

## AURELIA, a program for computer-aided analysis of multidimensional NMR spectra

Klaus-Peter Neidig<sup>a</sup>, Matthias Geyer<sup>b</sup>, Adrian Görler<sup>b</sup>, Christof Antz<sup>b</sup>, Rainer Saffrich<sup>b</sup>, Wolfgang Beneicke<sup>b</sup> and Hans Robert Kalbitzer<sup>b,\*</sup>

<sup>a</sup>*Bruker Analytische Messtechnik GmbH, Silberstreifen, D-76287 Rheinstetten, Germany*

<sup>b</sup>*Max-Planck-Institut für medizinische Forschung, Abteilung Biophysik, Jahnstrasse 29, D-69028 Heidelberg, Germany*

Received 29 May 1995

Accepted 25 July 1995

*Keywords:* Multidimensional NMR; Pattern recognition

---

### Summary

AURELIA is an advanced program for the computer-aided evaluation of two-, three- and four-dimensional NMR spectra of any type of molecule. It can be used for the analysis of spectra of small molecules as well as for evaluation of complicated spectra of biological macromolecules such as proteins. AURELIA is highly interactive and offers a large number of tools, such as artefact reduction, cluster and multiplet analysis, spin system searches, resonance assignments, automated calculation of volumes in multidimensional spectra, calculation of distances with different approaches, including the full relaxation matrix approach, Bayesian analysis of peak features, correlation of molecular structures with NMR data, comparison of spectra via spectral algebra and pattern match techniques, automated sequential assignments on the basis of triple resonance spectra, and automatic strip calculation. In contrast to most other programs, many tasks are performed automatically.

---

### Introduction

The extensive use of advanced computer programs for the analysis of NMR spectra becomes more and more important, since data evaluation, not data recording, often represents the time-consuming bottleneck before a scientific problem can be solved by NMR methods. This is especially apparent when multidimensional NMR spectroscopy is applied to biological macromolecules. Correspondingly, a number of groups have developed and published various more-or-less elaborate software packages (for a review, see Kalbitzer, 1994). However, many of these programs deal only with special aspects of data evaluation. The present paper describes the fundamental features of the general-purpose program package AURELIA, which provides all tools necessary for the evaluation of multidimensional frequency-domain data, and its application to a medium-sized protein, the phosphocarrier protein HPr, as well as a small neuropeptide, the head activator.

The analysis of multidimensional NMR data can be viewed within the scope of a general image analysis,

which usually comprises four different stages of operation: (1) improvement of the image quality with enhancement of important features; (2) detection and separation of relevant objects from the background; (3) classification of objects; and (4) interpretation of objects and classes of objects as part of a local or global analysis. This final step also includes the quantification of objects.

For the improvement of the spectral quality, a number of frequency-domain methods have been proposed. Probably the most important are methods for removing base plane distortions (Barsukov and Arseniev, 1987; Zolnai et al., 1989; Dietrich et al., 1991; Güntert and Wüthrich, 1992; Saffrich et al., 1993; Hare and Prestegard, 1994; Rouh et al., 1994). For spectra recorded with older instruments, procedures for the suppression of  $t_1$  ridges,  $t_1$  oscillations and antidiagonals are useful (Klevit, 1985; Glaser and Kalbitzer, 1986; Zolnai et al., 1986, 1988; Mitschang et al., 1990; Kalbitzer et al., 1991b; Manoleras and Norton, 1992). Symmetry enhancement methods can be used in spectra with inherent symmetries, in order to intensify these features relative to noise and artefact

---

\*To whom correspondence should be addressed.

signals which do not show symmetry (Baumann et al., 1981a,b; Bolton, 1985,1986a,b; Neidig and Kalbitzer, 1988,1991; Kalbitzer et al., 1991b).

The detection and separation of the relevant objects in the spectra, the multidimensional NMR peaks, is usually done by peak picking, which may be combined with methods exploiting peak shape information for a discrimination against noise and artefact peaks (Neidig et al., 1984,1990; Glaser and Kalbitzer, 1987; Cieslar et al., 1988; Novic and Bodenhausen, 1988; Pfändler and Bodenhausen, 1988a; Stoven et al., 1989; Kleywegt et al., 1990; Eccles et al., 1991; Garrett et al., 1991; Chylla and Markley, 1993; Rouh et al., 1994; Antz et al., 1995).

If signals are recognized as true resonance peaks, they are treated as valid objects described by a set of parameters such as chemical shifts, line shapes, multiplet structures, volumes, and statistical quality factors (classification and local interpretation). For the recognition of multiplets, a number of procedures have been proposed in the past. The multiplet structures searched for must be defined by the user or may be extracted by the program itself from training sets (Glaser and Kalbitzer, 1987; Hoch et al., 1987; Meier et al., 1987; Meier and Ernst, 1988; Novic et al., 1988; Novic and Bodenhausen, 1988; Neidig et al., 1990; Shen et al., 1990; Eccles et al., 1991; Kjaer and Poulsen, 1991; Shen and Poulsen, 1992). Integration of the cross peaks is essential for the calculation of interatomic distances and is sometimes used for the calculation of dihedral angles. Because of the importance of these parameters, various methods for volume determination have been published ranging from simple interactive to fully automated approaches (Denk et al., 1986; Holak et al., 1987; Olejniczak et al., 1989; Eccles et al., 1991; Güntert et al., 1993; Brown and Huestis, 1994; Geyer et al., 1995).

The global interpretation of the data ideally leads to a complete assignment of all resonances and the extraction of all relevant parameters that are necessary for the structural calculation. In the homonuclear case, these assignment procedures can be classified into two extremes. The first is the standard strategy first introduced by Wagner and Wüthrich (1982), where first as many general spin system patterns of elementary building blocks of the molecule as possible are assigned before the sequential NOEs are used for the sequence-specific assignments. The alternative is the main-chain-directed (MCD) strategy, which tries to assign the resonances of the backbone before assigning the spin patterns of the side chains (Englander and Wand, 1987; Wand and Nelson, 1991). In practice, hybrid techniques, which combine the advantages of the two methods, are often used (Cieslar et al., 1988; Eggenberger et al., 1988; Grahn et al., 1988; Eggenberger and Bodenhausen, 1989; Kleywegt et al., 1989,1991; Weber and Mueller, 1989; Van de Ven, 1990; Van de Ven et al., 1990; Eccles et al., 1991). With the

advent of multidimensional NMR in conjunction with isotope labeling, the number of possible NMR experiments and hence possible assignment strategies is exploding (Cieslar et al., 1990; Vuister et al., 1990; Oschkinat et al., 1991; Bernstein et al., 1993; Kraulis, 1994; Meadows et al., 1994; Oschkinat and Croft, 1994; Zimmerman et al., 1994). Since the development is not yet finished, a generally applicable computer program such as AURELIA cannot incorporate all possible strategies, but has to be flexible enough to be adapted to new developments.

Although a large number of publications exist where specific algorithms for data evaluation are presented, only a few programs have been described in the literature that are designed for a general and automated evaluation of multidimensional data sets (Kraulis, 1989; Kleywegt et al., 1990,1993; Zolnai et al., 1990; Eccles et al., 1991; Kjaer et al., 1991,1994; Markley et al., 1991; Darba, 1992). In this paper we will describe the program AURELIA, its general architecture and organization, as well as specific aspects of its application.

## Materials and Methods

### *Biological samples and NMR spectroscopy*

HPr proteins from *E. faecalis* and *S. aureus* were purified as described earlier (Kalbitzer et al., 1982). Recombinant HPr from *S. carnosus* was obtained according to Kruse et al. (1993). Assignments for HPr from *S. aureus* were taken from Kalbitzer et al. (1991a), and the three-dimensional structure was calculated as published earlier (Kalbitzer and Hengstenberg, 1993). The neuropeptide head activator was synthesized by an Fmoc solid-phase procedure; the resonance assignments were taken from Saffrich et al. (1989).

### *Software*

AURELIA (AUtomated REsonance Line Assignment) is written in C; the graphics is based on X-Windows (X11, R5). According to personal preferences, the user may run it with a pure X11 or with an X11/Motif user interface. The program runs on all types of SGI computers and Bruker ASPECT stations. A PC version, running under Windows NT, is available as well. The program is distributed by Bruker and is available on the ftp servers ftp.bruker.de and ftp.bruker.com.

## Results and Discussion

### *General organization of the program*

Following the ideas of a general image analysis, AURELIA is organized in different modules (Table 1). Module 0 is used for the general definition of input data, module 1 for display, plotting and artefact reduction, modules 2 and 3 for the local and global analysis, respect-

TABLE 1  
GENERAL ORGANIZATION OF THE PROGRAM AURELIA

Module no.	General description	Functions
0	Managing of input data	Adding to or deleting files in AURELIA's directory Listing and editing of spectral parameters
1	Enhancement of spectral quality	Mean ridge subtraction Mean row subtraction Antidiagonal suppression Removal of oscillatory artefacts Base-plane correction methods Symmetry enhancement methods Display and plotting
2	Local analysis of 2D spectra	Automated peak picking Automated multiplet recognition Interactive correction of peak and multiplet lists Volume calculation Classification of peaks by a Bayesian analysis Distance calculation from NOESY or ROESY data Simultaneous display of the 3D structure and the corresponding spectrum Simulation of NOESY spectra by relaxation matrix calculation Display and plotting
3	Nonlocal analysis of 2D spectra	Automated recognition of J-coupling networks Automated sequential assignment Transfer of peak and assignment lists Comparison of peak lists Spectral algebra Combining information from 2D and 3D spectra Display and plotting
4	Display and analysis of 3D spectra	Inspection of 3D data Mean base plane subtraction Base-plane correction methods Automated peak picking Automated integration of 3D peaks Automated sequential assignment in a set of 2D and 3D spectra Display and plotting
5	Display and analysis of 4D spectra	Inspection of 4D data sets Automated peak picking Automated integration of 4D peaks Display and plotting

ively, of two-dimensional NMR spectra, and modules 4 and 5 for the display and analysis of three- and four-dimensional NMR spectra, respectively. Some additional modules, like the module for the analysis of 1D and 2D NMR spectra of mixtures of low-molecular-weight compounds, are outside the scope of this paper and will not be described here.

