Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 175 (2005) 92–102
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Quantitative study of the effects of chemical shift tolerances and rates of SA cooling on structure calculation from automatically assigned NOE data

Michele Fossi [a], Hartmut Oschkinat [a,*], Michael Nilges [b], Linda J. Ball [a,c,*]

[a] *Forschungsinstitut für Molekulare Pharmakologie, Robert-Rössle-Str. 10, 13125 Berlin, Germany*
[b] *Institut Pasteur, 25-28 rue du Docteur Roux, 75015 Paris, France*
[c] *Structural Genomics Consortium, University of Oxford, Botnar Research Centre, OX3 7LD, UK*

Received 12 October 2004; revised 17 March 2005
Available online 4 May 2005

Abstract

The calculation of protein structures from nuclear magnetic resonance (NMR) data has been greatly facilitated by improvements in software for the automatic assignment of NOESY spectra. Nevertheless, for larger proteins, resonance overlap may lead to an overwhelming number of assignment options per peak. Although most software for automatic NOESY assignment can deal with a certain level of assignment ambiguity, structure calculations fail when this becomes too high. Reducing the number of assignment options per peak by reducing the chemical shift tolerances can lead to correct assignments being excluded, and thus also to incorrect structures. We have investigated, systematically, for three proteins of different size, the influence of the chemical shift tolerance limits (Δ) and of the number of simulated annealing (SA) cooling steps on the performance of the software ARIA. Large tolerance windows, and the correspondingly high levels of ambiguity, did not cause problems when appropriately slower cooling was used in our SA protocol. In cases where a high percentage of well-converged structures was not achieved, we demonstrate that it is more productive to calculate fewer structures whilst applying slow cooling, than to calculate many structures with fast cooling. In this way, high-quality structures were obtained even for proteins whose NMR spectra showed great degeneracy, and where there was much inconsistency in peak alignment between different samples. The method described herein opens the way to the automated structure determination of larger proteins from NMR data.

Keywords: Protein structure calculation; NMR; ARIA; Automated NOESY assignment; Simulated annealing

1. Introduction

Nuclear magnetic resonance (NMR) is the principal method for macromolecular structure determination in solution. The complete procedure consists of the consecutive steps of sample preparation, data acquisition, peak-picking, resonance assignment, collection of distance- and other structural restraints, and finally structure calculation and refinement [1].The traditionally slow step of assigning individual NOE peaks manually, provided until recently the major source of structural restraints. The introduction of automated NOE assignment software (reviewed recently [2]) has provided a viable alternative to this cumbersome and error-prone process. However, automated methods have limitations that require careful attention to avoid erroneous calculations.

NOESY spectra of even small proteins often contain significant signal overlap resulting in multiple assign-

* Corresponding authors. Fax: +49 30 94793 169 (H. Oschkinat), +44 1865 737231 (L.J. Ball).
E-mail addresses: oschkinat@fmp-berlin.de (H. Oschkinat), linda.ball@sgc.ox.ac.uk (L.J. Ball).

1090-7807/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2005.03.020

ment options for the majority of cross-peaks. The average number of NOE assignment options per peak (n_{av}) depends strongly on the chemical shift tolerance window (Δ) used. Assignment options which fall outside of this range are rejected. If the window is too large, too many options are considered and the simulated annealing (SA) process does not converge. Reducing Δ reduces the number of assignment options, but introduces the risk of excluding correct assignments, for example, in cases where peaks are shifted due to experimental effects (such as sample heating in a TOCSY experiment) or where there are small differences between samples used for the collection of different spectra [3]. Such exclusion of a correct assignment will result in an alternative, wrong assignment, which can distort the calculated structures.

An alternative way of reducing the average number of assignments per peak is to reject all peaks for which there are more than a user-defined number of assignment options. We call this the assignment cut-off or the 'maximum number of assignment options per peak' (n_{max}). Due to the intrinsic redundancy of structural information in NOESY spectra, the rejection of even a high number of peaks generally has little effect on the convergence of calculations. Recently, it has been shown that the omission of up to 50% of NOE cross-peaks has little influence on the quality of the calculated structures [4].

For smaller proteins (50–130 amino acids), a compromise can usually be found between degree of assignment ambiguity and number of rejected peaks. However, with increasing protein size, the increased signal overlap due to the sheer number of protons present, results in a dramatic increase in average number of assignments per peak. The same problem affects datasets of smaller proteins where alignment of the frequencies in the resonance assignment-list and the NOESY peak-list is particularly poor. In these cases, any attempt to reduce the number of assignment options by means of n_{max} results in the rejection of too many NOESY peaks, leaving insufficient distance restraints to define a high-resolution structure. For this reason, a strategy is needed for handling peaks with highly ambiguous assignments, rather than simply rejecting them.

