Journal of Biomolecular NMR, 7 (1996) 207–213
ESCOM

J-Bio NMR 338

Automated sequence-specific NMR assignment of homologous proteins using the program GARANT

Christian Bartels, Martin Billeter, Peter Güntert and Kurt Wüthrich*

Institut für Molekularbiologie und Biophysik, Eidgenössische Technische Hochschule-Hönggerberg, CH-8093 Zürich, Switzerland

Received 2 January 1996
Accepted 29 January 1996

Keywords: NMR structure determination; Automated sequence-specific resonance assignment; Homologous proteins

Summary

The program GARANT (General Algorithm for Resonance AssignmeNT) for automated sequence-specific NMR assignment of proteins is based on the mapping of peaks predicted from the amino acid sequence onto the peaks observed in multidimensional spectra [C. Bartels, P. Güntert, M. Billeter and K. Wüthrich (1996) *J. Comput. Chem.*, manuscript submitted for publication]. In this paper we demonstrate the potential of GARANT for the assignment of homologous proteins when either the three-dimensional structure or the chemical shifts of the parent protein are known. In these applications, GARANT utilizes supplementary information either in the form of interatomic distances derived from the three-dimensional structure, in order to add nuclear Overhauser effects reflecting the tertiary structure to the list of expected peaks, or in the form of the chemical shifts of the parent protein, in order to obtain a better estimate of the positions of the expected peaks. The procedure is illustrated with three different proteins: (i) a mutant form of Tendamistat (74 residues), using homonuclear 2D ^{1}H NMR spectra and either the three-dimensional structure or the chemical shifts of the wild-type protein; (ii) the mutant *Antp(C39S,W56S)* homeodomain (68 residues), using homonuclear 2D ^{1}H NMR spectra and the three-dimensional structure of the *Antp(C39S)* homeodomain; and (iii) free cyclophilin A (165 residues), using heteronuclear 3D NMR spectra and the three-dimensional structure of a cyclophilin A–cyclosporin A complex. In these three systems nearly complete assignment of the polypeptide backbone resonances and assignment of over 80% of the amino acid side-chain resonances was obtained without manual intervention.

Introduction

Projects using the techniques of structural biology often involve serial studies of closely related three-dimensional molecular structures, such as homologous proteins or complexes of a given protein or nucleic acid with different ligands. High-resolution NMR, one of the two principal methods for atomic-resolution structure determination of biological macromolecules (Wüthrich, 1986, 1995a,b), could greatly gain in its daily use by the introduction of tools that would automate the spectral interpretation once the structure determination for a parent protein in a homologous series has been completed. For example, the 3D structure of one or several homologous proteins may provide information on short ^{1}H-^{1}H distances that are likely to occur in the protein under investigation and thus enable more precise and extensive prediction of peaks expected in NOESY spectra, or the chemical shifts of a homologous protein may provide a more reliable estimate for the shifts of the new protein than could otherwise be obtained. To our knowledge, no systematic and fully automated way of using such additional information for sequence-specific NMR assignment of homologous proteins has so far been described. Here we propose such an approach, which is based on the program GARANT (General Algorithm for Resonance

*To whom correspondence should be addressed.

Abbreviations: 1D, 2D, 3D, one-, two-, three-dimensional; NOE, nuclear Overhauser enhancement; NOESY, nuclear Overhauser enhancement spectroscopy; TOCSY, total correlation spectroscopy; COSY, correlation spectroscopy; *Antp(C39S,W56S)*, mutant *Antennapedia* homeodomain with the residues Cys^{39} and Trp^{56} replaced by serine.

0925-2738/$ 6.00 + 1.00 © 1996 ESCOM Science Publishers B.V.