#### *Module 0: Managing the input data*

In this part of the program, the user selects the n-dimensional spectra that are to be used. These spectra should be stored in binary integer format and should be organized in n-dimensional submatrices with flexible dimensions to allow for the fastest mean access to any portion of the data in any dimension. Parameters describing the spectra are stored in separate files. NMR spectra

that have been Fourier transformed by any of the numerous existing software packages and do not follow the rules mentioned above may also be used. Proper conversions are offered for various formats. Different byte orders are handled automatically. Since a resonance assignment procedure most likely involves a group of spectra rather than just one, special care has been taken to access such a group in the most convenient way. This is achieved by running the program together with a process or history file, which holds a full description of all involved spectra. Transferring and comparing information of the different spectra can thus be handled automatically by the program. This history file also allows the user to interrupt the data evaluation at any point and resume it later. Spectra may be added or removed from the history file at any stage of the analysis.

### *Module 1: Display and enhancement of spectral quality*

The quality of the input data significantly influences any automated data analysis. Therefore, a careful processing is very important. AURELIA provides various tools to overcome the typical problems encountered, such as base plane artefacts,  $t_1$  ridges or antidiagonals (Table 1).

A flat base plane is not only important for the correct integration of multidimensional NMR spectra, where base plane variations can dominate the integral, but also for peak recognition, where a threshold is typically defined to sort out resonance peaks from noise spikes. The fundamental assumption used in base plane-correcting procedures is that the base plane is ideally flat and that slope variations of resonance peaks are greater than those of base plane artefacts. Published base plane correction methods differ in the functions used for approximating base line artefacts and in the way that regions are defined where the ideal base line should be zero. Those regions which contain no cross peaks can either be defined by the user (Barsukov and Arseniev, 1987; Zolnai et al., 1989; Saffrich et al., 1993) or identified automatically by the program (Dietrich et al., 1991; Güntert and Wüthrich, 1992; Chylla and Markley, 1993). The latter methods are more convenient for the user; however, in interactively driven methods external information can be incorporated, e.g. evidence derived from other experiments, such as well-resolved 1D spectra. AURELIA provides two approximations of the base line, the cubic spline fit (Zolnai et al., 1989) and the sectionally linear interpolation (Saffrich et al., 1993). The cubic spline interpolation has the principal advantage that it approximates the base line by a smooth function; the linear interpolation, where the base plane is approximated by short sections of straight lines, has the advantage of being computationally very fast. In addition, it avoids overcorrections which often result from cubic interpolation. Both methods can be applied to two- and three-dimensional NMR spectra. Base line points where no signals are expected are read from a file. They can be interactively defined by inspecting suitable 1D spectra or projections of the 2D (and 3D) spectra. In the case of symmetric homonuclear spectra, one set of points can be used for a base plane correction in all dimensions. Other types of spectra require the definition of different sets of base line points for the individual dimensions.

Typical artefacts that often dominate spectra recorded with older instruments are  $t_1$  ridges. A simple method for their attenuation, the mean row subtraction, was devised by Klevit (1985). In this method, the user defines one or several regions in the 2D spectrum where no cross peaks but  $t_1$  ridges are present. The mean of these rows is calculated and then subtracted from all other rows. Initially, this method was devised for absolute-value spectra and it leads only to small spectral improvements in phase-sensitive NMR spectra where positive and negative intensities

partly cancel. AURELIA provides two additional variations of the original method, well suited to phase-sensitive spectra (Zolnai et al., 1986, 1988; Kalbitzer et al., 1991b). Positive and negative data points are treated individually during the calculation, and subtraction of the correction profile and different types of profiles, mean and projection, may be chosen.

The oscillatory components of ridges, which originate from truncation effects, can be effectively removed by a frequency-domain filter developed to suppress periodic features (Mitschang et al., 1990). The only information needed for this process is the degree of zero-filling, which determines the basic frequency of the oscillatory artefacts.

Attenuation of the diagonal peaks leads to better resolution of the cross peaks in their immediate vicinity. In phase-sensitive COSY spectra, diagonal peaks can be suppressed in the mixed time-/frequency-domain spectrum  $S(t_1, \omega_2)$  after transforming signals in the  $t_1$  direction into pure exponentials (shifting their frequencies to 0) and applying a filter that selectively suppresses nonoscillating signals (Friedrichs et al., 1991; Pelczer, 1991). AURELIA provides an alternative method that is useful for absolute-value spectra. An appropriately weighted mean diagonal peak is constructed by projecting the whole spectrum parallel to the diagonal. The resulting profile is then subtracted columnwise from the spectrum (Glaser and Kalbitzer, 1986). This procedure has the advantage of also significantly reducing  $t_1$  ridges that show strong intensity modulations parallel to the diagonal, as often found in ROESY spectra.

Antidiagonals in 2D NMR spectra, which originate from incorrect pulse phases (Simorre and Marion, 1990), are often found in spectra recorded with instruments using analogous phase shifters. Since their general structure is known, they can be simply subtracted from absolute-value spectra (Glaser and Kalbitzer, 1986). The algorithm implemented in AURELIA automatically determines the weighting factors necessary for this calculation from the ratio of the diagonal-to-antidiagonal intensities.

Inherent symmetries in multidimensional spectra provide redundant information that is useful for discriminating resonance signals from noise and artefacts. A more powerful method of enhancing pertinent information is the exploitation of global symmetries, such as cross-peak symmetries about the main diagonal present in many homonuclear spectra. The first symmetry enhancement procedures consisted of replacing the intensities  $I(i,j)$  and  $I(j,i)$  in symmetry-related positions by the minimum  $\min(I(i,j), I(j,i))$  (Baumann et al., 1981b). This method is computationally very fast and suppresses asymmetries in the spectrum very efficiently. However, it often suppresses useful information as well and must be applied with great care. A generalization of the method is implemented in AURELIA. This much more advanced symmetrization technique also takes into account information about the

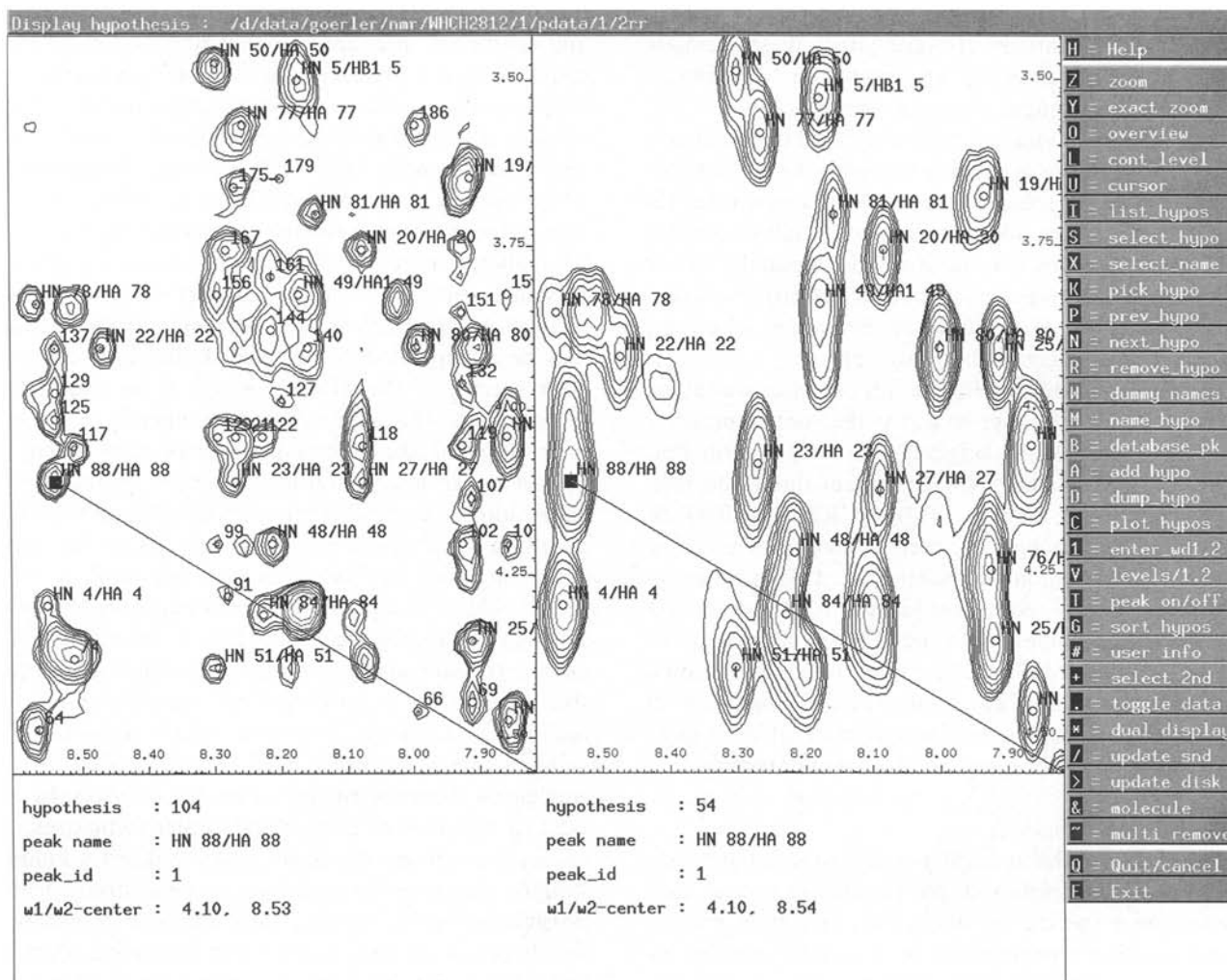


Fig. 1. Simultaneous display and handling of two NMR spectra. A NOESY and a TOCSY spectrum of HPr from *S. carnosus* are displayed. Annotations from the TOCSY spectrum were automatically transferred to the NOESY spectrum. Every peak in the spectrum is characterized by its hypothesis number, its name, its peak\_id (which defines the multiplet pattern), and its frequency coordinates. The cursor designates the peak under consideration.

environment of each data point. Instead of only comparing pairwise points  $I(i,j)$  and  $I(j,i)$ , symmetry-related two-dimensional regions containing such points are arranged as one-dimensional vectors. The cosine of the angle between each pair of vectors represents a normalized match factor, which describes the correlation of the corresponding relative intensity distributions. This match factor can be used in various ways (e.g., proportional, difference proportional, additive) (Neidig and Kalbitzer, 1988,1990; Kalbitzer et al., 1991b) to correct the intensities at positions  $(i,j)$  and  $(j,i)$ . This method is computationally intensive, but much less error prone than the simple symmetrization methods. By a proper selection of the correction technique, it may be carefully adapted to obtain the information desired. If symmetrization techniques are to be applied, the data processing should be chosen such that inherent symmetries are conserved (e.g., by using different window functions in each dimension).

