sequence the mapping number of the spin system fragments is set to the residue numbers. Then the mapping numbers are exchanged with the fragment numbers. The result are peak lists with peaks assigned to the atoms of the amino acid residues.

First the intra-residual peaks in the NOESY spectrum must be identified:

[ns]
> Load the NOESY spectrum.

[lp],[ls],[lc]
> Load the TOCSY peaks assigned to the spin system fragments (refer to Spin System Assignment using [1H,1H] COSY, TOCSY on page 109).

[mr]
> Adjust the TOCSY cross peaks such that they match the intra-residual NOESY cross peaks.

[wp]
> Save the peak list as your NOESY peak list. Now the search for sequentially neighboring spin systems can start. Here it helps to record stretches of spin systems and the connectivities found in a separate text file or in the comment entry of the fragments (which can be edited using the [ed] command).

[sp]
> Select a spin system to start with by specifying the fragment number.

[ds] or [df]
> Draw lines at the frequencies of the HN, Ha and Hb's.

[za]
> Zoom those regions that could correspond to sequential neighbors.

[sw]
> If necessary make a detailed comparison of the line-shapes of the intra residual peaks and the potential sequential peaks.

[ed]
> Decide on the most probable sequential neighbor. Record your decision in your notes or in the comment entry of the spin system fragment.

Search for the following spin system using the same method.
Try to map your subsequence of spin systems to your amino acid sequence.

[ed]
> Enter the residue numbers of the mapped spin system into the mapping number field in the edit window.

[sn]
> Exchange the fragment numbers with the mapping numbers. Now the fragment numbers correspond to the amino acid residue numbers.

Continue to look for other subsequences of spin systems

[ws][wc][wp]
> Write out the final fragment, atom and peak lists.

16.4 Sequential Assignment using 3D 15N-Correlated [1H,1H]-NOESY and 3D 15N-Correlated [1H,1H]-TOCSY

The general idea is similar to that explained in section 16.3 on page 111. First stretches of sequentially neighboring spin systems are identified by searching for the sequential neighbors of spin systems. The found stretches of spin systems are mapped onto the amino acid sequence and the fragment numbers are changed from spin system numbers into residue numbers.

In a 3D 15N-Correlated [1H,1H]-NOESY strips can be defined at the position of the HN and N atom frequencies of the amide groups of the different spin systems. These strips can be used to identify sequentially neighboring spin systems. A strip corresponding to a sequential neighbor of another strip should have a) sequential peaks and b) have a similar cross peak pattern as the neighboring spin system.

Both features can be exploited by the spectral correlation method using the peak pattern of the strip from the NOESY and the picked peaks from the TOCSY spectrum (C. Bartels and K. Wüthrich, to be published).

Once a stretch of spin systems is found and mapped onto the amino acid sequence the mapping number of the spin system fragments is set to the residue numbers. Then the mapping numbers are exchanged with the fragment numbers. First the strips must be identified and the 15N-Correlated [1H,1H]-NOESY must be picked:

[ns]
Load the 2D [15N,1H]-HMQC spectrum.

[pp]
Pick the peaks.

[ls]
Load a fragment list defining generic spin systems (i.e., ss.seq).

[ar]
Assign the peaks to the HN and N atoms of different spin systems.

[wp]
Write out the peak list.

[ns]
Load the 3D 15N-Correlated [1H,1H]-TOCSY.

[lp]
Load the saved peak list, use w2(1H) of the loaded peak list for w2(1H) and w3(1H) of the spectrum. This results in a list of peaks which lie on the diagonal of the proton planes.

[se]
Define a strip for each peak.

[gs][fs][bs]
Display the strips.

[mp]
Adjust the peak positions in the w2(1H) and w3 (1H) dimensions, search for additional overlapping strips that were missed in the 2D spectrum.

[es][se]
Redefine the strips at the modified peak positions.

[pm]
Turn the spectrum around: w1(15N) horizontally and w2(1H) vertically.

[gs][fs][bs]
Redisplay the strips.