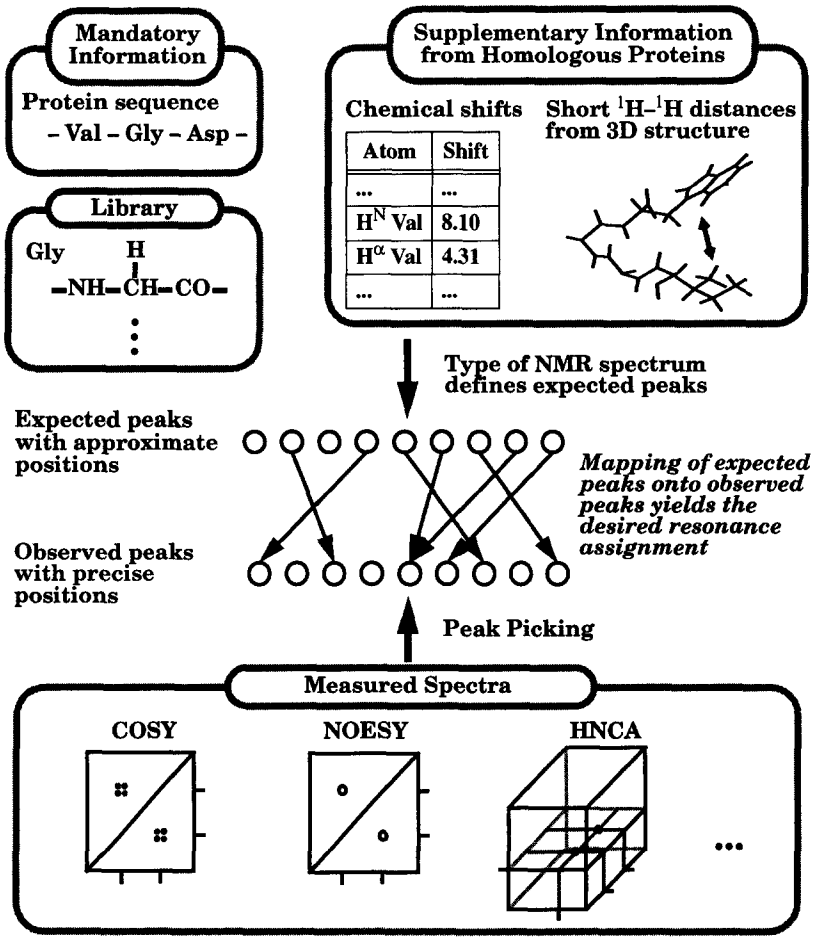

Fig. 1. The general concept of the program GARANT for automated sequence-specific NMR assignment of proteins. Boxes at the top represent the sources upon which prediction of expected peaks is based. The two boxes in the upper left represent mandatory information that is provided by a program library and by the amino acid sequence. The box in the upper right lists supplementary information that may be obtained from homology considerations. The box at the bottom represents the experimental NMR spectra from which the observed peak positions are extracted. The central part of the drawing indicates the search for optimal mapping of expected and observed peaks, which eventually yields the desired assignments (see text).

AssignmeNT) (Bartels et al., 1996). GARANT allows the combined use of NMR cross-peaks observed in different NMR spectra and supplementary information derived from homologous proteins. In addition to a conceptual outline of the use of GARANT for this purpose, the potential of the method is illustrated with applications to several systems of protein analogs.

Methods

The algorithm GARANT for automated sequence-specific resonance assignment

The search for sequence-specific resonance assignments in proteins implemented in GARANT involves firstly the construction of a list of expected peaks based on the amino acid sequence, and secondly the mapping of these peaks onto the experimentally observed peaks (Fig. 1; Billeter, 1991; Bartels et al., 1996). Cross-peaks due to scalar couplings can be predicted from knowledge of the amino acid sequence, the covalent structure of all amino acid residues, and the magnetization pathways of a given NMR experiment. For example, ^{1}H-^{1}H COSY peaks are expected for all pairs of protons (or groups of equivalent protons) that are separated by not more than three bonds in the covalent structure. In contrast, in NOESY spectra only intraresidual and sequential connectivities can, with certain limitations, be predicted from the sequence alone. Here, GARANT offers the possibility of using additional information taken from homologous proteins to improve the prediction of expected NOESY peaks; these particular applications are the theme of the present work. Lists describing the observed peaks are obtained by peak picking of the NMR spectra, for example, with the help of the spectral interpretation program XEASY (Bartels et al., 1995). While the observed peaks are characterized by their precise positions, the expected peaks are given by approximate positions (e.g., peaks involving a methyl group are found at the high-field end of the corresponding frequency axis) and an indication of their likelihood to be observed.

The resonance assignments obtained by GARANT are defined as the optimal mapping of the expected peaks onto the observed peaks (Fig. 1). The search for this optimal mapping is complicated by the fact that both peak lists may be incomplete and may contain errors. Thus, NOE cross-peaks arising due to the close proximity of pairs of hydrogen atoms in the 3D structure of the protein cannot be predicted from the sequence alone, or entries in the list of observed peaks may be missing, for example because of spectral overlap or due to incomplete mixing in experiments such as TOCSY. Spurious entries may result, for example, when reported observed peaks actually represent spectral noise.