Finally, resolution enhancement and reduction of unwanted components of the data can be achieved by a technique known from general image analysis, segmentation and local rescaling. According to the selected parameters, an automated peak picking is combined with an iterative region growing based on analyzing local intensity distributions. The method determines data points belonging to local peaks and zeros all others. Broad humps, noise spikes and the bottom parts of the peaks may be removed from the spectra. The intensities of the remaining data points may be mapped into a well-defined small range of values (Neidig and Kalbitzer, 1990). Segmentation provides a much better visual access to spectra that contain many overlapping regions and show large intensity variations. It also represents a powerful method for data compression, since only the peak areas having intensities different from zero have to be stored.

To achieve the optimum quality of a spectrum, any of

the methods provided in AURELIA may be combined and applied sequentially. However, since these methods are based on the detection and evaluation of patterns, which are then changed during a corrective process, the order of the operations is significant. Based on the underlying theories, in our experience the following order yields the best results: (1) removal of oscillatory components; (2) removal of  $t_1$  ridges; (3) reduction of antidiagonals; (4) removal of the diagonal and oscillations parallel to the diagonal; (5) base line correction; (6) symmetrization; and (7) segmentation. Depending on the actual spectrum, typically two or three methods are applied.

The display of the NMR data, for example as contour diagrams, allows the user to access the spectra properly and to cross-check the original data with information derived from it. As any NMR program should do (see, e.g., Johnson and Blevins, 1994), AURELIA offers all standard display techniques, such as contour smoothing by Bezier interpolation (Kalbitzer et al., 1991b), zooming techniques, overlay techniques (different 2D spectra, 1D and 2D spectra, planes of 3D and 4D spectra), magnification by real-time contour calculation, multiple windows, multiple cursors, and automatic spectral alignment in multiple overlay techniques if spectra with different spectral widths and calibrations are used simultaneously.

#### *Module 2: Local analysis*

Although it is theoretically possible to simultaneously recognize all multiplets and spin patterns in a set of multidimensional spectra by fitting the data to a general model function characterized by a suitable number of parameters, in practice this approach is only possible for very simple systems, involving only a few variables. The more economic way is to try to quickly reach a level of abstraction and reduce the size of the input data to be handled by the computer. Since the publication of the first program for performing pattern recognition in 2D NMR spectra of polypeptides (Neidig et al., 1984), this general strategy has been used almost exclusively. In practical applications, the signal-to-noise and signal-to-artefact ratios are usually insufficient to unambiguously observe all theoretically expected cross peaks. Even after artefact reduction, spectra may still contain invalid signals that erroneously may be regarded as true peaks. Whether a program can handle missing and erroneous information and still yield nonambiguous solutions of a problem finally decides whether it is useful in practice. Even just allowing the human expert to interfere at a critical stage might be the best strategy. Module 2 covers applications ranging from simple peak picking to automatic multiplet analysis and volume calculations with the following iterative strategy.

The most elementary objects AURELIA recognizes are local extrema (often simply called peaks), which are defined as data points having a greater absolute intensity

than their closest neighbouring data points. In 2D, 3D and 4D spectra there are 8, 26 and 80 closest neighbours, respectively. For reducing the possible number of such simple peaks, usually a minimum threshold is required (Neidig et al., 1984; Glaser and Kalbitzer, 1987; Novic and Bodenhausen, 1988; Pfändler and Bodenhausen, 1988a; Stoven et al., 1989; Eccles et al., 1991). With the help of the program, a suitable threshold can be obtained by analysing a set of defined regions where no true signals are expected. The simple peak picking results in a list of local peaks, which is fully accessible to the user and may be manipulated in detail. Peak lists from homonuclear spectra of the TOCSY type can for example be symmetrized. This approach can be superior to the symmetrization of the spectra itself, since now it may be limited to certain spectral areas and the degree of symmetrization is controlled interactively. The allowed variations of simple peak positions may always be chosen such that losing real signals is extremely unlikely. Areas that should be completely free of signals may be defined interactively. Depending on the type of spectrum, simple (singlet) peaks may or may not be sufficient for a further analysis. For example, in DQF spectra well-resolved multiplet patterns may be visible. Prior to any analysis of substructures of peaks, AURELIA first applies a direct, single-pass cluster technique to the list of peaks. In contrast to statistical or hierarchical cluster techniques, this means that groups of signals may be directly found if suitable clustering parameters can be defined. This is possible in NMR spectra, since pairwise distances of simple peaks are directly related to J-coupling constants and relative intensity ratios are related to ratios of couplings and chemical shift differences (e.g., weakly and strongly coupled systems). In both cases, reasonable upper limits can readily be specified. If no substructures of peaks are expected, special clusters containing just one single simple peak are calculated.

Only in ideal spectra with an optimum signal-to-noise ratio and nonoverlapping peaks, each cluster would correspond to just one multiplet and may be identified directly. In reality, this is almost never the case and clusters usually consist of groups of signals which may contain peaks from various multiplets or even partial multiplets, as well as various artefact peaks. A more sophisticated multiplet recognition procedure has therefore been implemented in AURELIA (Neidig et al., 1990; Kalbitzer et al., 1991b). First, arbitrary types of multiplets (currently containing up to 64 components) may be defined in a database. AURELIA reads and analyses the database entries and translates them into an internal description, containing representative features such as symmetry, shape, and sign distribution. For each multiplet these features are mapped to any cluster partitions that hold the same number of peaks as the multiplet (Neidig et al., 1990). This mapping procedure is independent of transla-

tions, homogeneous scaling and homogeneous distortions. It results in a match factor that describes the similarity of a group of experimental simple peaks and a defined multiplet. If the match factor exceeds a certain threshold value, the result is accepted. To overcome problems arising from the technical upper limits of cluster sizes, the multiplet search algorithm is implemented in an iterative way. After each iteration, successfully determined multiplets are removed from the list and the remaining peaks undergo the same procedure (starting with the cluster algorithm) again. This is repeated until a certain number of iterations has been performed or a certain amount of CPU time has been used. The results may be inspected by the user and verified, completed, annotated, fitted with Gaussian or Lorentzian functions to get J-couplings, or cross-checked with other spectra to any detail. Assuming that no procedure is perfect, AURELIA allows the interactive deletion, redefinition and creation of multiplets. Figure 1 shows a dual display of two corresponding regions of a TOCSY and a NOESY spectrum, which have automatically been aligned to the same ppm values. All peak annotations that were entered in the TOCSY spectrum have been transferred automatically to the NOESY spectrum and combined with proper annotations.

The basis for macromolecular structure determination by NMR is the distance calculation from NOE data. The magnitude of the NOE is proportional to the volume of the NOESY cross peak or, in well-resolved spectra, the volume of the cross-peak multiplet. The simplest method for estimating the relative volumes of cross peaks is based on the determination of the peak amplitudes. As long as the widths and shapes of two peaks are identical, the volumes are proportional to the amplitudes; here, the amplitudes can be used for the calculation of relative volumes and distances derived from the volumes. Unfortunately, this condition usually is not fulfilled since the cross peaks in NOESY spectra often show inhomogeneous broadening, which depends on the magnitude and number of J-couplings, and homogeneous broadening, which depends on the distribution and mobility of the nearby nuclei. Therefore, a reliable definition of peak areas and integration of such areas is needed. Most simply, peak integration areas are defined as rectangles, circles or ellipses using the mouse. All data points of such areas may then be summed up to yield the volumes. This method turns out to be very cumbersome if many thousands of peaks have to be integrated, and can hardly be used in 3D or 4D spectra. Therefore, it has not been chosen in AURELIA. A number of different methods, where the integration areas are determined automatically, are proposed in the literature (Denk et al., 1986; Stoven et al., 1989; Shen et al., 1990; Eccles et al., 1991). AURELIA uses an optimized version (Geyer, 1995; Geyer et al., 1995) of an iterative segmentation method combined with a region-growing algorithm (Glaser and

Kalbitzer, 1987; Neidig and Kalbitzer, 1990). It recognizes all data points that are part of a given cross peak by analysing local intensity distributions. A parallel local analysis of neighbouring peaks allows the analysis of overlapping peaks, provided their maxima are still separated. The segmentation may be fine-tuned by specifying suitable segmentation parameters, such as segmentation threshold, number of iterations and maximum expected integration areas. The volume is calculated as the sum of intensities of the segmented data points. This method works fully automatically and is very efficient. In AURELIA it is implemented for 2D, 3D and 4D spectra.