The success of previous workers in obtaining high-quality structures for the melanoma inhibitory activity (MIA) protein by using a slow cooling phase during simulated annealing to handle highly ambiguous NOE peak-lists [5] showed that the rate of cooling is an important parameter which influences convergence in calculations from very ambiguous data. Allowing more time for thermal equilibration minimises the risk of the structure becoming trapped in local energy minima with insufficient kinetic energy to escape. Slow cooling therefore increases the probability of finding the global minimum of the target function.

We have expanded this idea to study systematically, the effects of slow-cooling on calculations carried out using automatically assigned NOE data from ARIA [6–10]. Peak-lists from datasets of varying quality and from proteins of different size were used, together with appropriately chosen chemical shift tolerances Δ. In this work, we aimed to quantify the benefits of slow-cooling on convergence when very ambiguous or very incomplete peak-lists are used. We have also identified optimal numbers of cooling steps for NMR datasets with different chemical shift tolerance requirements, whilst minimising use of CPU time.

Finally, we show that, whenever there are problems with calculation convergence, this approach is a far more productive way of using available CPU-time than previous suggestions to simply calculate more structures using a fast-cooling protocol [11].

2. Theory

Following the assignment of NOESY spectra, unambiguous and ambiguous NOE-derived distance restraints (plus additional restraints like dihedral angle restraints, hydrogen bonds, and disulphide bridges, if present) are used to calculate an ensemble of structures. In the automated NOE assignment program ARIA, these are refined in an iterative manner, each time using the latest round of structures to find new assignments for input into the next round of calculations.

Structure calculation algorithms aim to find the global minimum of a hybrid energy function E (target function), which includes a priori chemical knowledge of the system (the force field, defining bond lengths, bond angles, improper angles, and non-bonded interactions) and experimental data (the structural restraints):

$$\begin{aligned} E &= E_{\text{chem}} + E_{\text{exp}} = \sum_i w_i E_i \\ &= w_{\text{covalent}} E_{\text{covalent}} + w_{\text{angle}} E_{\text{angle}} + w_{\text{float}} E_{\text{float}} \\ &\quad + w_{\text{vdW}} E_{\text{vdW}} + w_{\text{unambig}} E_{\text{unambig}} + w_{\text{ambig}} E_{\text{ambig}} + \cdots \end{aligned} \quad (1)$$

The standard minimisation methods [12] tend to steer the system into local minima and frequently fail to reach the global minimum if the starting model is far away from the correct one. The introduction of simulated annealing (SA) optimisation techniques [11,13] have therefore had a large impact on structural biology. Monte Carlo-based and molecular dynamics-based SA have been used in structure prediction, molecular modelling, X-ray refinement and NMR structure determination for many years. The goal of these methods is to find the global minimum of the target function E by reducing the temperature of the system during a molecular dynamics simulation. The parameter 'temperature' in this context has no physical meaning, but is simply a

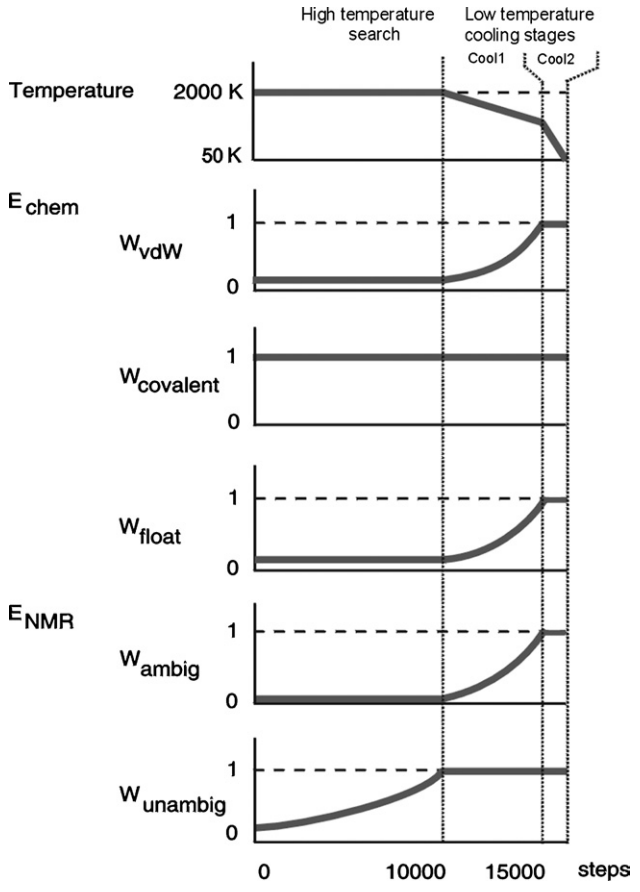

Fig. 1. Temperature and energy constants (with weighting, w_i) in a typical torsion angle dynamics SA protocol. w_{covalent} and w_{vdW} are the energy weightings for bonded and non-bonded interactions, respectively. w_{float} is the weighting of the angular energy term for chemical groups which undergo floating chirality assignment.