In addition to the representation of resonance assignments as a mapping of expected peaks onto the observed peaks, GARANT includes a scheme for evaluating the quality of non-final assignments in order to drive the search for the optimal mapping. This optimization procedure is a combination of a genetic algorithm with a local optimization routine, as described in detail elsewhere (Bartels, 1995; Bartels et al., 1996).

Mandatory input for GARANT

To derive the list of expected peaks, GARANT utilizes the amino acid sequence of the protein, a general library describing the covalent structure of the 20 proteinogenic amino acid residues, and knowledge about individual types of NMR spectra (Fig. 1). Furthermore, a peak list is required that represents the experimental NMR data recorded for the protein under investigation. For each peak this list contains the type of spectrum in which it was observed, and its position in this spectrum. Additional input parameters describe the accuracy of the peak positions within each spectrum, and provide an estimate for maximal expected chemical shift differences between corresponding peaks in different spectra.

Supplementary input for GARANT derived from homologous proteins

For the applications described here, supplementary data derived from the 3D structures and/or the chemical shifts of homologous proteins are used (Fig. 1). The 3D structures of homologous proteins serve to improve the prediction of the expected peaks in NOESY spectra. From short proton–proton distances in these structures, expected NOE cross-peaks are derived and corresponding entries are added to the list of expected peaks for all NOESY spectra considered. Additional parameters, such as a threshold for short proton–proton distances and a lower limit on the number of input structures of homologous proteins that must contain the short distance, determine when to include an NOE between two particular hydrogen atoms into the list of expected peaks.

A different type of information from homology considerations affects the probability distributions for expected chemical shifts. In the basic operation of GARANT (Bartels et al., 1996), these are assumed to be normal distributions about a mean value with a specified standard deviation, which are taken from a library based on statistical data from unrelated proteins (Gross and Kalbitzer, 1988) and on random coil chemical shifts (Richarz and Wüthrich, 1978; Bundi and Wüthrich, 1979). When shifts for homologous proteins are known, the mean values of the afore-mentioned distributions are set equal to the corresponding 'homologous' chemical shifts, and much smaller, user-defined standard deviations are used. In the examples of assignments supported by chemical shifts in related proteins as presented in the Results section, the standard deviations were set to 0.05 ppm for carbon-bound hydrogen atoms and to 0.15 ppm for amide protons. These standard deviations are significantly smaller than the corresponding values from the general program library, where, for example, the standard deviation for H^{α} of glutamic acid is 0.42 ppm (for details, see Bartels, 1995; Bartels et al., 1996). In this way, knowledge of the shifts in homologous proteins results in a much more accurate description of the expected peak positions.

Preparation of the lists of observed peaks

Lists of observed peaks were generated with the following automated tools. Homonuclear COSY peaks were picked using a simple symmetry-based algorithm (Meier et al., 1987) implemented in the program XEASY (Eccles et al., 1991; Bartels et al., 1995), where peaks close to the diagonal or close to the water resonance in ω_1, as well as peaks having no symmetric counterpart relative to the diagonal are eliminated from the list. For TOCSY and NOESY spectra, we used another simple algorithm (Schäfer, 1992) implemented in the program XEASY, which identifies all local maxima in the spectra, and discards those entries that are either close to the diagonal, have very narrow line shapes, or belong to noise bands. The latter are identified based on the high number of local maxima present in a narrow frequency range. The lists from heteronuclear resolved 3D spectra were further processed by comparing them to the peaks in corresponding 2D [^{1}H,^{13}C]-COSY or 2D [^{1}H,^{15}N]-COSY spectra recorded with large frequency ranges, using the restriction that within some tolerances each peak of the 3D spectrum has to correspond to a peak in the corresponding heteronuclear COSY spectrum; we used band widths of 0.02 ppm in the direct ^{1}H dimension and 0.3 ppm in the heteronuclear dimension. This processing by comparison with 2D spectra is mainly used to define whether and how often a peak in the 3D spectrum is folded, and to identify artefacts in the list of 3D peak positions that do not correspond to any COSY peak. As a result of this processing step all real peak positions of the 3D spectra are set to their unfolded values, and whenever a peak

from a 3D spectrum corresponds to several 2D COSY peaks, multiple entries are included in the list picked from the 3D spectrum to ensure that all possible unfolded positions of this peak will be considered.