Distances can be calculated from the volumes in NOESY or ROESY spectra using different approaches. If no dependence on the mixing time is available, they can be calculated according to the initial slope approximation from the ratio of cross-peak volumes. The reference cross peak and the functional dependency are defined by the user. Alternatively, AURELIA allows the calculation of distances from the cross-peak/diagonal-peak ratio (Esposito and Pastore, 1988). This requires knowledge of the correlation time of the participating nuclei. If applied to a pair of nuclei with a known distance, this method may reversely be used to estimate the correlation time. If the dependence of the NOE on the mixing time has been measured, these data can be fitted with three different approximations, i.e., the quadratic polynomial approximation (Macura and Ernst, 1980), the exponential approximation, and the biexponential approximation. In the case of (noncompensated) ROESY

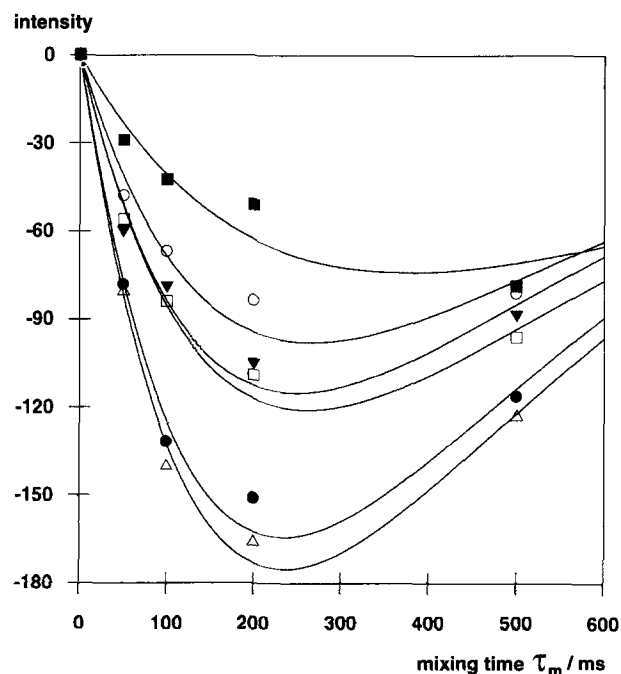


Fig. 2. Calculation and plotting of ROESY buildup curves. Shown are biexponential fits of ROESY buildup curves of the neuropeptide head activator. The curves are for sequential  $H^{\alpha}$ - $H^N$  ROEs of Pro<sup>3</sup> to Gly<sup>4</sup> (□), Ser<sup>6</sup> to Lys<sup>7</sup> (○), Lys<sup>7</sup> to Val<sup>8</sup> (▼), Val<sup>8</sup> to Ile<sup>9</sup> (Δ), Ile<sup>9</sup> to Leu<sup>10</sup> (●), and Leu<sup>10</sup> to Phe<sup>11</sup> (■) (Saffrich, 1991).



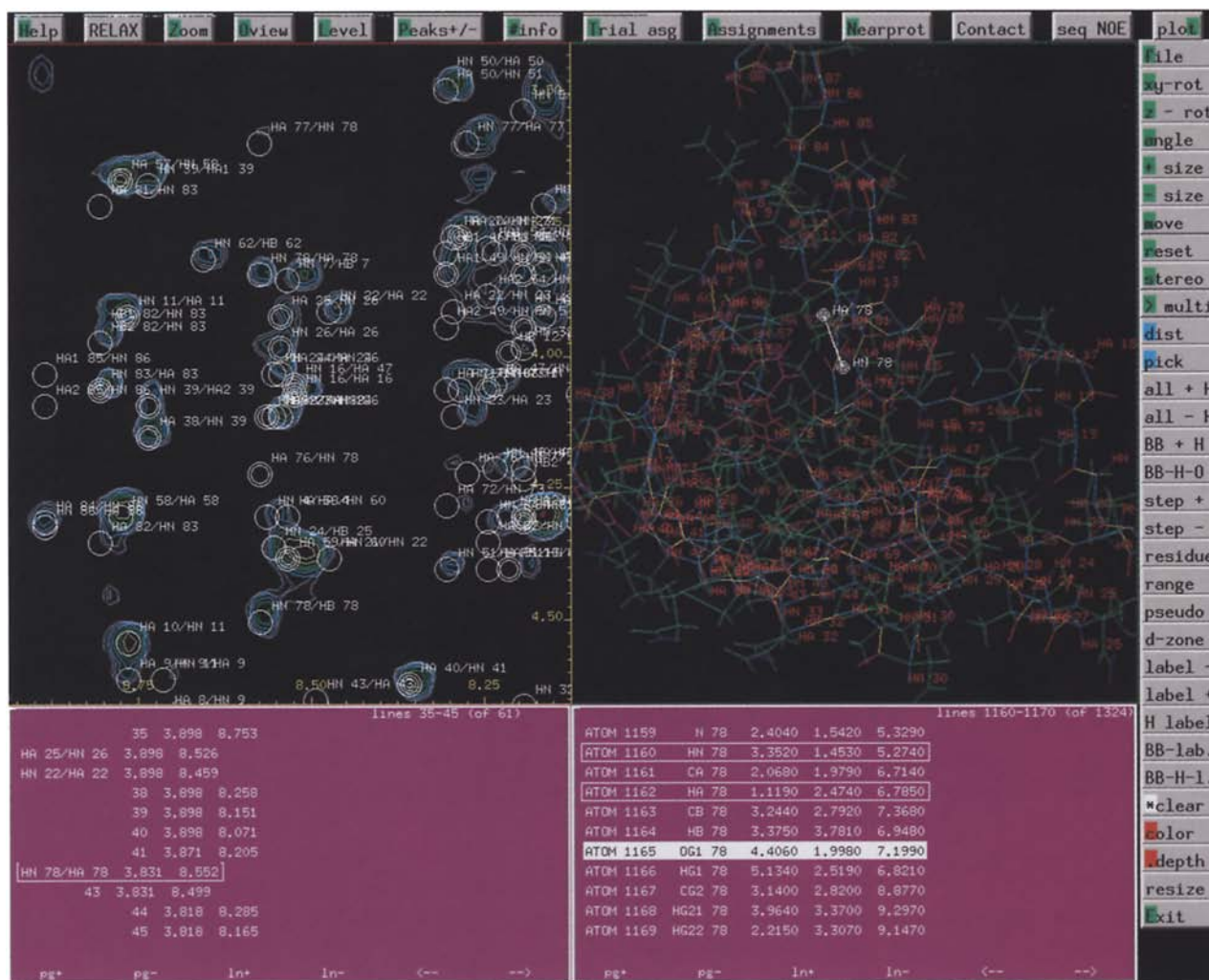


Fig. 3. Backcalculated NOESY spectrum. The figure shows a display of a NOESY spectrum of HPr from *S. carnosus* (mixing time 100 ms), the backcalculated cross peaks, and the actual structure used for the calculation.

spectra, the observed volumes are also dependent on the effective magnetic field experienced by the nuclei. Therefore, the strength and frequency of the  $B_1$  field relative to the observed resonances (assuming isolated spin pairs and predominantly dipolar relaxation) must be taken into account by calculating the offset correction (Bull, 1988). Figure 2 shows the ROESY buildup curves as calculated with AURELIA for some backbone protons of the neuropeptide head activator (Saffrich, 1991). As soon as the cursor is navigated to a cross peak and a mouse button is clicked, the buildup curve to which this peak contributes is shown. Any set of curves may be plotted on paper as well.

Especially in NOESY-type spectra, the question often arises whether certain signals are true peaks or not. If no other rules apply, the human user decides intuitively: the stronger the signals are and the more they compare in shape to other already known true signals, the more likely they may be accepted and used as additional restraints in

structure calculations. Apart from the fact that misinterpretations occur, this task may be very cumbersome. Therefore, various methods for automated peak assessment have been proposed in the past (e.g., Glaser and Kalbitzer, 1987; Kleywegt et al., 1990). We have implemented a new Bayesian method that is based on a statistical line shape analysis (but can be extended to any measurable quantity). It makes no special assumption about line shapes, although for an efficient discrimination the probability distributions for some of these features must differ for signals and noise. After a training phase, in which peaks of a noise and a signal area are analysed, a multivariate linear discriminant analysis is applied to obtain linear combinations of line shape parameters that discriminate best between true signals and noise. The final Bayesian analysis allows the estimation of the probability for a given peak to belong to either the signal or the noise class (Antz et al., 1995).

Frequently, the structure determination turns out to be





an iterative process. After most of the resonance lines have been assigned and sufficient volumes and distances have been obtained, one may try to calculate potential structures. Sometimes structures are created elsewhere, or structures of homologous proteins are available. In order to verify the consistency of a structure, even with NOESY peaks not yet assigned, the structure can be displayed simultaneously with the NMR spectra. Coordinate files of proteins and DNA/RNA stored in the PDB format may be loaded into AURELIA. Correlations between pairs of nuclei and peaks are obtained by combining peak lists, peak annotations and the coordinate files. Even better is a backcalculation from structure to NMR spectrum using a full relaxation matrix approach (Keepers and James, 1984). The algorithm implemented in AURELIA can handle a number of different relaxation models, such as the Lipari-Szabo model (Lipari and Szabo, 1982a,b), and is optimized for minimal computing time (Görler, 1995). For efficiency, it can be applied either to all or only to specific sets of nuclei of the molecule. Figure 3 shows a part of a NOESY spectrum of the HPr protein together with the three-dimensional structure. To reveal differences between the experimental data and the backcalculated spectrum, the two spectra are shown in dual display. Such differences can arise from errors in the sequential assignments of the resonances or from a discrepancy between the calculated and the true structure; the comparison of the calculated and experimental spectra can help to correct the possibly wrongly assigned resonances and to refine the three-dimensional structure. Many useful tools, such as creation of graphs of different kinds of NOEs versus the sequence and contact maps, are also available in this part of the program.