measure of the probability of the macromolecule to cross an energy barrier (i.e., its kinetic energy). Depending on temperature, a system is able to overcome local minima, thus allowing for the sampling of a larger conformational space. In ARIA, the user can choose between Cartesian (see [12] for a review) and Torsion Angle dynamics [14]. Both of these SA strategies have been optimised for ambiguous NOE-derived restraints. We use torsion angle dynamics for all structure calculations in this work as this generally produces an increased convergence radius and leads to better local geometries.

The standard protocol for torsion-angle molecular dynamics in ARIA used in the present work consists of an initial high-temperature conformational search at 2000 K, followed by two cooling stages in which the temperature is decreased linearly from 2000 to 1000 K and then again from 1000 to 50 K. During these three stages, the weights w_i in Eq. (1), defining the relative weights of the different energy terms of the target function, are varied (see Fig. 1).

3. Materials and methods

3.1. Proteins and datasets used in these calculations

The NMR test data used for our structure calculations came from four different proteins: the C-terminal PB1 domain of yeast CDC24p (Leitner et al., 2005, personal communication) (PB1 in the text), the N-terminal domain of the arginine repressor [15,16] (ArgR in the text), the HRDC domain of RecQ [17] (HRDC in the text), and the Ena/VASP homology 1 (EVH1) domain of human vasodilator stimulated phosphoprotein (VASP) [18] (EVH1 in the text). The important characteristics of these proteins and their spectra, relevant to this work, are summarised in Table 1.

3.2. Structure calculations

Automatic assignment of NOESY spectra and structure calculations were performed using the program ARIA v.1.2 on a Dual Athlon M1800+ cluster at the Pasteur Institute. Unless otherwise specified, the number of structures calculated was 20 for iterations 0–7 and 100 for the final iteration 8. The number of cooling steps used in each of the five cooling protocols tested herein was 9000, 18,000, 36,000, 72,000, and 144,000, respectively. In each protocol, the lengths of the first and second cooling stages were fixed in the ratio 5:4. The results

Table 1
The four different NMR datasets used for the calculations discussed

Name	PDB entries	Chain length	Secondary structure	Type of data	Number of peaks
PB1	1PQS	77	$4\beta + 2\alpha$	3D NOESY	^{13}C-edited 3D: 1909 ^{15}N-edited 3D: 766
ArgR	1AOY	78	$3\alpha + 2\beta$	2D NOESY	2D (H_2O): 1403 2D (D_2O): 1245
HRDC	1D8B	91	3α	3D NOESY	^{13}C-edited 3D: 2455 ^{15}N-edited 3D: 824
EVH1	1QC6	115	$7\beta + 1\alpha$	3D NOESY	^{13}C-edited 3D: 3506 ^{15}N-edited 3D: 2772

The NOESY peak-lists of ArgR, HRDC, and EVH1 were obtained by manual peak-picking, while those of PB1 were generated automatically with the internal peak-picking algorithm of Sparky v.3.1 (http://www.cgl.ucsf.edu/home/sparky/). Diagonal- and obvious noise peaks were removed manually. Unless specified in the text, no manual NOE assignments were included.

of the calculations were evaluated by computing a pairwise rmsd (precision) and an rmsd to a reference structure (accuracy) of the 20 lowest-energy structures. For references, we used the published X-ray structure for EVH1 [19] and the deposited NMR solution structures for ArgR [15,16], HRDC [17], and PB1 (Leitner et al., 2005, personal communication).