Results

Three systems are used here as an illustration of the extent to which sequence-specific resonance assignments can be obtained when using GARANT for serial studies of related proteins. Firstly, homonuclear 2D ^{1}H NMR spectra of Tendamistat(R19L), a 74-residue mutant protein containing extensive β-sheet structure, were automatically assigned using data that were derived either from the 3D structure of wild-type Tendamistat (Kline et al., 1988) or from its chemical shifts (Kline and Wüthrich, 1986) in addition to the NMR spectra and the standard program libraries (Fig. 1). Secondly, for the *Antp-(C39S,W56S)* homeodomain, an α-helical protein of 68 residues with significant unstructured polypeptide segments at both chain termini, data from homonuclear 2D ^{1}H NMR spectra were combined with information from the structure of a closely related homeodomain, *Antp-(C39S)* (Güntert et al., 1991). Thirdly, studies with the free form of human cyclophilin A, a protein with 165 amino acid residues, using supporting data from a cyclophilin A–cyclosporin A complex (Spitzfaden et al., 1994), are included to demonstrate the potential of GARANT for assignments of larger proteins with the use of heteronuclear 3D NMR spectra.

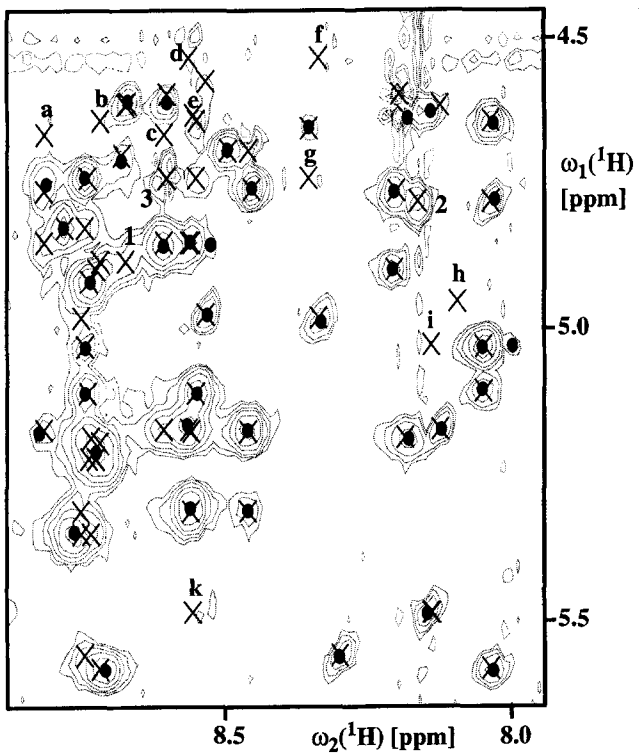

Fig. 2. A region of a 2D ^{1}H NOESY spectrum of Tendamistat(R19L) (O'Connell et al., 1994). Filled circles identify the positions of the peaks that were picked by the automatic peak picker used in XEASY (Bartels et al., 1995) and entered into the list of observed peaks. Crosses mark positions of entries in the list of expected peaks obtained using experimental chemical shifts in GARANT (Fig. 1). Lower-case letters indicate expected peaks that were not observed, and numbers identify positions of peaks that were missed by the peak picker; these cases are further discussed in the text.

Tendamistat(R19L)

The experimental input for the resonance assignment of Tendamistat(R19L) consisted of the automatically picked peak positions (see Methods) from 2D homonuclear ^{1}H COSY, TOCSY and NOESY spectra. The quality of the lists of observed peaks is illustrated with a region from the NOESY spectrum, where all the peak positions identified by the automatic peak picker are marked by filled circles (Fig. 2). Several peaks were missed in regions of overlap (e.g., peak 1), in noise bands (e.g., peak 2), or because of distorted line shapes (e.g., peak 3). The list of expected peaks obtained from the sequence was augmented by considering the three-dimensional NMR structure of wild-type Tendamistat as represented by an ensemble of 20 conformers (Kline et al., 1988; O'Connell et al., 1994). For all pairs of atoms closer than 3.0 Å to each other in all 20 wild-type conformers, an additional entry was generated with a probability of 1.0 of being observed. For all pairs of atoms closer than 4.0 Å to each other in five or more of the 20 wild-type conformers, an entry was added with a probability of 0.2 of being observed.