[mp]
Adjust the peak position in w1(15N), search for additional overlapping strips that were missed in the 2D spectrum.

[pm]
Turn the spectrum to view again both proton dimensions.

[es][se]
Redefine the strips at the modified peak position.

[gs][fs][bs]
Redisplay the strips.

[cm]
Pick the peaks and if possible assign them to the side chain atoms of the spin system. Keep notes about the probable type of the spin system. You can use the comment entry of the fragment which can be set in the peak edit window.

[wp],[ws],[wc]
Save the TOCSY peak list. The TOCSY peak list produced can now be loaded onto the NOESY spectrum. Since the peak positions in the NOESY and the TOCSY spectrum may differ slightly, the peaks may have to be re-positioned in the NOESY spectrum. The TOCSY peaks define the position of the intra-residual peaks in the NOESY spectrum and can be used to identify the sequential peaks in strips that could correspond to the sequentially neighboring spin systems.

[ns]
Load the 3D 15N-Correlated [1H,1H]-NOESY spectrum.

[lp]
Load the peak list of the 3D 15N-Correlated [1H,1H]-TOCSY.

[pm]
Display w1(15N) horizontally and w2(1H) vertically.

[mr]
Position the peaks in the NOESY as already described for the TOCSY.

[sp]

90

Select the diagonal peaks. For example by setting the field Assignments (e.g. HA HN) to N HN HN.

`[se][gs][fs]`
Define and display the strips.

`[da]`
Redisplay all peaks. This is necessary for the calculation of the spectral correlation coefficients since only the displayed peaks are considered.

`[cp]`
Define the noise level by drawing contours. All peaks and as little as possible noise should be visible.

`[cd]`
Calculate the spectral correlation between the strips in the current strip list.

`[sf]`
Start with a strip that corresponds to a spin system of known type (i.e., Gly or Ala).

`[cc]`
Display potential sequential neighbors of the strip.

`[fc],[bc]`
Visually inspect the displayed strips to find the one that most likely corresponds to the sequential neighbor.

`[sw]`
If necessary, make a detailed comparison of the line-shapes of the intra residual peaks and the potential sequential peaks.

Decide on the most probable sequential neighbor. Record your decision in your notes or in the comment entry in the fragment list.

`[es],[sa]`
Define a new strip representing the sequence of sequentially neighboring spin systems.

`[cc][sa]`
Look for the strips that corresponds to the following spin system.

`[gs]`
Display the strip list with the sequence of sequentially neighboring spin systems.

Try to map your subsequence of spin systems to your amino acid sequence. Keep notes about the used spin systems and their connectivities.

`[ed]`
Enter the residue numbers of the mapped spin system into the mapping number field in the edit window.

`[sn]`
Exchange the fragment numbers with the mapping numbers. Now the fragment numbers correspond to the amino acid residue numbers.

`[ss]`
Save the strip list containing the assigned subsequence.

`[ws][wc][wp]`
Write out the changed fragment, atom and peak lists.

`[es],[se],[sf]`
Continue to look for other subsequences of spin systems.

16.5 Sequential Assignment using HNCA and HN(CO)CA

The general idea is discussed in section 16.3 on page 111. First one builds a stretch of sequentially neighboring spin systems by searching for sequential neighbors of spin systems. This stretch of spin systems is then mapped onto the amino acid sequence and the fragment numbers are changed from spin system numbers into residue numbers.

In 3D HNCA and HN(CO)CA experiments, strips can be defined at the position of the HN and N atom frequencies of the amide groups of the different spin systems. These strips can be used to identify sequentially neighboring spin systems. First, for each spin system, the frequency of the Ca of the same and the previous spin system must be identified. In small molecules this is possible using only the HNCA experiment. For larger molecules, the HN(CO)CA experiment might help.

Strips corresponding to sequential neighboring spin system should have one Ca peak in common. Starting from one Ca peak of a strip one can search for all the strips that have a peak at the same position. The lineshapes of these peaks must then be compared carefully.

91

Once a stretch of spin systems is found and mapped onto the amino acid sequence the mapping number of the spin system fragments is set to the residue numbers. The mapping numbers are then exchanged with the fragment numbers.