### *Module 3: Nonlocal analysis*

The nonlocal data analysis comprises a large number of useful tools, such as comparing peak lists of different spectra, transferring peaks or annotations between spectra, importing external peak information, performing spectral algebra with 2D spectra or defining and comparing spectral patterns. Central goals are the search for spin systems and the sequential assignment. Following the ideas of the standard procedures, AURELIA provides a number of different routines for automatic or interactive spin pattern recognition. To be most flexible, the user may first describe homonuclear spin patterns in a database as a set of diagonal and cross peaks having a certain number of links. For each diagonal peak an expected range of possible resonance positions is given. These can for example be derived from the means and 5th and 95th percentiles of all published assignments (Gross and Kalbitzer, 1988; Wishart et al., 1995). The number of ambiguous solutions can be reduced when the same spin system is simultaneously searched in different types of 2D spectra. The user can interactively define the cross-peak

patterns expected and the experimental spectra used for the search. Figure 4 shows as an example the three cross-peak patterns of the amino acid valine as they would be expected to occur in a DQF, an RCT and a TOCSY spectrum. The automatic spin system search algorithm first translates the database entries into an internal representation. The translation involves a decomposition of any pattern into a set of subpatterns: (1) the primary branch, which is the largest set of diagonal peaks that are contiguously connected by cross peaks; (2) secondary branches, starting out from diagonal peaks of the primary branch; (3) a set of internal cross peaks, each connecting two diagonal peaks of the primary branch; and (4) the remaining cross peaks, which connect diagonal peaks of different branches. The algorithm (which is not error-tolerant) starts out searching for the primary branch in a recursive procedure, using the lists of identified multiplets obtained in module 2. For any identified primary branch the secondary branches are searched for in the same way. The remaining cross peaks of patterns (3) and (4) can be identified by a simple look-up in the lists of multiplets. The check of peaks contained in other variations of a pattern is finally performed in a second pass. Cross peaks in symmetric positions are taken into account. Patterns are only accepted when the complete primary branch is found. A rich set of tools allows the user to verify, clean up and annotate the results. The spin system search may also be conducted interactively. In this case the user first navigates the cursor to a particular cross peak (e.g., an  $H^N-H^\alpha$  cross peak in a protein spectrum), which is then used as the starting point of the search. Automatic sequential assignments can be performed by combining the obtained patterns with NOE spectra. The spectra are overlaid, for example in the fingerprint region, and lists of NOE peaks and COSY/TOCSY peaks (or spin systems if available) are used to perform sequential walks. Following a standard strategy in  $\beta$ -sheet structures, the underlying algorithm starts from ( $H^N, H^\alpha$ ) peaks in the TOCSY spectrum and searches for peaks in the NOESY spectrum having the same  $\omega_1$  coordinate (giving a link to the  $\alpha$ -proton of the previous residue). If such a cross peak can be found, the algorithm switches back to the TOCSY spectrum and now searches for peaks having the same  $\omega_2$  coordinate, and so on. Similar strategies exist for other secondary structures. However, a main difficulty are ambiguities which arise in crowded regions of the three spectra. During the comparison of coordinates, the program may find more than one peak to follow the sequential walk. Whenever this occurs, the program stores all possible solutions whose number increases exponentially. A number of different strategies have been published that try to solve this problem and to find the most reasonable solution (Englander and Wand, 1987; Billeter et al., 1988; Cieslar et al., 1988; Eads and Kuntz, 1989; Kleywegt et al., 1989, 1990, 1991; Van de Ven, 1990; Van de Ven et al.,

1990; Billeter, 1991; Eccles et al., 1991; Nelson et al., 1991; Wand and Nelson, 1991; Craig and Kuntz, 1993; Wehrens et al., 1993; Hare and Prestegard, 1994; Olson and Markley, 1994). An alternative, which is used most often in practice, consists in the introduction of further information at critical stages. One such possibility has been implemented in AURELIA. The program is able to correlate different 3D peak lists, taken from homonuclear NOESY-TOCSY spectra, with the lists of NOE peaks and spin systems. The 3D peaks correlate three different nuclei, e.g.,  $H^N(i)$ ,  $H^\alpha(i)$  and  $H^N(i+1)$  or  $H^\alpha(i)$ ,  $H^N(i)$  and  $H^\alpha(i-1)$ . The 3D peak coordinates can be projected into two dimensions and can be represented as vertical or horizontal lines that connect NOESY and TOCSY peaks in the 2D spectra (Fig. 5). Since the overlap problem in three dimensions is less severe than in two dimensions, most of the ambiguous sequential pathways can be eliminated interactively by inspection of the data.

#### *Module 4: Display and evaluation of 3D NMR spectra*

The concepts of an image analysis applied to 2D NMR spectra as described so far apply for 3D and 4D spectra as well. In addition, more advanced display tools are now needed to allow a visual inspection of the data and the information derived from it. Furthermore, a data analysis is typically not only based on 3D or 4D spectra, but on the combination of 2D, 3D and 4D spectra. Therefore, more effort has been put into techniques that correlate NMR spectra of different dimensions.

For a general assessment of the basic quality of the 3D spectra, the localization of global artefacts and the first steps of a data analysis, it is important to be able to create a 3D view of the data set. AURELIA provides the necessary routines for the calculation, display, rotation, projection and zooming of 3D objects. Depending on the spectrum size, these routines may work on compressed data in order to achieve a satisfactory real-time performance. Moving 2D slices through a 3D dataset is a typical tool that requires real-time calculations (Fig. 6). Accessing orthogonal 2D planes, diagonal planes and projections can also be achieved by reading the proper data from disk. In this case no data compression is applied. The sub-cube organization of the 3D spectra on disk (Vuister et al., 1990) and an internal AURELIA cache mechanism allow an almost interactive walking through any dimension of the 3D data set.

In 3D spectra, the image enhancement preceding the data evaluation is even more important than in 2D spectra. For example,  $t_1$  noise ridges in 2D spectra may now appear as bundles of  $t_1$  noise planes, which could overload 3D peak lists. AURELIA provides a generalization of the mean row subtraction procedure described in module 1. It is called mean-plane subtraction and works quite similarly. The user defines a so-called mean plane area, i.e., a part of the 3D spectrum where no real peaks are

expected. Different types of projections or mean planes may then be calculated from this area and the resulting 2D plane is subtracted from the complete 3D spectrum (Fig. 7). The base line correction procedures described for 2D spectra (Saffrich et al., 1993) are also available for 3D spectra. Any combination of dimensions, or just the full spectrum, may be base-line corrected using a set of base line points that have for example been defined in corresponding 1D spectra. The 3D analysis, starting from simple 3D peak picking and ending with 3D peak annotation and 3D volume calculation, is implemented similar to the 2D case as described above. Simple peaks are defined as data points having a greater intensity than the closest 26 neighbours. Lists of such peaks may be interactively manipulated, based on proper displays. Since in common 3D applications multiplet splittings are usually suppressed by decoupling and cancelled by the low digital resolution, the cluster and multiplet analyses have only been implemented implicitly, assuming that the signals can be described as 3D singlets. Annotating 3D peaks is usually done by accessing 2D planes. The volume determination is done by an optimized iterative segmentation procedure (Geyer et al., 1995). Like in the case of two dimensions, it is completely automated, independent of the peak shape, accurate and fast.

While the tools described above are independent of the type of 3D spectrum, it has been found useful to implement specialized strategies that are applicable to different types of 3D spectra. In all cases a parallel usage of 2D spectra is possible. We want to describe two typical examples, the tracing of J-coupling networks via sets of triple resonance experiments and the application of 'strip' techniques (see, e.g., Kjaer et al., 1991) applied to 3D NOESY-HSQC spectra. Up to eight different 2D and 3D spectra may be correlated to yield sequential assignments. The 2D spectrum used could, for example, be a ( $^{15}\text{N}$ ,  $^1\text{H}$ )-correlated spectrum and the 3D set could consist of spectra such as HNCA, HNCOC and HN(CO)CA (Kay et al., 1990; Bax and Ikura, 1991). Picking correlated planes, multi-cursors and tracing orthogonal slices are the interactive tools to do this job. We have also implemented an automatic technique that is applicable to polypeptides. As found in some other published procedures (e.g., Olson and Markley, 1994), the central idea of our procedure is first to find suitable pairs of HNCA peaks representing correlations of  $N^H(i)$  peaks to  $C^\alpha(i)$  and  $C^\alpha(i-1)$  peaks. Pairs of such peaks having one of the  $C^\alpha$  frequencies correspond to neighbouring residues in the primary sequence. Since degeneracies either lead to collapsed pairs or to ambiguous NH or  $C^\alpha$  frequencies, it is almost impossible in practice to find unique sets of peaks corresponding to a group of sequential residues. The ambiguities may be reduced, but not fully overcome, by incorporating additional spectra, especially two-dimensional ( $^{15}\text{N}$ ,  $^1\text{H}$ )-correlated and three-dimensional HN(CO)CA