4. Results and discussion

4.1. Effect of SA cooling rate on the accuracy of structures obtained with increased chemical shift tolerance windows Δ

The resonance assignment-list and peak-lists from three proteins, ArgR, HRDC, and EVH1 were used as input data for all calculations described herein (see Table 1). Recently, we investigated the effects of different chemical shift tolerances Δ on the accuracies of structures obtained from these datasets using the program ARIA [3]. In these calculations, the ARIA default simulated annealing protocol was used, which consists of 9000 cooling steps (5000 for the first and 4000 for the second cooling phase, respectively). For each of these datasets, we performed two sets of structure calculations (one at each of two set values of the parameter n_{max}), in which only the sizes of the tolerance windows Δ were varied. Calculations were performed using 16 different sets of Δ. In each of these Δ-sets, Δ^{pro2} (direct proton dimension), Δ^{pro1} (indirect proton dimension), and Δ^{het1} (heteronuclear dimension) were fixed in the ratio 0.5:1:12.5. In the first set of calculations, n_{max} was fixed at 20 (the default setting of ARIA). In this case, all peaks with more than 20 assignment possibilities were discarded. The number of peaks discarded due to excess ambiguity increases dramatically as the tolerance window Δ is opened up. Hence, when n_{max} is limited, large windows Δ result in significant rejection of data. The second set of calculations was performed with $n_{max} = 200$ (i.e., effectively including all cross-peaks in the calculations, regardless of their degree of ambiguity). In these calculations, no peaks were rejected due to excess assignment options, allowing larger tolerance windows to be used without loss of restraints. The price paid for retaining these highly ambiguous data was a dramatic increase in n_{av} with increasing Δ, which often led to failure of convergence when using standard SA protocols.

The results of these previous structure calculations for ArgR, HRDC, and EVH1 are shown in Figs. 2A′, 3A′, and 4A′, respectively. Each of the test datasets displays a different optimal range for Δ, within which structures could be obtained with average backbone rmsds to previously published structures of less than 2 Å. This optimum Δ range depends on the protein size and on the cleanness and accuracy of the peak-list used as input data. With large values for Δ, the increased ambiguity led either to the rejection of large numbers of peaks in the case where n_{av} was limited to 20 (Figs. 2A′, 3A′, 4A′, top panel) or a dramatic increase in n_{av} in the case where this limit was relaxed (Figs. 2A′, 3A′, 4A′, bottom panel), as Δ was increased. In all cases, above a certain critical Δ, the number of rejected peaks in calculations with $n_{max} = 20$ and the number of assignment possibilities per peak in calculations with $n_{max} = 200$ eventually prevents convergence to accurate structures.

In this work, we investigated systematically the effect of slower cooling on datasets of different quality, to assess whether it may be possible to improve convergence at higher Δ. For each domain, three sets of Δ values were selected for which unsatisfactory structures had been obtained using the standard 9000 SA cooling steps. The selected sets are indicated by the vertical lines (labelled A–T) in Figs. 2A′, 3A′, and 4A′. The chemical shift tolerance for the indirect proton dimension (Δ^{pro1}) is given on the x-axis. The calculations were repeated four times with increasing numbers of cooling steps to a maximum of 144,000. In the case where n_{max} was limited to 20 (Figs. 2A′, 3A′, and 4A′; top panel), many peaks were thus rejected and peak-lists were very incomplete. Despite this, high quality, well-converged structures could still be obtained, provided that the number of cooling steps was sufficiently increased. The results of calculations carried out with $n_{max} = 20$, for Δ values labelled: A, B, C (ArgR); G, H, J (HRDC); and N, P, Q (EVH1) are summarised in the upper panel of Figs. 2B′, 3B′, and 4B′. It can be seen that simply doubling the default number of 9000 cooling steps can lead to dramatic improvements in rmsd to the reference structure, as in the cases of ArgR (2D dataset) and the EVH1 domain (3D dataset). The HRDC domain (3D dataset) showed comparable improvements with 36,000 steps.

In the case where n_{max} was effectively unlimited (i.e., set to 200; Figs. 2A′, 3A′, and 4A′; bottom panel), similarly dramatic improvements in the final structures were obtained (Figs. 2B′, 3B′, and 4B′; bottom panel). In this case, increasing the number of cooling steps enabled the ARIA software to cope with the larger number of assignment options per peak, rather than simply excluding large numbers of highly ambiguous peaks from the calculations. The results of calculations carried out with $n_{max} = 200$ for Δ values labelled: D, E, F (ArgR); K, L, M (HRDC); and R, S, T (EVH1) are summarised in the lower panel of Figs. 2B′, 3B′, and 4B′. For all 18 sets of calculations, regardless of whether n_{max} was restricted or not, the same input data which gave badly defined structures with 9000 cooling steps produced structures with greatly improved rmsd to the published, reference structures (<2 Å), when the number of cooling steps was increased.