First strips must be identified. This is the same procedure as when doing the sequential assignments using the 15N-Correlated [1H,1H]-TOCSY:

[ns]
>   Load the 2D [15N,1H]-HMQC spectrum with a large sweep width.
[pp]
>   Pick the peaks.
[ls]
>   Load a fragment list defining the generic spin system fragments (i.e., ss.seq).
[ar]
>   Assign the peaks to the HN and N atoms of different spin systems.
[wp]
>   Write out the peak list.
[ns],[as]
>   Load the 3D HNCA and HN(CO)CA spectra.
[ic],[rc]
>   Display the HNCA spectrum as intensity plot and the HN(CO)CA spectrum as contour plot.
[lp]
>   Load the saved peak list, use w2(1H) of the loaded peak list for w3(1H) of the spectrum and w1(15N) for w1(15N). Set w2(13C) to a constant value corresponding to the approximate chemical shift of the Ca atoms. This generates a list of peaks lying on one plane in w2(13C).
[se]
>   Define a strip for each peak.
[gs][fs][bs]
>   Display the strips.
[mp],[ap]
>   Move the picked peaks to the observed peak positions and assign them to the Ca of the corresponding spin system.
[cm]
>   Pick the second peak, corresponding to the Ca of the previous spin system.
[es][se][rd]
>   Define a reference strip list with strips for each of the peaks in the spectrum.
[sp]
>   Select the intra-residual peaks. For example by setting the field Assignments (e.g. HA HN) to N CA HN and the field Sequential range (intra and seq. = 0 1): to 0.
[es][se]
>   Define a strip sequence containing one strip per spin system.
[sf]
>   Start with a strip that corresponds to a spin system of known type (i.e., Gly or Ala).
[pc]
>   Display potential sequential neighbors of the strip (i.e., strips that have a cross peak at the same position as the displayed strip).
[fc],[bc]
>   Visually inspect the displayed strips to find the one that most likely corresponds to the sequential neighbor.
[sw]
>   If necessary make a detailed comparison of the line-shapes of the intra residual peak and the potential sequential peak.
[ed]
>   Decide on the most probable sequential neighbor. Record your decision in your notes or in the comment field of the fragment list.
[es],[sa]
>   Define a new strip list containing the sequence of sequentially neighboring spin systems.
[pc][sa]
>   Look for strips corresponding to the following spin systems.
[gs]
>   Display the sequence of sequentially neighboring spin systems.

Try to map your subsequence of spin systems to your amino acid sequence. Keep notes on the used spin systems and the observed connectivities.

[ed]
>   Enter the residue numbers of the mapped spin system into the mapping number field in the edit

92

window.

[sn]

Exchange the fragment numbers with the mapping numbers. Now the fragment numbers correspond to the amino acid residue numbers.

[ws]

Save the strip list of this subsequence.

[es],[se],[sf]

Continue to look for other subsequences of spin systems.

[ws][wc][wp]

Write out the final fragment, atom and peak lists.

16.6 NOESY cross peak assignment in 2D

For collecting a set of distance constraints necessary for the determination of the 3D structure of a biomacromolecule, NOESY cross peak assignments are crucial. Other necessary steps include NOESY cross peak integration, conversion of the peak intensities to distance constraints, determination of additional constraints (e.g., coupling constants and amide exchange rates) and checks to ensure the consistency of the entire set of constraints. The following two sections treat the procedure of NOESY cross peak assignment.

Based only on the known frequencies a number of assignment possibilities can be proposed for each peak. These possibilities can be used by the spectroscopist to judge the most likely assignment. The assignments are usually checked in the spectrum by comparing the lineshapes and by ensuring that they are part of a consistent set of constraints. The following commands represent a likely sequence used to assign cross peaks in a NOESY spectrum.

[ns]

Load the NOESY spectrum.

[ls],[lc],[lp]

Load the NOESY peak list originating from the sequential assignment.

[zo],[cp]

Zoom into a region of the spectrum.