An overview of the extent of the automated resonance assignment obtained from the above input for Tendamistat(R19L) is shown in Fig. 3A. The backbone assignment is complete except for the flexibly disordered N-terminal tripeptide segment, and two amide protons and six α-protons in other regions of the sequence. Overall, of the total of 393 proton frequencies (number of protons in the protein, counting each methyl group once and excluding hydroxyl and carboxyl protons) that could in principle be observed in NMR experiments, 321 were correctly determined. Of the missing assignments, 18 resonances were assigned erroneously to a particular side-chain atom position in the correct sequence position. This may be explained by the fact that side-chain assignments are often dependent on a single COSY peak, and the complete lack of redundant experimental data then makes the assignment very sensitive to inaccuracies of the peak picker.

Since residue 19 has been mutated, its resonance assignment was of particular interest. The H^N resonance frequencies of Leu^{19} and Tyr^{20} are degenerate. This could in principle lead to difficulties in the sequential assignment because of overlap of the sequential connectivities with intraresidual peaks of Leu^{19}. Nonetheless, as indicated in Fig. 3A, the H^N and H^α frequencies were correctly assigned. The incomplete assignment for Leu^{19} (Fig.

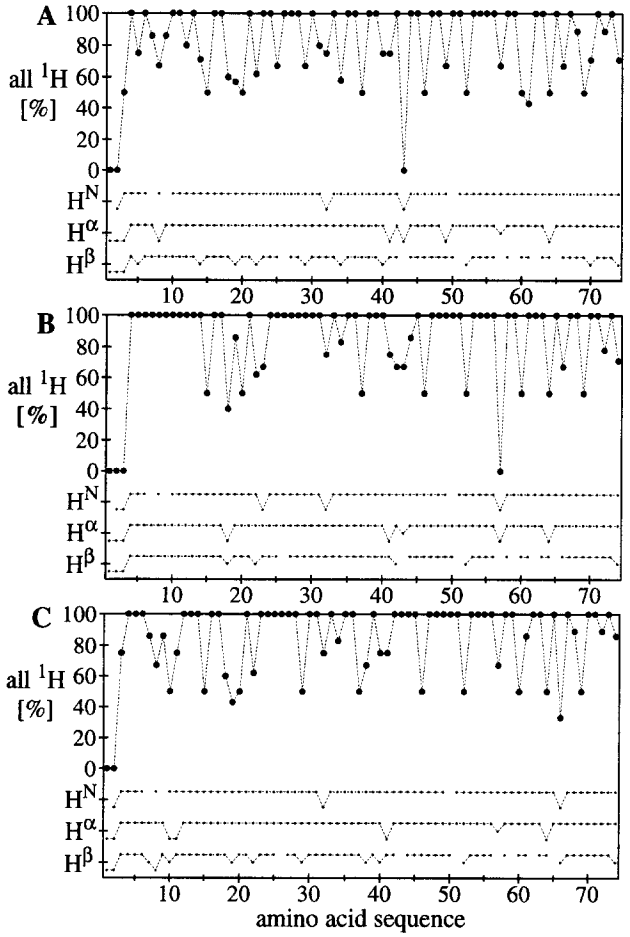

Fig. 3. Analysis of the automatically determined resonance assignment of Tendamistat(R19L) by plots against the amino acid sequence. In the upper part of each plot the large dots indicate the percentage of correctly determined proton resonance frequencies for the individual residues. In the lower part the positions of the small dots indicate whether the H^N, H^α and H^β frequencies were correctly assigned: a high position of the dot depicts a correct frequency; a low position depicts an incorrect frequency; and an intermediate position indicates that one of two frequencies is correct (e.g., αCH_2 of glycine or βCH_2); no dot means that the atom type is absent (H^N in proline, H^β in glycine). (A) Assignment using the NOEs derived from the 3D structure of wild-type Tendamistat; (B) Assignment using the chemical shifts of wild-type Tendamistat; (C) Assignment using both the 3D structure and the chemical shifts of the wild-type protein. Observed peaks were automatically picked from the 2D ^{1}H COSY, TOCSY and NOESY spectra (O'Connell et al., 1994). All spectra were recorded with a 4-mM solution of Tendamistat(R19L) in H_2O at 50 °C, pH 3.2, on a Bruker AM 600 MHz spectrometer.

3A) is due to chemical shift degeneracy of one H^β with both δCH_3 groups, a difficulty that was not resolved by GARANT.