A closer look at Figs. 2A′, 3A′, and 4A′ (bottom panel) allows quantification of these benefits. When

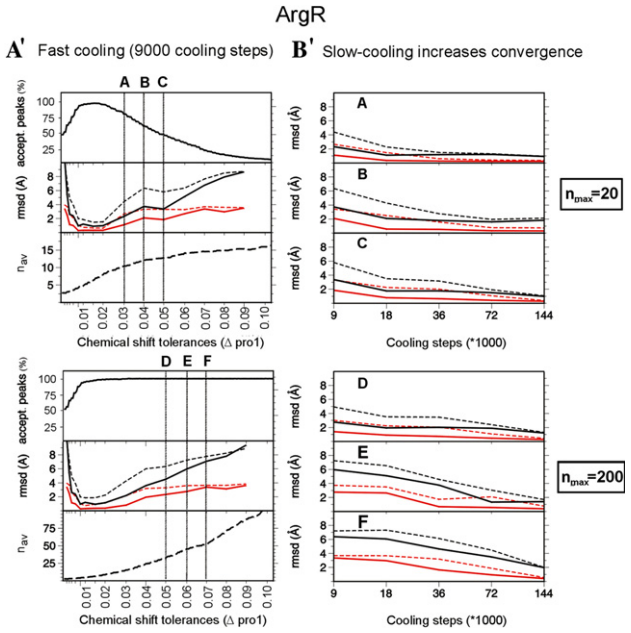

Fig. 2. Effects of chemical shift tolerances and slow-cooling on ArgR. (A′) Effects of chemical shift tolerances Δ on calculations with 9000 cooling steps. For each protein, two plots are presented corresponding to $n_{\max} = 20$ (upper plots) and $n_{\max} = 200$ (lower plots). Each plot contains three panels, representing three different parameters as a function of Δ. The top panels show the percentage of peaks accepted by ARIA in the calculations. The middle panels show (i) the accuracy (black lines) and (ii) the precision (red lines) of the calculated structures relative to the reference structures. These rmsds were calculated for the 20 lowest-energy structures (solid lines) and for all the 100 final structures (dashed lines). The lower panels show the average number of assignment options per peak (n_{av}). Numbers on the x-axis represent the values of Δ^{pro1}, The three capital letters on the top of each plot indicate calculations using larger chemical shift tolerances, for which poor quality structures were obtained. (B′) Effect of slow cooling on the 18 calculations indicated by capital letters in (A′). Black and red, solid and dashed lines are as in (A′). The top and bottom panels show calculations performed with $n_{\max} = 20$ and $n_{\max} = 200$ as in (A′). The plots show consistently that, from the same NOE input data and Δ-sets, greatly improved structures can be obtained if the number of cooling steps is increased. In most cases, 72,000 cooling steps were sufficient to obtain accurate and precise high-resolution structures. In addition, with slower SA cooling, the dashed and solid lines converge, i.e., the statistics for the entire ensemble become comparable to those for the best 20 structures. This reflects the higher percentage success rate of the algorithm when slower cooling is used.

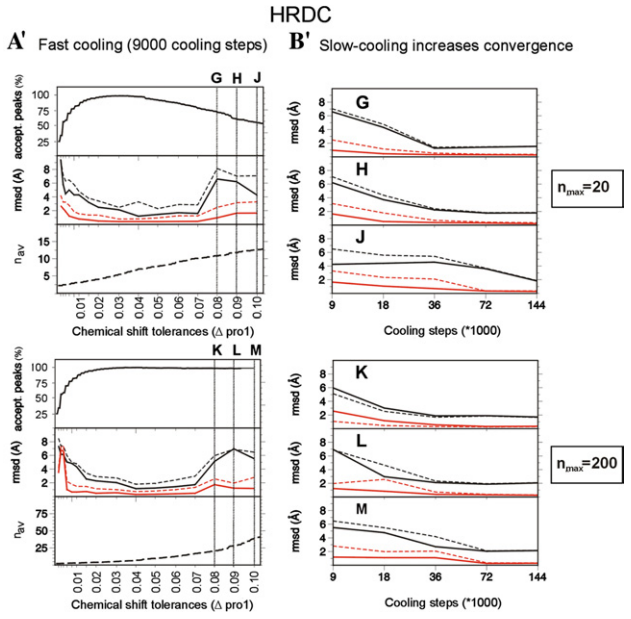

Fig. 3. Effects of chemical shift tolerances and slow-cooling on HRDC (see Fig. 2 for description of plots).