[pp]

Pick the remaining peaks.

[gr][gl][gu][gd]

Go to an adjacent region of the spectrum and continue the peak picking.

[pp]

Continue the peak picking.

[ac]

Calculate the average chemical shift from the assigned peaks in the NOESY spectrum.

[pm]

Redisplay the whole spectrum.

[pa]

Generate possible assignments for the peaks in the spectrum.

[zo]

Zoom into a region of the spectrum, for example the amide region.

[ea]

Look at the possible assignments of a cross peak and select the most likely one.

[vr]

Compare the lineshapes of all the peaks with the same assignment as the selected peak.

[ea]

If the line-shapes are inconsistent select an other assignment.

Continue with another peak.

An alternative method to search for valid assignments is given by the routine to search for strips close to a selected position. After loading the spectrum and peak list, the following commands can be used to find valid assignments for a cross peak:

[sp]

Select all assigned peaks by setting the field Assignments (e.g. HA HN) to * *.

[es][se][rd]

For each peak define a reference strip.

[zo]

93

[pc]

Zoom into a region around the peak to be assigned.

Display strips close to the peak by clicking the left mouse button at the position of the peak. The peak to be assigned is displayed together with all peaks that are already assigned and have a similar chemical shift.

16.7 NOESY cross peak assignment in 3D

This section describes NOESY cross peak assignments using heteronuclear correlated 3D spectra.

In 2D spectra, checks of cross peak assignments have mainly relied on the comparison of peak line shapes and position. In heteronuclear correlated 3D spectra, it is possible to check assignments by looking for "reverse" peaks: if a peak involving dipolar magnetization transfer from atom A to atom B exists its reverse peak involving magnetization transfer from atom B to atom A must also exist. This fact is used by the following method to assign cross peaks in heteronuclear correlated 3D NOESY.

First for each atom strips must be defined. For example, in the 15N-correlated [1H,1H]-NOESY, strips for each amide proton can be defined. Similar in the 13C-correlated [1H,1H]-NOESY spectrum strips can be defined for each proton bound to a carbon. The defined strips can then be used to search for valid assignments of cross peaks. Strips of atoms possibly coupled to the atom of the selected strip can be displayed by searching for strips with a diagonal peak at the position of one of the cross peaks. The displayed strips can then be screened for the one containing the reverse peak.

[ns]

Load the 15N-correlated [1H,1H]-NOESY.

[ls],[lc],[lp]

Load a peak list of the 15N-correlated [1H,1H]-NOESY with assigned diagonal peaks.

[sp]

Select all the diagonal peaks, for example by setting the field Diagonal peaks (tolerance in ppm, e.g 0.02) to 0.02.

[es][se]

Define for each of the diagonal peaks in the 15N-correlated [1H,1H]-NOESY one strip.

[ns]

Load the 13C-correlated [1H,1H]-NOESY.

[lp]

Load a peak list of the 13C-correlated [1H,1H]-NOESY with assigned diagonal peaks.

[sp]

Select all the diagonal peaks, for example by setting the field Diagonal peaks (tolerance in ppm, e.g 0.02) to 0.02.

[se]

Define for each of the diagonal peaks in the 13C-correlated [1H,1H]-NOESY one strip.

[rd]

Copy the current strip list to the reference strip list.

[sf],[zo]

Display the strip containing the peak to be assigned.

[pc]

Click the left mouse button at the position of cross the peak. Strips are displayed from either of the two spectra with a diagonal peak at the position of the selected cross peak.

[fc][bc]

Search for the strip that contains the reverse cross peak.

[sw]

If necessary, make a comparison of the line shapes.

[ap]

Assign the cross peak.

Continue the assignments with other cross peaks.

16.8 Measuring Rate Constants

94

For measuring rate constants (e.g., relaxation rates or amide exchange rates), one usually records a series of 2D spectra under identical conditions at different time points. The peak intensities vary in such spectra according to the rate constant of the measured process. To analyze these spectra, the same regions must be integrated in the different spectra. The series of 2D spectra might either be present as a set of individual 2D spectra or as a single 3D spectrum.