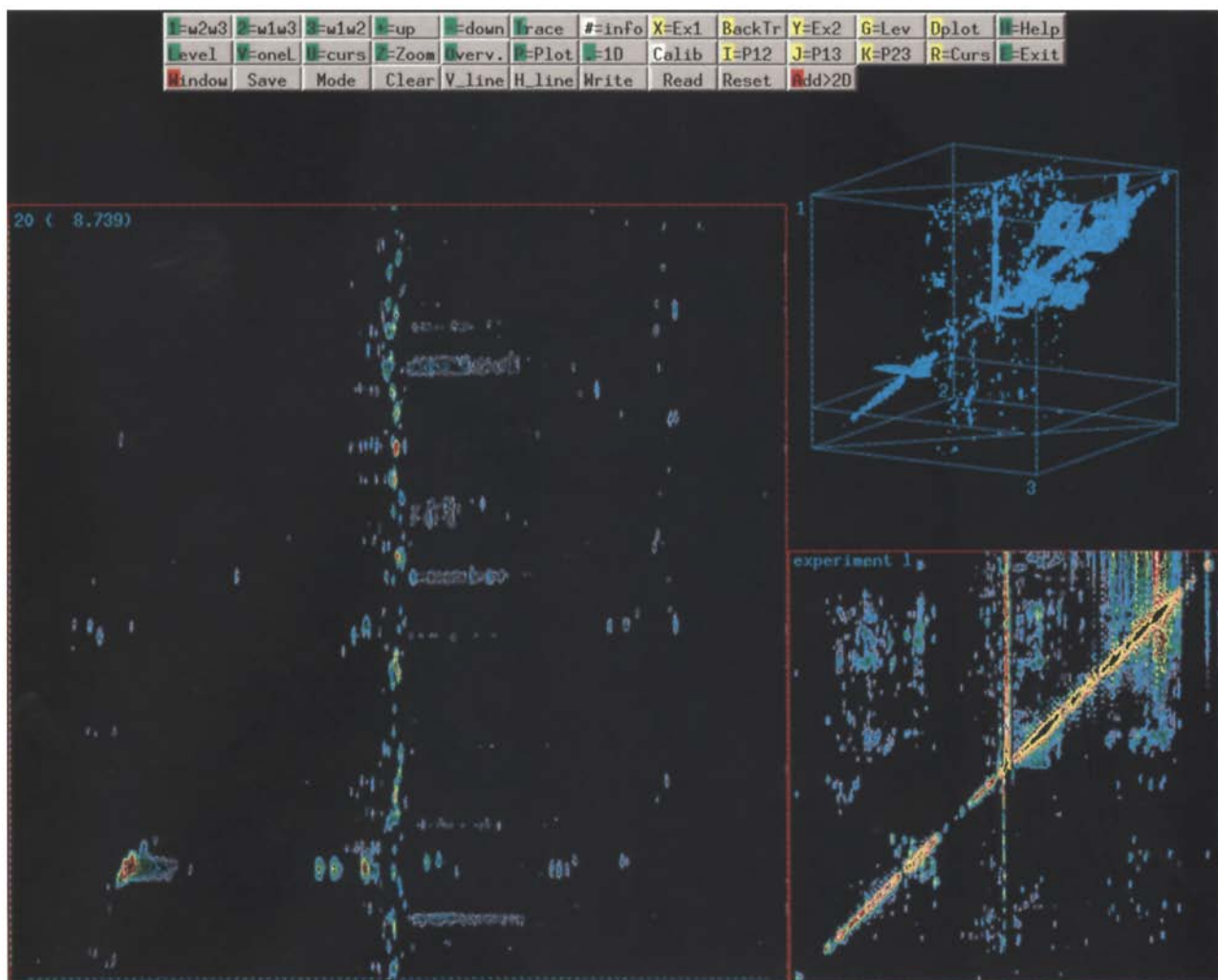


Fig. 6. Display of 3D NMR data. The figure shows a NOESY-TOCSY spectrum of HPr protein from *E. faecalis*. The data size before and after Fourier transformation was  $128 \times 128 \times 512$  points (as in Figs. 5 and 7); a sine multiplication was applied. The planes shown are indicated in the three-dimensional representation of the data.

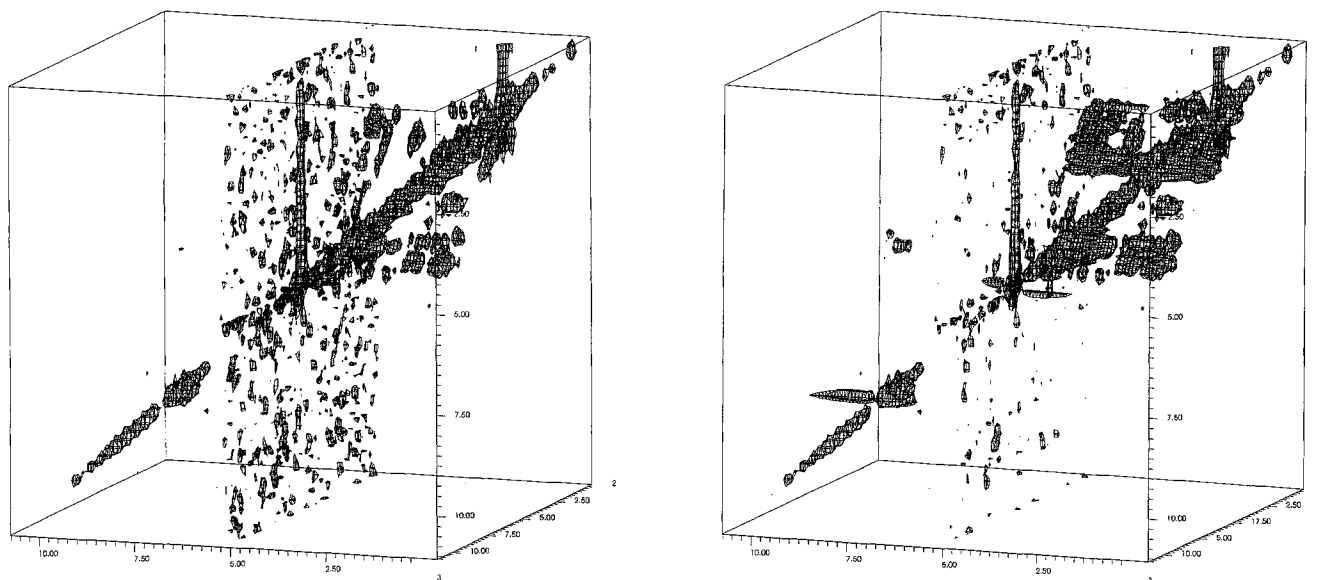


Fig. 7. Mean-plane subtraction of 3D data, used for noise suppression in a NOESY-TOCSY spectrum of HPr protein from *E. faecalis*. Left: spectrum before correction; right: spectrum after subtraction of an  $(\omega_2, \omega_3)$  mean plane. Data are plotted at the same contour levels.

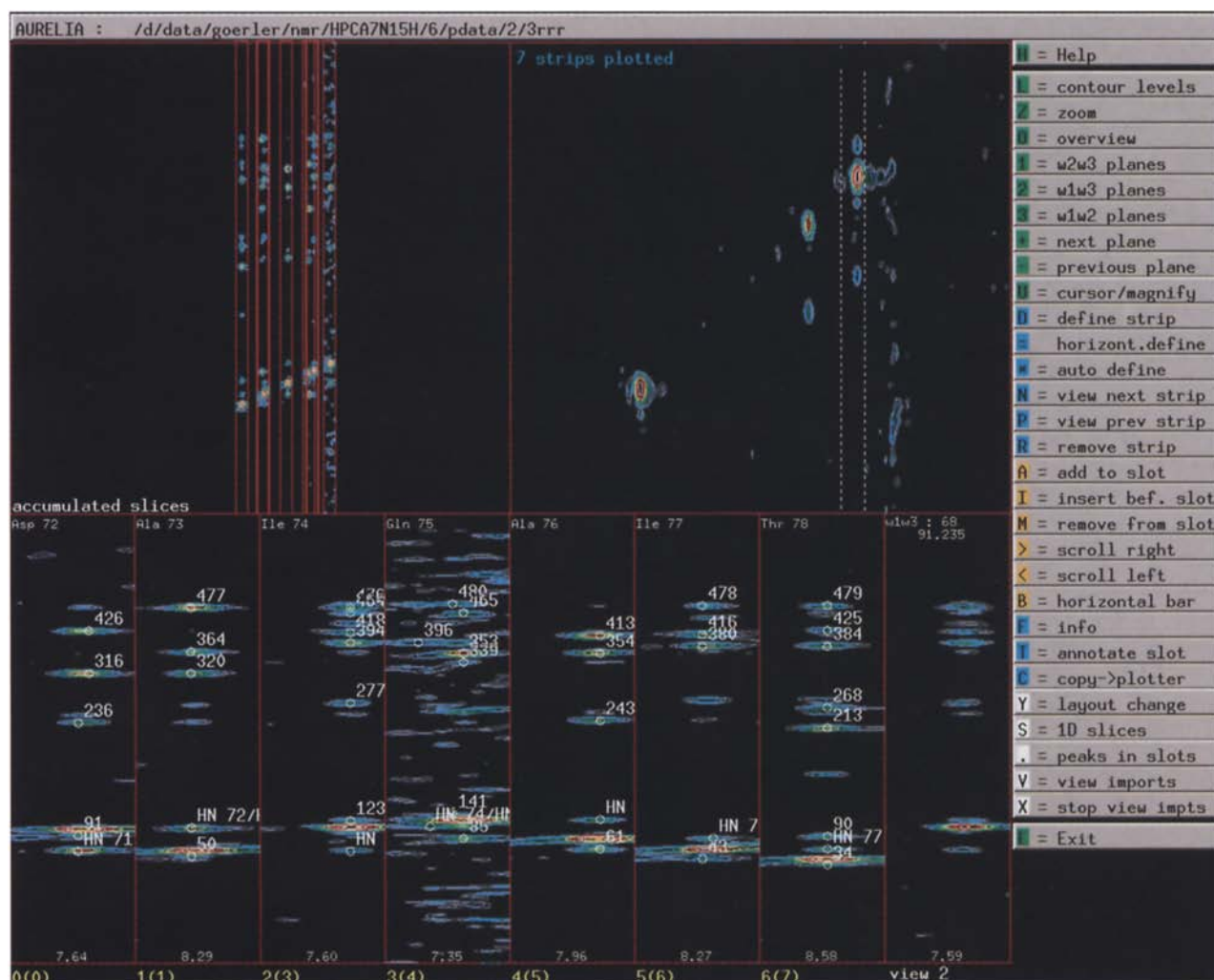


Fig. 8. Strip editor, used to sort a pool of strips. Sequential NOEs are shown in a NOESY-HSQC-spectrum of  $^{15}\text{N}$ -enriched HPr protein from *S. carnosus*. The data size before Fourier transformation was  $256 \times 112 \times 1024$  points, with spectral widths of 14.08 ppm, 40.00 ppm, and 14.08 ppm, respectively. A cosine multiplication was applied in the  $t_1$  direction and exponential multiplications in the  $t_2$  and  $t_3$  directions. The data size after Fourier transformation was  $256 \times 128 \times 1024$  points.