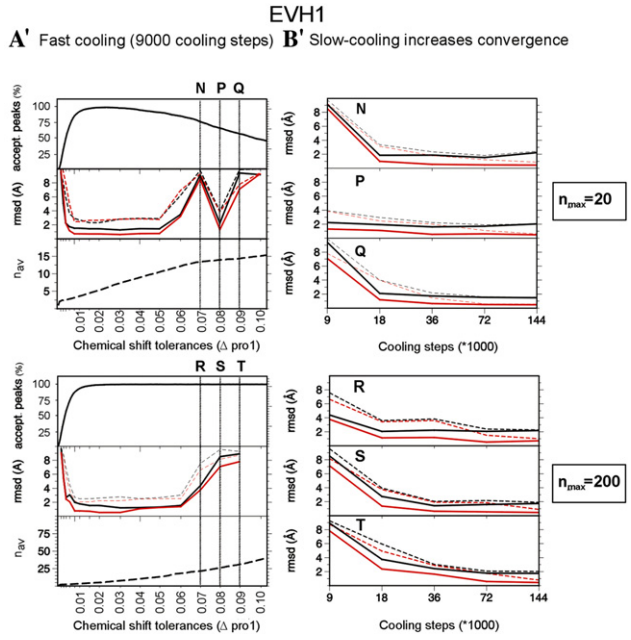

Fig. 4. Effects of chemical shift tolerances and slow-cooling on EVH1 (see Fig. 2 for description of plots).

$n_{\max} = 200$, no peak is rejected due to $n_{\max}$. When Δ was increased over a certain critical value, peak lists became too ambiguous and calculations produced incorrect results. With the default value of 9000 cooling steps, ARIA produced incorrect structures already with tolerances as high as $\Delta^{\text{pro1}} = 0.03$ for ArgR ($n_{\text{av}} = 15.4$), $\Delta^{\text{pro1}} = 0.08$ for HRDC ($n_{\text{av}} = 20.8$), and $\Delta^{\text{pro1}} = 0.07$ for EVH1 ($n_{\text{av}} = 22$). Nevertheless, with sufficiently slow cooling, ARIA handled much larger values of n_{av}, such as 52.01, 39.27, and 32.15, for ArgR, HRDC, and EVH1, respectively, without detrimental effects (see calculation F in Fig. 2B′, calculation M in Fig. 3B′ and calculation T in Fig. 4B′). Notable is the case of ArgR, where a simple increase in the number of cooling steps from 9000 to 144,000 enabled ARIA to handle 3.5 times more average ambiguities per peak. In calculations with limited ambiguity ($n_{\max} = 20$; Figs. 2A′, 3A′, and 4A′; top panel), incorrect structures were obtained when Δ was increased over a critical value because too many peaks were rejected due to the imposed limitation on $n_{\max}$. With 9000 cooling steps, poor results were obtained already with $\Delta^{\text{pro1}} = 0.03$ for ArgR ($n_{\text{av}} = 10.4$,

18.7% rejected peaks), $\Delta^{pro1} = 0.08$ for HRDC ($n_{av} = 10.8$, 27.4% rejected peaks), and with $\Delta^{pro1} = 0.07$ for EVH1 ($n_{av} = 12.7$, 14.7% rejected peaks). When a sufficient number of cooling steps was provided, good results were obtained even with $\Delta^{pro1} = 0.05$ for ArgR ($n_{av} = 12.7$, 48% rejected peaks), with $\Delta^{pro1} = 0.10$ for HRDC ($n_{av} = 12.5$, 45% rejected peaks) and $\Delta^{pro1} = 0.09$ for EVH1 ($n_{av} = 14.3$, 43% rejected peaks). It is important to note that these percentages refer to the removal of the most ambiguous peaks due to the imposed cut-off n_{max} and not to the random deletion of peaks from the peak-list. Taken together, the results show that slower cooling during the simulated annealing process increases the general robustness of the calculation procedure, allowing for the use of more ambiguous and less complete peak lists.

Slower cooling increases the probability of obtaining the correct fold, which corresponds theoretically to the global minimum of the target function. The solid lines in Figs. 2B′, 3B′, and 4B′ show the accuracy (black lines) and precision (red lines) of the 20 lowest-energy structures from the 100 calculated in each ensemble. For comparison, the dashed lines indicate the accuracy and precision of the total ensemble of 100 structures, respectively. The general trend is that the solid and dashed lines tend to become closer and sometimes even overlay, as the number of cooling steps increases. This shows that, in general, the larger the number of cooling steps, the higher the percentage of accurate calculated structures (with backbone rmsd values within 2 Å of the published reference structures). The boxed numbers in Fig. 5 indicate, for calculations E, L, and S (representing the datasets for ArgR, HRDC, and EVH1, respectively), the number of structures out of the total 100 calculated, which were accurate to within 2 Å backbone rmsd of the respective reference structures, for calculations performed using different numbers of cooling steps. All 100 structures in each ensemble calculated are superimposed, clearly illustrating that slowing down the cooling rate dramatically increases the robustness of the calculation procedure and hence the probability of obtaining the correct structure.