`[ns]`
> Load the spectrum or the plane with the largest peak intensities.

`[lp]`
> Load the peak list of the spectrum.

`[dr]`
> Draw rectangles around each peak. The rectangles define the size of the region to integrate. In order to copy the integral to the picked peaks, only a single peak should lie within one rectangle.

`[wg]`
> Save the rectangles in a file

`[ir]`
> Integrate the rectangles.

`[wp]`
> Write out the peak list containing the peak volume of the first spectrum.

`[ns] or [fp]`
> Load the next spectrum or go to the next plane.

`[ir], [wp]`
> Integrate the peaks and write them out.

Continue with the remaining planes.

In this way, a peak list containing the volumes of the peaks integrated is generated for each 2D spectrum or each plane of the 3D spectrum. The volumes present in these peak lists can then be used to calculate the desired rate constants.

next: Installation / contents

Looking

95

Installation from source code

For a complete installation of the program package the following steps are necessary. First all the programs must be compiled. This is done by typing

`./configure` *system release*

where *system* stands either for sgi - if you have a silicon graphics work station, sun - if you have a sun sparc station, or solaris - if you have SunOS 5.x. The program should run also on other systems (I was told that somebody has it running under LINUX), however we have only tested the program on SGI and Sun work stations. To compile and run the XEASY program Motif libraries must be properly installed on your system.

release can be either "120" for release 1.2.0 (Sept. 95) or "new" for the most recent release, which was 1.3.11 in December 96.

Meanwhile the 1.3.x release is much more stable than 1.2.0, so I recommend everybody who installs XEASY to use the newest release.

Stability problems have been observed with release 3 of the motif libraries on Sun computers. I could not identify the problem - link Xm release 2 or 1 if you have have the choice.

Installation of executables

Executables for SunOS 4, SunOS 5 (=Solaris 2) and IRIX are provided. If you install one of these XEASY executables you still have to compile the external programs that you may need.

XEASY 1.3.x

The newest release is in /src.new. When you "configure *system* new" src.new is simply linked as src. Depending on *system* some default paths are inserted in the Makefile. If your system is neither SGI nor solaris, or if the default doesnt compile you should have a look at Makefile and xe_sys.h.

read XEasy! Modify the file according to your needs and put it somewhere, where X-Window will find it - for example to your home directory. You don't need any environment variables to run release 1.3.x

XEASY 1.2

In your .cshrc file two environmental variables must be defined and the path can be changed to include the ./bin directory that contains all the executables of the XEASY program package. The two environmental variables are XEASY_GROUP and XEASY_LIB. XEASY_GROUP specifies the groups home directory, i.e., the directory in which the home directories of the different users are located. So, if your home directory is /home/mydir, you must add to your .cshrc file

`setenv XEASY_GROUP /home/`

`XEASY_LIB` specifies the default library, the provided library is `./src/ecepp.lib`. So, if for example you have installed the XEASY program package in the directory /home/programs, you should add the line `setenv XEASY_LIB /home/program/src/ecepp.lib` to your .cshrc file.

The files with the extension ".c" contain the basic source code and files with the extension ".h" are the included files defining the global constants and variables for the program XEASY. The file "Makefile" is used for compiling and linking XEASY. To produce the executable file "xeasy", type the command

96

"make". On SUN the option "-lPW" should be removed for the library flag "LIBS" in the file "Makefile". Since XEASY links in the libraries of the Motif widget set., these libraries must be installed on the used computer system in order to compile XEASY. However the compiled XEASY program can be run without these libraries. Compiled version for SUN sparc station and for SG Indigo computers are provided, they are called xeasy.sun and xeasy.sgi respectively.

Each user must create a subdirectory called "easy" in their home directory, where all files except the spectrum files will be stored. In the subdirectory "easy" there should be another subdirectory called "data" where the spectrum files are located. It is often convenient to make the "data" directory a link to a directory in a partition that is capable of holding large files.