and CBCA(CO)NH spectra (Grzesiek and Bax, 1992). However, if such spectra have been obtained under different experimental conditions, this may lead to additional problems such as differences of resonance frequencies in the spectra. One of these factors is represented by differences in digital resolution from spectrum to spectrum and from dimension to dimension. This problem can partly be solved by allowing different deviations of the chemical shifts, depending on whether peaks in a spectrum or peaks of different spectra are compared. For cases where this is not sufficient for unique solutions, another strategy has been implemented as the second part of the sequential assignment procedure, based on some fixed, predefined information and on the primary sequence of the polypeptide. The user may load a set of already assigned resonances, such as the N,  $\text{H}^{\text{N}}$  and  $\text{C}^{\alpha}$  resonances of the residues Gly, Ala, Thr and Ser. These assignments do not need to be sequence specific, i.e., it is not necessary to

know which  $\text{C}^{\alpha}$  belongs to which glycine. Typical sources to obtain such information are HCCH-TOCSY, HCACO and 2D correlated experiments. If a possible sequence of HNCA pairs is generated during the analysis, it first gets cross-checked with the preloaded resonances. If some of them are identified, a hypothetical piece of sequence is generated like Xxx Xxx Gly Xxx Xxx Xxx Ala Xxx Xxx Xxx Xxx Xxx Gly Xxx, where Xxx represents the unidentified spin systems. This pattern is now matched to the primary sequence of the polypeptide. If a match is obtained, i.e., in the above example if the primary sequence contains a set of residues with Gly at position  $i$ , Ala at  $(i+4)$  and another Gly at  $(i+10)$ , then the result is assigned a high score value, otherwise it is assigned a low score value. The exact values are derived from the length of the pattern and the number of fixed residues contained in the pattern. The results are sorted by their score values. The limitations of this strategy are that (1) patterns must be



sufficiently long; (2) enough fixed resonances must be available; and (3) the fixed residues must be distributed suitably over the full sequence.

Strips are small portions of 2D planes derived from 3D spectra. For example, a strip in a NOESY-HSQC spectrum might be taken at a certain proton frequency in a nitrogen plane. The width of the strip usually is small, e.g., 20 data points, whereas the height is the full height of the plane. Strips may also have a depth. A depth of 1 means that a strip is taken from just one plane. Since 3D peaks extend in all dimensions and since strips should contain the best representations of such peaks, one may decide to sum up a few (=depth) planes before taking the strips. Representing 3D spectra as a set of strips eliminates empty spaces in 3D spectra. We have implemented interactive and automated tools to define such strips. All defined strips are collected in a pool. A strip editor may then be used to sort them into any order (Fig. 8). The sorting is done by comparing positions of peaks of different strips. Again, the program can do most of the job automatically. For example, reading out strips can be done on the basis of suitable 2D peak lists or on the basis of 3D peak lists. The user may specify additional criteria, like reading strips at the strongest peaks or reading only strips that contain a minimum number of peaks. Families of strips coming from different 2D and 3D spectra may be identified, stored and accessed as a group.

#### Module 5: Display and evaluation of 4D NMR spectra

Most recently, the 4D module of AURELIA has been added. It offers all tools to analyze 4D spectra. Most simply, the user may read 2D planes and display and plot them. Six different types of such planes are of importance. Apart from the type, two fixed frequencies along the two other perpendicular directions must be specified. This makes a visualisation of all planes very tedious. To check in advance which planes should be visualised at all, one may first look at one of the 12 types of 2D projections. Accessing any 2D plane is very fast; in a 128 MB 4D spectrum it takes typically less than a second. The calculation of projections requires scanning through larger portions of the data set and typically needs a few seconds. It turns out that the larger the dimensionality of a spectrum, the more important it becomes that the program assists the analysis. In this sense, the 4D peak picking is mandatory. We have implemented it such that data points having intensities greater than their closest 80 neighbours are taken as 4D singlets. 4D peak lists can be sorted, cleaned, annotated and printed. The 4D volume calculation again applies the iterative segmentation procedure already mentioned in the 2D and 3D cases.

### Conclusions and Outlook

Computer-aided evaluation of multidimensional NMR

data will continue to grow in importance. It is clear that this approach is superior to a data analysis performed exclusively by hand, especially if 3D and 4D spectra are involved. A fully automated data analysis for small molecules can be expected in the foreseeable future. To our knowledge, the currently existing programs that support an automated data analysis, including AURELIA, need some aid from the human expert and the results are not free of ambiguities and need manual checking. Therefore, the strategy of AURELIA is to use the possibilities of the computer and the human expertise combined and most efficiently. The fact that AURELIA has been developed on the basis of the ideas of the general image analysis, rather than providing limited solutions for specific cases of the data analysis, has been found to be very important.

### Acknowledgements

We are grateful to the Deutsche Forschungsgemeinschaft and the Bundesministerium für Forschung und Technologie for supporting this work.

### References

- Antz, C., Neidig, K.-P. and Kalbitzer, H.R. (1995) *J. Biomol. NMR*, **5**, 287–296.
- Barsukov, I.L. and Arseniev, A.S. (1987) *J. Magn. Reson.*, **73**, 148–149.
- Baumann, R., Kumar, A., Ernst, R.R. and Wüthrich, K. (1981a) *J. Magn. Reson.*, **44**, 76–83.
- Baumann, R., Wider, G., Ernst, R.R. and Wüthrich, K. (1981b) *J. Magn. Reson.*, **44**, 402–406.
- Bax, A. and Ikura, M. (1991) *J. Biomol. NMR*, **1**, 99–104.
- Bernstein, R., Cieslar, C., Ross, A., Oschkinat, H., Freund, J. and Holak, T.A. (1993) *J. Biomol. NMR*, **3**, 245–251.
- Billeter, M., Basus, V.J. and Kuntz, I.D. (1988) *J. Magn. Reson.*, **76**, 400–415.
- Billeter, M. (1991) In *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy* (Eds, Hoch, J.C., Redfield, C. and Poulsen, F.M.), Plenum Press, New York, NY, pp. 279–290.
- Bolton, P.H. (1985) *J. Magn. Reson.*, **64**, 352–355.
- Bolton, P.H. (1986a) *J. Magn. Reson.*, **67**, 391–395.
- Bolton, P.H. (1986b) *J. Magn. Reson.*, **70**, 344–349.
- Brown, J.W. and Huestis, W.H. (1994) *J. Biomol. NMR*, **4**, 645–652.
- Bull, T.E. (1988) *J. Magn. Reson.*, **80**, 470–481.
- Chylla, R.A. and Markley, J.L. (1993) *J. Magn. Reson. Ser. B*, **102**, 148–154.
- Cieslar, C., Clore, G.M. and Gronenborn, A.M. (1988) *J. Magn. Reson.*, **80**, 119–127.
- Cieslar, C., Holak, T.A. and Oschkinat, H. (1990) *J. Magn. Reson.*, **89**, 184–190.
- Craig, E.C. and Kuntz, I.D. (1993) *J. Magn. Reson. Ser. B*, **102**, 265–284.
- Darba, P. (1992) *Chem. Design Autom. News*, **7**, 16–22.
- Denk, W., Baumann, R. and Wagner, G. (1986) *J. Magn. Reson.*, **67**, 386–390.
- Dietrich, W., Rüdell, C.H. and Neumann, M. (1991) *J. Magn. Reson.*, **91**, 1–11.
- Eads, C.D. and Kuntz, I.D. (1989) *J. Magn. Reson.*, **82**, 467–482.