Fig. 6 shows, in histogram format, the effect of the SA cooling rate on the percentage of accurate structures obtained for each of the datasets studied. To exemplify this, we chose three calculations carried out with $n_{max} = 200$ (calculations F (ArgR), L (HRDC), and S (EVH1)), where the benefits of slow-cooling were particularly dramatic. Using the standard 9000 cooling steps, the structure calculations showed an almost zero probability of success in converging to within 2 Å of the reference structure (white bars). However, increasing the number of cooling steps resulted in a dramatic increase in the percentage of highly accurate structures. For all 3D datasets tested, almost all the final structures were accurate to within 2 Å of the reference structure when

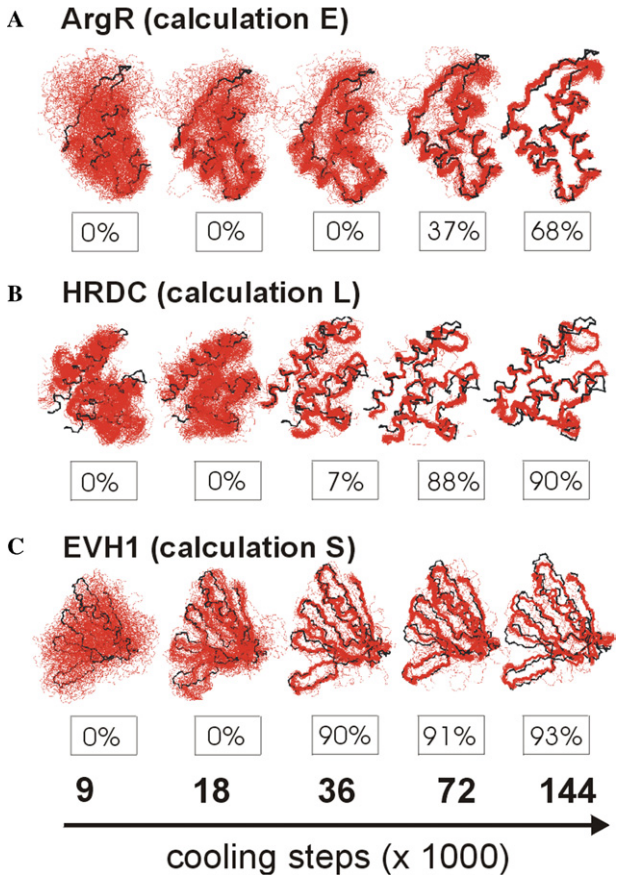

Fig. 5. Ensembles of 100 structures obtained from calculations E (ArgR), L (HRDC), and S (EVH1) with different numbers of cooling steps. The number in the box below each ensemble represents the percentage of structures with an accuracy within 2 Å of the reference structure, and shows that slower cooling dramatically increases the probability of each individual SA run leading to the correct fold.

using 36,000 cooling steps or more (mid-grey, dark grey, and black bars). Even with the extremely ambiguous calculation F ($n_{av} = 52.0$, highly degenerate 2D ArgR dataset), similar accuracies could still be achieved when the cooling was slowed to 144,000 steps.

It should be noted that structures calculated using very slow cooling show a tendency to appear over-restrained in regions where there are few or no experimental restraints, such as loops or terminal regions. In the absence of experimental restraints, slower equilibration leads to preferences between almost equivalent minima in the potential surface created by the a priori component of the target function alone (E_{chem} in Eq. (1)). Therefore, ^{15}N T_1, ^{15}N T_2, and heteronuclear NOE measurements are particularly important to identify genuinely flexible regions within the structure.

4.2. Slower cooling versus increased number of structure calculations

To illustrate the effects of slow cooling for a protein with very highly ambiguous NOE peak lists, we