To further illustrate some special features of GARANT we use Fig. 2, which shows a contour plot of the NOESY spectrum of Tendamistat(R19L), from which the observed NOE peaks (Fig. 1) were extracted. Peaks indicated by filled circles were picked by the automatic peak picker and are available to GARANT as observed peaks. The program can map entries from the list of expected peaks only to these observed peaks. Expected peaks cannot be shown directly in the spectrum of Fig. 2 since, as explained previously, their positions are known only approximately. However, at the end of a GARANT run, many expected peaks are mapped onto observed peaks, which provides the precise chemical shifts for the corresponding nuclei. With this additional information, unmapped entries in the list of expected peaks can be positioned in the spectrum, yielding the crosses in Fig. 2. Crosses coinciding with filled circles identify correct peak assignments, except for the peak at ($\omega_1 = 5.48$ ppm, $\omega_2 = 8.15$ ppm); this error will be discussed below. One peak at ($\omega_1 = 5.01$ ppm, $\omega_2 = 8.01$ ppm) was picked but no expected peak was mapped onto it; in this case the same result was obtained from manual assignment. Crosses that are not near filled circles represent either correct assignments to peaks missing in the list of observed peaks or assignment errors by GARANT, as discussed in the next section.

The following details provide more intimate insight into the operation of GARANT. Ten expected peaks, labelled with lower case letters in Fig. 2, were assigned to positions in the spectrum where no intensity is observed. Of these, a and c involve one α-proton of Gly^{26}; for the other α-proton, peaks are both expected and observed. The error arose because no stereochemical assignment was assumed, and therefore peaks for both α-protons were entered into the list of expected peaks by GARANT. Peak b corresponds to a distance that is short in wild-type Tendamistat but not in the mutant protein. The observed peaks corresponding to d and f are hidden under the water line, as was revealed by measurements at a different temperature. Peak h indicates a peak that was assigned a wrong shift for the α-proton of residue 64, with the correct peak being hidden under the water line. Peak i represents the expected sequential $d_{\alpha N}$ peak connecting the N-terminal two residues; no peak involving the α-proton of residue 1 could be observed in this NOESY spectrum. Peak g was assigned to one β-proton of Asp^{40}; the resonance of this proton is, however, degenerate with the other β-proton, which was correctly assigned by GARANT. Peaks e and k are incorrectly placed due to interchange of the amide proton frequencies of Thr^{32} and Thr^2 by GARANT. The correct H^N frequency of Thr^{32} is 8.15 ppm, and the expected peak k should be mapped onto the peak observed at ($\omega_1 = 5.48$ ppm, $\omega_2 = 8.15$ ppm), which was erroneously assigned to Thr^2. The corresponding cross-peak of Thr^2 is not observed in the spectrum due to its proximity to the water resonance frequency.

Figure 3B shows the resonance assignment resulting from GARANT when the chemical shifts of the wild-type protein were used as supplementary data. A total of 332 resonance frequencies were correctly assigned, which is nearly identical to the result of Fig. 3A and thus shows

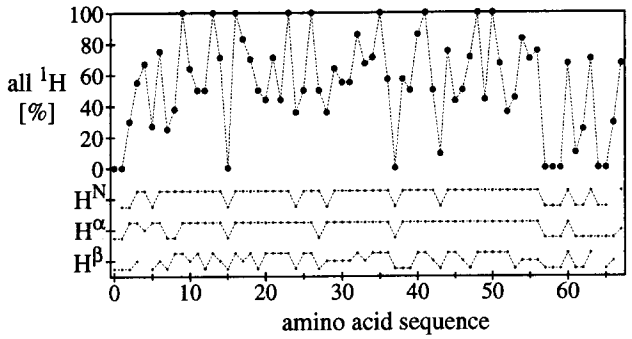

Fig. 4. Assignment with GARANT for the mutant homeodomain *Antp(C39S,W56S)* when using automated peak picking for obtaining the observed peaks, and the 3D NMR structure of *Antp(C39S)* as a source of supplementary input data. 2D ^{1}H TOCSY and NOESY spectra recorded in H_2O and a TOCSY spectrum recorded in D_2O were used as the primary data. All spectra were recorded at 10 °C, protein concentration 1.5 mM, pH 4.3. The TOCSY spectrum in D_2O was recorded at 500 MHz on a Bruker AMX spectrometer, the spectra in H_2O at 750 MHz on a Varian Unity-plus spectrometer. The same presentation as in Fig. 3A is used.

that not only the information from the 3D structure of homologous proteins but also the chemical shifts are efficiently used as supplementary input by GARANT. Combined use of the chemical shifts and the 3D structure of the homologous protein did not yield further significant improvement (Fig. 3C).