- Eccles, C., Güntert, P., Billeter, M. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 111–130.
- Eggenberger, U., Pfändler, P. and Bodenhausen, G. (1988) *J. Magn. Reson.*, **77**, 192–196.
- Eggenberger, U. and Bodenhausen, G. (1989) *Anal. Chem.*, **61**, 2298–2306.
- Englander, S.W. and Wand, A.J. (1987) *Biochemistry*, **26**, 5953–5958.
- Esposito, G. and Pastore, A. (1988) *J. Magn. Reson.*, **76**, 331–336.
- Friedrichs, M.S., Metzler, W.J. and Mueller, L. (1991) *J. Magn. Reson.*, **95**, 178–183.
- Garrett, D.S., Powers, R., Gronenborn, A.M. and Clore, G.M. (1991) *J. Magn. Reson.*, **95**, 214–220.
- Geyer, M. (1995) Ph.D. Thesis, University of Heidelberg, Heidelberg.
- Geyer, M., Neidig, K.-P. and Kalbitzer, H.R. (1995) *J. Magn. Reson. Ser. B*, **109**, 31–38.
- Glaser, S. and Kalbitzer, H.R. (1986) *J. Magn. Reson.*, **68**, 350–354.
- Glaser, S. and Kalbitzer, H.R. (1987) *J. Magn. Reson.*, **74**, 450–463.
- Görler, A. (1995) Diploma thesis, University of Heidelberg, Heidelberg.
- Grahn, H., Delaglio, F., Delsuc, M.A. and Levy, G.C. (1988) *J. Magn. Reson.*, **77**, 294–307.
- Gross, K.-H. and Kalbitzer, H.R. (1988) *J. Magn. Reson.*, **76**, 87–99.
- Grzesiek, S. and Bax, A. (1992) *J. Am. Chem. Soc.*, **114**, 6291–6293.
- Güntert, P. and Wüthrich, K. (1992) *J. Magn. Reson.*, **96**, 403–407.
- Güntert, P., Berndt, K.D. and Wüthrich, K. (1993) *J. Biomol. NMR*, **3**, 601–606.
- Hare, B.J. and Prestegard, J.H. (1994) *J. Biomol. NMR*, **4**, 35–46.
- Hoch, J.C., Hengxi, S., Kjaer, M., Ludvigsen, S. and Poulsen, F.M. (1987) *Carlsberg Res. Commun.*, **52**, 111–122.
- Holak, T.A., Scarsdale, J.B. and Prestegard, J.H. (1987) *J. Magn. Reson.*, **74**, 546–549.
- Johnson, B.A. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.
- Kalbitzer, H.R., Hengstenberg, W., Rösch, P., Muss, P., Bernsmann, P., Engemann, R., Dörschug, M. and Deutscher, J. (1982) *Biochemistry*, **21**, 2879–2885.
- Kalbitzer, H.R., Neidig, K.-P. and Hengstenberg, W. (1991a) *Biochemistry*, **30**, 11186–11192.
- Kalbitzer, H.R., Neidig, K.-P., Geyer, M., Saffrich, R. and Lorenz, M. (1991b) In *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy* (Eds, Hoch, J.C., Redfield, C. and Poulsen, F.M.), Plenum Press, New York, NY, pp. 175–190.
- Kalbitzer, H.R. and Hengstenberg, W. (1993) *Eur. J. Biochem.*, **216**, 205–214.
- Kalbitzer, H.R. (1994) In *Computer-Aided Analysis of Multidimensional NMR Spectra in Two-Dimensional NMR Spectroscopy: Applications for Chemists and Biochemists* (Eds, Croasmun, W.R. and Carlson, R.), Verlag Chemie, Weinheim, pp. 581–618.
- Kay, L., Ikura, M., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **89**, 496–514.
- Keepers, J.W. and James, T.L. (1984) *J. Magn. Reson.*, **57**, 404–426.
- Kjaer, M. and Poulsen, F.M. (1991) *J. Magn. Reson.*, **94**, 659–603.
- Kjaer, M., Anderson, K.V., Ludvigsen, S., Shen, H., Windekilde, D., Sorensen, B. and Poulsen, F.M. (1991) In *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy* (Eds, Hoch, J.C., Redfield, C. and Poulsen, F.M.), Plenum Press, New York, NY, pp. 291–302.
- Kjaer, M., Andersen, K.V. and Poulsen, F.M. (1994) *Methods Enzymol.*, **239**, 288–307.
- Klevit, R.E. (1985) *J. Magn. Reson.*, **62**, 551–555.
- Kleywegt, G.J., Lamerichs, R.M.J.N., Boelens, R. and Kaptein, R. (1989) *J. Magn. Reson.*, **85**, 186–197.
- Kleywegt, G.J., Boelens, R. and Kaptein, R. (1990) *J. Magn. Reson.*, **88**, 601–608.
- Kleywegt, G.J., Boelens, R., Cox, M., Llinas, M. and Kaptein, R. (1991) *J. Biomol. NMR*, **1**, 23–47.
- Kleywegt, G.J., Vuister, G.W., Padilla, A., Knegt, R.M.A., Boelens, R. and Kaptein, R. (1993) *J. Magn. Reson. Ser. B*, **102**, 166–176.
- Kraulis, P.J. (1989) *J. Magn. Reson.*, **84**, 627–633.
- Kraulis, P.J. (1994) *J. Mol. Biol.*, **243**, 696–718.
- Kruse, R., Hengstenberg, W., Beneicke, W. and Kalbitzer, H.R. (1993) *Protein Eng.*, **6**, 417–423.
- Levy, G.C., Delaglio, F., Macur, A. and Begemann, J. (1986) *Comput. Enh. Spectrosc.*, **3**, 1–12.
- Lipari, G. and Szabo, A. (1982a) *J. Am. Chem. Soc.*, **104**, 4546–4558.
- Lipari, G. and Szabo, A. (1982b) *J. Am. Chem. Soc.*, **104**, 4559–4570.
- Macura, S. and Ernst, R.R. (1980) *Mol. Phys.*, **41**, 95–117.
- Manoleras, N. and Norton, R.S. (1992) *J. Biomol. NMR*, **2**, 485–494.
- Markley, J.L., Darba, P., Fejzo, J., Krezel, A.M., Macura, S., McNemar, C.W., Mooberry, E.S., Seavey, B.R., Westler, W.M. and Zolnai, Z. (1991) In *Computational Aspects of the Study of Biological Macromolecules by NMR Spectroscopy* (Eds, Hoch, J.C., Redfield, C. and Poulsen, F.M.), Plenum Press, New York, NY, pp. 39–50.
- Meadows, R.P., Olejniczak, E.T. and Fesik, S.W. (1994) *J. Biomol. NMR*, **4**, 79–96.
- Meier, B.U. and Ernst, R.R. (1988) *J. Magn. Reson.*, **79**, 540–546.
- Meier, B.U., Mädi, Z.L. and Ernst, R.R. (1988) *J. Magn. Reson.*, **74**, 565–573.
- Mitschang, L., Neidig, K.-P. and Kalbitzer, H.R. (1990) *J. Magn. Reson.*, **90**, 359–362.
- Neidig, K.-P., Bodenmüller, H. and Kalbitzer, H.R. (1984) *Biochem. Biophys. Res. Commun.*, **125**, 1143–1150.
- Neidig, K.-P. and Kalbitzer, H.R. (1988) *Magn. Reson. Chem.*, **26**, 848–851.
- Neidig, K.-P. and Kalbitzer, H.R. (1990) *J. Magn. Reson.*, **88**, 155–160.
- Neidig, K.-P., Saffrich, R., Lorenz, M. and Kalbitzer, H.R. (1990) *J. Magn. Reson.*, **89**, 543–552.
- Neidig, K.-P. and Kalbitzer, H.R. (1991) *J. Magn. Reson.*, **91**, 155–161.
- Nelson, S.J., Schneider, D.M. and Wand, A.J. (1991) *Biophys. J.*, **59**, 1113–1122.
- Novic, M. and Bodenhausen, G. (1988) *Anal. Chem.*, **60**, 582–591.
- Novic, M., Eggenberger, U. and Bodenhausen, G. (1988) *J. Magn. Reson.*, **77**, 394–400.
- Olejniczak, E.T., Gampe, R.T. and Fesik, S.W. (1989) *J. Magn. Reson.*, **81**, 178–185.
- Olson, J.B. and Markley, J.L. (1994) *J. Biomol. NMR*, **4**, 385–410.
- Oschkinat, H., Holak, T.A. and Cieslar, C. (1991) *Biopolymers*, **31**, 699–712.
- Oschkinat, H. and Croft, D. (1994) *Methods Enzymol.*, **239**, 308–318.
- Pelczar, I. (1991) *J. Am. Chem. Soc.*, **113**, 3211–3212.
- Pfändler, P. and Bodenhausen, G. (1988a) *Magn. Reson. Chem.*, **26**, 888–894.
- Pfändler, P. and Bodenhausen, G. (1988b) *J. Magn. Reson.*, **79**, 99–123.
- Pfändler, P. and Bodenhausen, G. (1990) *J. Magn. Reson.*, **87**, 26–45.
- Rouh, A., Louis-Joseph, A. and Lallemand, J.-Y. (1994) *J. Biomol. NMR*, **4**, 505–518.
- Saffrich, R., Kalbitzer, H.R., Bodenmüller, H., Muhn, P., Pipkorn, R. and Schaller, H.C. (1989) *Biochim. Biophys. Acta*, **997**, 144–153.
- Saffrich, R. (1991) Ph.D. Thesis, University of Heidelberg, Heidelberg.
- Saffrich, R., Beneicke, W., Neidig, K.-P. and Kalbitzer, H.R. (1993) *J. Magn. Reson. Ser. B*, **101**, 304–308.

- Shen, H. and Poulsen, F.M. (1992) *J. Magn. Reson.*, **97**, 385–390.
- Shen, H., Ludvigsen, S. and Poulsen, F.M. (1990) *J. Magn. Reson.*, **90**, 346–354.
- Simorre, J.-P. and Marion, D. (1990) *J. Magn. Reson.*, **89**, 191–197.
- Sole, P., Delaglio, F. and Levy, G.C. (1988) *J. Magn. Reson.*, **80**, 517–519.
- Stoven, V., Mikou, A., Piveteau, D., Guittet, E. and Lallemand, J.-Y. (1989) *J. Magn. Reson.*, **82**, 163–168.
- Van de Ven, F.J.M. (1990) *J. Magn. Reson.*, **86**, 633–644.
- Van de Ven, F.J.M., Lycksell, P.-O., Van Kammen, A. and Hilbers, C.W. (1990) *Eur. J. Biochem.*, **190**, 583–591.
- Vuister, G.W., Boelens, R., Padilla, A., Kleywegt, G.J. and Kaptein, R. (1990) *Biochemistry*, **29**, 1829–1839.
- Wagner, G. and Wüthrich, K. (1982) *J. Mol. Biol.*, **155**, 347–366.
- Wand, A.J. and Nelson, S.J. (1991) *Biophys. J.*, **59**, 1101–1112.
- Weber, P. and Mueller, P.L. (1989) *J. Magn. Reson.*, **81**, 430–434.
- Wehrens, R., Lucasius, C., Buydens, L. and Kateman, G. (1993) *J. Chem. Inf. Comput. Sci.*, **33**, 245–251.
- Wishart, D.S., Bigam, C.G., Holm, A., Hodges, R.S. and Sykes, B.D. (1995) *J. Biomol. NMR*, **5**, 67–81.
- Zimmerman, D., Kulikowski, C., Wang, L.Z., Lyons, B. and Montelione, G.T. (1994) *J. Biomol. NMR*, **4**, 241–256.
- Zolnai, Z., Macura, S. and Markley, J.L. (1986) *Comput. Enh. Spectr.*, **3**, 141–145.
- Zolnai, Z., Macura, S. and Markley, J.L. (1988) *J. Magn. Reson.*, **80**, 60–70.
- Zolnai, Z., Macura, S. and Markley, J.L. (1989) *J. Magn. Reson.*, **82**, 496–504.
- Zolnai, Z., Westler, W.M., Ulrich, E.L. and Markley, J.L. (1990) *J. Magn. Reson.*, **88**, 511–512.