In the user's ".cshrc" file there should be two environment variables defined, one is "XEASY_GROUP" for the group directory and the other is "XEASY_LIB" for XEASY fragment library. For example, if a user wants to read a sequence file called mysequence.seq, XEASY will try to construct the file name as follows: "XEASY_GROUP/userloginname/easy/sequence.seq". The "XEASY_LIB" should be the complete file name of the fragment library, for example, `"/home/project/ecepp.lib"`.

External programs

If the installation runs well, the programs in external/*/ will be compiled, and the executables copied to bin/. If, for some reason, this does not work: go to the subdirectories, check the Makefiles, and type make, or make *program*, respectively.

Bruker X32 to XEASY

If you want to use the cpx32 shell script, that allows you to directly fetch spectra files form a Bruker X32 computer and transform them into the file format needed by XEASY, you need to install the getpar program on the X32 computer and adapt the getx32 and getx32rrr shell scripts in your ./bin directory. Copy the source code ./external/conversion/getpar.c to the X32, say into the directory /v/local/bin. Compile it with

`cc -o getpar getpar.c` Change the paths in the getx32 and getx32rrr shell scripts from /u/local/bin into /v/local/bin, i.e. change the lines

`rsh $1 -n /u/local/bin/getpar /$2/data/$3/nmr/$4/1/pdata/1`

into

`rsh $1 -n /v/local/bin/getpar /$2/data/$3/nmr/$4/1/pdata/1`

[contents](#)

Looking

97

[contents](#)

Summary of changes since Release 1.2 (Sept. 95)

In September 1995 [I](#) have taken over responsibility for XEASY from Christian Bartels. I refer to the version that I got in September as Release 1.2 (somewhat arbitrarily, a documentation of the changes made before is not available.) It is **not** my job to make major improvements to XEASY, but I will

- adapt xeasy to exchange information with other programs
- make some small modifications that become neccessary for the group (and in some cases also for external users, if the problem is a general one).
- integrate additions and improvements made by the users, and make them available to other users.

Release 1.2.4 (13.10.95)

- "Save all" button in [slice window](#) (to print content of sw, see [pslices](#))
- nr. of occupied colors reduced to 105 (the other 15 have not been used anyway)
- colors changed, meaning of sliders changed
- bugfixes made by Christian Bartels to drawspectra2, 3, menubuttons, readspectra.
- automatic inclusion of release-nr.

Release 1.2.11 (10.11.95)

- alternative colormaps: It is now possible to run any number of XEASY instances on one screen. Interference between colormaps can be minimized by setting the appropriate [user resources](#)
- xe_sys.h added for system-specific compilation.
- beep can be swiched off in the [resource file](#). This works on all machines
- "EPSF-1.2" added in postscript file header
- ignores comments (#...) info in peaklist files that cannot be related to a peak (earlier releases crashed)
- limited support for dimension identifiers ("iname")

Release 1.3

=1.2.11 has been made to be able to "go back" to this release

Release 1.3.4 (25.11.95)

- Default paths are taken from [user resources](#), not from $XEASY_GROUP...
- started to add support for [macros](#)
- xeasy.default has been made much smaller. (The format has changed twice, in 1.2.11 and 1.3.3.)
- xeasy.default and xeasy.txt are written in the startup directory, not in ~/easy/. This makes it much easier to have different defaults: Start your copies from different directories.

Release 1.3.5 (12.12.95)

- indirect dimension in [plotfiles](#)

Release 1.3.6 (31.1.96)

- bug fixed: dimension ID's when loading and writing out peaklists (there is still a bug !)
- timeout between keypressed events. default 2000 ms. adjust with XEasy*timeout
- a new command: "wd" to write an xeasy.default file immediately

98

font can now be defined in the resources

- pw has a new option: Assignment y
- some more macro support: "lp filename w1 w2 w3" and "gs - first number width" seem to work properly.
- A new numbering scheme for the proton list can be selected with new_atom_numbers: true. "false" gives compatibility with the previous scheme. "true" gives compatibility with SPSCAN. But you may run into problems with the long proton numbers you get when you read the lists into FORTRAN programs, e.g. ASNO!