Antp(C39S,W56S) homeodomain

Figure 4 shows the result of GARANT calculations using automatically picked peaks from homonuclear 2D ^{1}H TOCSY and NOESY spectra recorded for the *Antp(C39S,W56S)* homeodomain (C. Bartels, D. Resendez-Perez, P. Güntert, D. Braun, W.J. Gehring and K. Wüthrich, unpublished results), and supplementary NOE predictions based on the 3D structure of the *Antp(C39S)* homeodomain (Güntert et al., 1991). The *Antp* homeodomain consists only of α-helices and connecting loops, which increases the overlap for cross-peaks involving α-protons when compared to β-proteins such as Tendamistat (Wüthrich, 1986). Extensive overlap in the region of the side-chain resonances occurs also because the protein contains 20 arginine and lysine residues. The spectra were recorded at 10 °C and exhibited unfavourable relaxation times for the TOCSY mixing. Another limitation for the present approach is that the N-terminal heptapeptide segment and the C-terminal octapeptide segment are flexibly disordered, so that no useful homology information for these terminal segments was available. This resulted in poor assignments of the chain termini, but did not significantly interfere with the assignment for the central, well-structured part, where nearly complete assignments for H^N and H^α were obtained (Fig. 4).

Cyclophilin A

Three-dimensional structures of several variants of human cyclophilin A in complexes with small ligands have been determined by X-ray crystallography and by NMR, and a crystal structure of the free protein is also available (for a review see Braun et al., 1995). The present illustration is taken from an NMR structure determination of free cyclophilin A. The resonance assignment of uncomplexed cyclophilin A is described as obtained from combined analysis of a 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum, a 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY spectrum and a 3D CBCA(CO)NHN spectrum (M. Ottiger, O. Zerbe and K. Wüthrich, unpublished results); these data were supplemented with NOE predictions based on the 3D NMR structure of cyclophilin A in a complex with cyclosporin A (Spitzfaden et al., 1994). Automatically picked experimental peaks were used, and supplementary expected NOESY peaks were generated from the conformers representing the NMR structure of the complex in the same way as for Tendamistat(R19L) and the *Antp(C39S,W56S)* homeodomain. An overview of the resulting resonance assignments is given in Fig. 5, where data for backbone amide ^{15}N resonances have also been in-

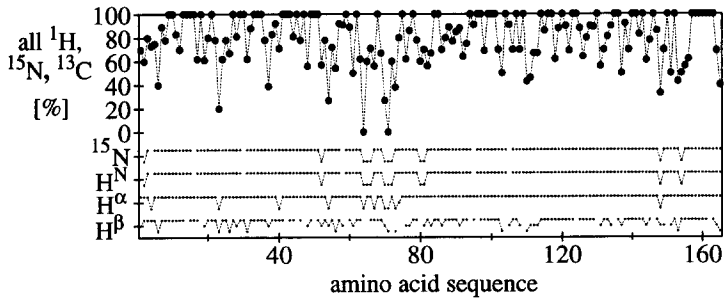

Fig. 5. Assignments with GARANT for human cyclophilin A when using NOE predictions derived from the structure of the cyclophilin A–cyclosporin A complex (Spitzfaden et al., 1994) as supplementary input. The primary input included a 3D ^{13}C-resolved [^{1}H,^{1}H]-NOESY spectrum and a 3D ^{15}N-resolved [^{1}H,^{1}H]-NOESY spectrum recorded in 4 days each on a Varian Unity-plus 750 MHz spectrometer, and a CBCA(CO)NHN spectrum recorded in 3 days on a 600 MHz Bruker AMX spectrometer (M. Ottiger, O. Zerbe and K. Wüthrich, unpublished results). All spectra were recorded at 26 °C, protein concentration 1.5 mM, pH (or pD) 6.5. The same presentation as in Fig. 3A is used, except that an additional row has been added to represent assignments of backbone ^{15}N nuclei, and that all ^{13}C and ^{15}N resonances are included in the percentage of correctly determined resonances given in the upper part of the figure.

cluded. The result for this larger protein is very similar to those obtained for the previous examples in Figs. 3A and 4. Nearly complete assignments were obtained for backbone ^{15}N, H^N and H^α, and a high percentage of assignments was obtained also for H^β (Fig. 5). Limitations arise mostly in flexible regions of the protein, where NOESY cross-peaks are scarce and peak intensities are reduced by line broadening. In particular, the assignment of Val^2, His^{70} and Glu^{81}, for which no cross-peaks are observable either in 3D ^{15}N-resolved [^1H,^1H]-NOESY or in the 3D CBCA(CO)NHN spectrum, was difficult. Residues 64 to 67 are located in a loop region and show only few NOEs. For Phe^{67} no peaks are observable in the CBCA(CO)-NHN experiment, and both Gly^{65} and Gly^{66} have one H^α resonance close to the water resonance line. All these localized difficulties could not be properly resolved by GARANT, as can be seen from Fig. 5.