Have a look at [XEasy](#)

xeasy.lib

/opt/group/lib/xeasy.lib is a library based on the most recent library used in diana. It has, in addition to the normal amino acids, SS and BB the residues SRD and SSP.
SRD defines the atoms CA+-HA, CA+-HA2, CB+-HB2, CB+-HB3 in a reproducable manner. You have to use this residue type to correct automatically evaluated reduced dimensionality spectra in xeasy.
SSP defines the additional atoms Cp, CAp, HAp, HA2p, CBp, HB2p, HB3p and QBp. These are the respective atoms of HN(CO)... spectra. To assign these atoms correctly is a prerequisite for semi-automated assignment and consistency checks.

xeasy.lib has no shift information, i.e. the "rw" commands cannot (yet) be used with this library.

To use the library, set
XEasy.residue_library: /opt/group/lib/xeasy.lib

Release 1.3.7 (27.2.96)

- some debugging of 4D commands. Should be possible to use XEASY on 4D spectra now, if it is compiled with "#define DIMENSION 4"
- another bug fixed in adapting dimension ID's

Release 1.3.8 (13.5.96)

- Changes in xeasy.lib, residues SSP and SRD. Some atom numbers may be different when you create new proton lists from *.seq (Only for residues SSP and SRD)
- "gp" folding debugged
- "ap" proposals
- 2 instabilities fixed
- macro commands debugged

Release 1.3.9 (3.6.96)

- bug in [li] fixed

Release 1.3.10 (22.7.96)

- small bug in the library reading routine fixed. It was undetected, because usually all residues have "O" as last atom - this isn't read anyway. Only in some pseudoresidues the last atom had not be read.
- bug fixed which made xeasy sometimes crash when slices in the indirect plane were extracted.
- There is a new command "[nc](#)" which does the same as "ac", but does not set back shift to 999.0 and error to 0.0 for atoms that are not in the peaklist or peaklists.
- **xeasy.lib** has some new pseudo-residues for RNA work. Some old-style pseudoatom numbers were corrected.

99

pipe2xeasy NMR Pipe to xeasy converter (test release 8.10.96)

Release 1.3.11 (18.10.96)

- plotfiles have now only right and bottom scales, and no separation between margin and axes. Some defaults are changed. Nothing basic.

pipe2xeasy NMR Pipe to xeasy converter (bugfixes 12.11.96)

Release 1.3.12 (group)

- some more bugs fixed in uxnmr_to_xeasy.
- commands "is", "ps" and "sp" to integrate spscan commands into xeasy.
- inames of peaklists are always adapted to the inames of the spectrum when the peaklist is written out. (in earlier versions they were only adapted if they had not been set before.)
- DIMENSION is set to 4 (you cannot load xeasy.default files created with 3D executables)
- the old (Release 2) Motif library is statically linked for reasons of stability.

Release 1.3.13 (group)

- command "wm" to write a 4D peaklist from a 3D reduced-dimensionality experiment.
- Meaning of "Scaling factor" in "co", "di" commands changed (by a factor of 2): Scaling factor is now 1/kappa * sf_c/sf_s, where sf_c is the frequency for the nucleus that defines the center of the doublet and sf_s is the frequency for the nucleus that defines the splitting.

---

contents / bugs

Last modified: Feb 17th, 1997
Ralf W. Glaser, xeasy@mol.biol.ethz.ch

Looking

100

References

In all publications for which data were processed with XEASY, please refer to

Ch. Bartels, T.-H. Xia, M. Billeter, P. Güntert and K. Wüthrich (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. J. Biomolecular NMR 5, 1-10

XEASY has been licensed to about 90 groups worldwide (August 1996) and in many of these groups it is used as the only, or as the most frequently used program to evaluate NMR spectra of biological macromolecules.

Availability

Academic licences for XEASY are available for a handling fee of US$ 200.- per group, independent of the number of computers. You get the source code, and you can also get executables for Sun and SGI. However, on some machines the executables are incompatible with local libraries, and you have to compile the program on your machine. Academic licences can **not** be used for projects that are supported by a "for profit" organization. To apply for an academic licence send e-mail including your full mail adress (or the adress of the head of the group) to [xeasy@mol.biol.ethz.ch](). A licence form will be sent to you. After receipt of the signed licence form and payment you get access to crypted files on our ftp-server.

For other licences please contact [Prof. Kurt Wüthrich]().

System requirements

XEASY is written in C to run on X-Window systems. It uses the motif libraries, i.e. **motif must be installed**. For SunOS 4.x and 5.x executables with statically linked motif library are available, so that you may be able to run it without motif. XEASY is extensively tested on various Sun workstations and on SGI. No basic problems are known that would prevent it from running on other UNIX systems supporting X-Window.

Upgrades

Those who have a valid licence for any release of XEASY can upgrade to Release 1.2 (September 1995), or to the most recent release, free of charge. 1.2 is the release described in

Ch. Bartels, T.-H. Xia, M. Billeter, P. Güntert and K. Wüthrich (1995) The program XEASY for computer-supported NMR spectral analysis of biological macromolecules. J. Biomolecular NMR 5, 1-10

To find out, whether it is beneficial for you to install a new release, check the [Recent Changes section on the server of the Wüthrich Group]().

HTML Manual

The HTML manual describes the most recent release. External users who access the manual frequently are kindly requested to install the html-manual on their own server. The current version of the manual is available by ftp from "ftp.mol.biol.ethz.ch" under the directory "pub/software/" as "xeasy.html-manual.gz" (gzip'd, but not crypted. Don't forget to set binary mode for transfer.)

The "postscript manual" describes Release 1.2.0 and will not be updated any longer.

[contents]()

Ralf W. Glaser, xeasy@mol.biol.ethz.ch

Looking

102

Known Bugs

back to [XEASY home page](#)

- When you assign a peak with "ap" to an atom name, which is not in the proton list (i.e. the library does not have this atom name for the current residue, and it has not been used before), then you heve to call the "ap" command twice: in the first run the new atom is added to the proton list, and in the second run the peak is really assigned to the new atom. (However, you should avoid using atoms that are not in the library, anyway.)
- sm + cc doesnt display anything. Use ri or ms to see strips and symmetrize them individually with zn
- If a *.sla file is written from a spectrum that is not completely loaded, i.e. if in "pm" the range starts from something other than 0, referencing of the *.sla file is wrong.
- If you have two different spectra as intensity and contour spectrum, and you make a plot, the result may not always be what you expect. FIX: make "rs" before plotting, so that you have the same spectrum as intensity and contour.
- don't overwrite or delete a spectrum file with another program, or change a peaklist with a text editor while xeasy uses the list.
- "cd" - "cc" works only with 'Use absolute value of intensities = y'.

Fixed bugs:

- font usage was a bit chaotic before release 1.3.6.
- The 4D commands had some bugs before release 1.3.7. (They are still not very well tested - I'm interested in reports from people who work a lot on 4D spectra with xeasy)
  The infinite loop of tree is not a bug: it appears when you try to load a 4D spectrum into an executable that was compiled with "#define DIMENSIONS 3".
- If you load a proton list (lc) and a peaklist (lp) and then try to edit a peak (ed) xeasy will crash. Workaround: load the sequence (ls) from which the proton list originates before loading the proton list. Bug fixed in 1.3.8.
- If a cluster of peaks is integrated with "li" (Denk lineshape integration), and for one of the peaks in this cluster the integration method was 'r', 'a', 'e' or 'm', then all following peaks **within this cluster** have the wrong volume (the volume of the previous peak). Fixed in 1.3.9. (Bug report + bug fix courtesy of Dieter H.W. Kastrau)
- A number of stability problems has been solved.

Dr. Ralf W. Glaser
Institut für Molekularbiologie & Biophysik
ETH Hönggerberg
CH-8093 Zürich
Switzerland
E-mail: xeasy@mol.biol.ethz.ch
Fax: 0041-1-633 1073
Tel.: 0041-1-633 2469

Last modified: Jan 25, 1996

Looking