Discussion and Conclusions

GARANT was designed for simultaneous use of the data from various recorded NMR spectra, thus avoiding the difficulty of treating, in subsequent steps, errors made at the outset. Early applications showed that, similar to other routines available, GARANT cannot correct for poor input data resulting from inadequate peak-picking routines (Bartels et al., 1996). For example, for Tendamistat(R19L) only 166 assignments were obtained when based on automatic peak picking of the NMR spectra without supplementary information. Fortunately, the basic concept readily allows the incorporation of additional information, such as homology data in the form of tertiary structure contacts or chemical shifts. The mapping algorithm of GARANT is then capable of substituting information missing in the automatically picked peak lists by data obtained from supplementary sources. An input of manually prepared lists of observed peaks yields similar results in the absence of homology data as the automatically picked peak list combined with supplementary information, like in the present paper (Bartels et al., 1996). This indicates that further improvement of automated resonance assignments may primarily depend on the development of peak-picking procedures with improved handling of spectral overlap and noise bands, possibly by adding line shape information into the lists of observed peaks to improve the reliability of automated peak picking, or eventually by the development of assignment programs with direct feedback loops to the original spectra.

Acknowledgements

We thank M. Ottiger for providing unpublished spectra of cyclophilin A, and R. Marani for the careful processing of the manuscript. Financial support was obtained from the Kommission zur Förderung der wissenschaftlichen Forschung (project 2223.1) and the Schweizerischer Nationalfonds (project 31.32033.91). The use of the C4 workstation cluster of the ETH Zürich is gratefully acknowledged.

References

Bartels, C. (1995) Ph.D. Thesis, ETH Zürich, Zürich, Switzerland.

Bartels, C., Xia, T.H., Billeter, M., Güntert, P. and Wüthrich, K. (1995) *J. Biomol. NMR*, **6**, 1–10.

Bartels, C., Güntert, P., Billeter, M. and Wüthrich, K. (1996) *J. Comput. Chem.*, manuscript submitted for publication.

Billeter, M. (1991) In *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance* (Eds, Hoch, J.C., Poulsen, F.M. and Redfield, C.), Plenum Press, New York, NY, pp. 279–290.

Braun, W., Kallen, W., Mikol, V., Walkinshaw, M.D. and Wüthrich, K. (1995) *FASEB J.*, **9**, 63–72.

Bundi, A. and Wüthrich, K. (1979) *Biopolymers*, **18**, 299–311.

Eccles, C., Güntert, P., Billeter, M. and Wüthrich, K. (1991) *J. Biomol. NMR*, **1**, 111–130.

Gross, K.H. and Kalbitzer, H.R. (1988) *J. Magn. Reson.*, **76**, 87–99.

Güntert, P., Qian, Y.Q., Otting, G., Müller, M., Gehring, W.J. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 531–540.

Kline, A.D. and Wüthrich, K. (1986) *J. Mol. Biol.*, **192**, 869–890.

Kline, A.D., Braun, W. and Wüthrich, K. (1988) *J. Mol. Biol.*, **204**, 675–724.

Meier, B.U., Mádi, Z.L. and Ernst, R.R. (1987) *J. Magn. Reson.*, **74**, 565–573.

O'Connell, J.F., Bender, R., Engels, J.W., Koller, K.P., Scharf, M. and Wüthrich, K. (1994) *Eur. J. Biochem.*, **220**, 763–770.

Richarz, R. and Wüthrich, K. (1978) *Biopolymers*, **17**, 2133–2141.

Schäfer, N. (1992) Diplomarbeit, ETH Zürich, Zürich, Switzerland.

Spitzfaden, C., Braun, W., Wider, G., Widmer, H. and Wüthrich, K. (1994) *J. Biomol. NMR*, **4**, 463–482.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York, NY.

Wüthrich, K. (1995a) *Acta Crystallogr.*, **D51**, 249–270.

Wüthrich, K. (1995b) *NMR in Structural Biology*, World Scientific, Singapore.