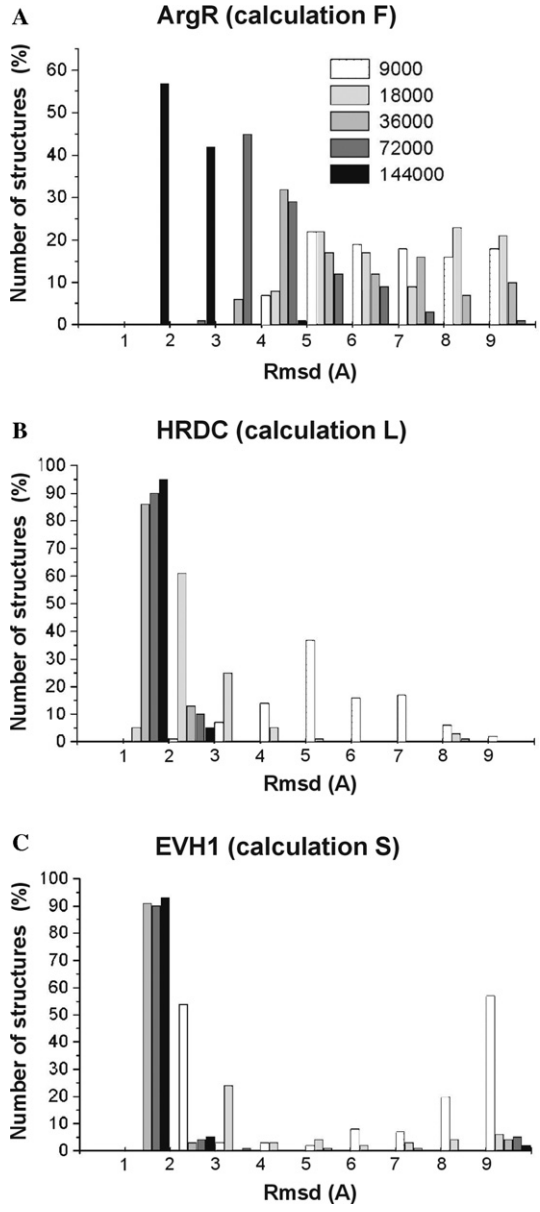

Fig. 6. The percentage of accurate structures obtained improves with the number of cooling steps used. The probability of success of calculations increases dramatically with slower cooling. When a sufficiently large number of cooling steps were used, almost all the final structures were accurate to within an rmsd of 2 Å to the reference structure (dark bars). This is in direct contrast to similar calculations which used 9000 cooling steps (white bars), where few, if any, of the final structures were close to the reference structure. (A) Calculation F (ArgR). (B) Calculation L (HRDC). (C) Calculation S (EVH1) (see Fig. 2).

analysed an NMR dataset from the 8.9 kDa C-terminal PB1 domain of yeast CDC24p. These data suffered from particularly poor alignment of the frequencies in the NOESY peak-lists and the resonance assignment-list: a problem that often arises when working with samples which are unstable over the period of time needed to acquire all of the experiments. Any attempt to obtain de novo structures of this domain using a protocol with 9000 cooling steps failed completely, even when a limited number of unambiguous, manually assigned, long-range NOEs were included as additional input data for the calculations. In cases like this, there is no choice but to use larger chemical shift tolerances Δ. However, in the case of the PB1 domain, this led to excessive ambiguity (>16 assignment possibilities per peak) and very cumbersome, non-convergent calculations. Simply increasing the number of structures calculated in each iteration, as has been previously suggested [11], did not result in satisfactory numbers of correct structures. We calculated 100, 200, and 500 structures per iteration (far in excess of the typical default value of 20 calculated structures) using the fast-cooling protocol. The backbone rmsds of best 20 structures of the final ensembles in each of these three cases were 4.10, 3.41, and 3.35 Å, respectively, and the accuracies with respect to the reference structures were described by backbone rmsds of 7.2, 7.0, and 7.0 Å, respectively. Regardless of the numbers of structures calculated, convergence was poor. This shows clearly that a better strategy is needed to calculate structures from such highly ambiguous data than simply increasing the number of calculated structures.

We therefore investigated whether using a slow-cooling SA protocol on this dataset could rescue the structure calculations in ARIA. The result was a successful de novo automatic structure determination of the PB1 domain when large tolerance values ($\Delta^{\text{prot}} = 0.07$) were combined with 72,000 cooling steps. Using this combination, an average of 16.4 possible assignments per peak in the first iteration occurred but did not prevent convergence. The results are represented in Fig. 7. For 9000 cooling steps, the convergence was extremely poor (Fig. 7A). The backbone rmsd of any individual calculated structure to the reference structure was never less than 7 Å. However, when 72,000 cooling steps were used with this same set of tolerance values, the 20 lowest-energy structures showed a precision of <1 Å and an accuracy of <3 Å to the reference structure (Fig. 7B). This result is encouraging as the calculations in Fig. 5B were performed in the absence of any additional structural information such as hydrogen bonds, dihedral angle restraints and manually assigned NOE-based distance restraints. Significantly, it shows that the correct fold can be obtained based entirely on automatically assigned NOE data even in the presence of inconsistencies within the datasets. In Fig. 7C, we show the results obtained when supplementing the dataset used in Fig. 7B with a set of experimental hydrogen bond restraints. This led to further improvements in precision and accuracy (rmsd to reference of 1.45 Å). The plots in Figs. 7B and C show clearly that correct results can be obtained only when sufficiently large values for Δ are chosen, as this is the only way to avoid exclusion of correct NOE assignments for this dataset. The wider tolerance

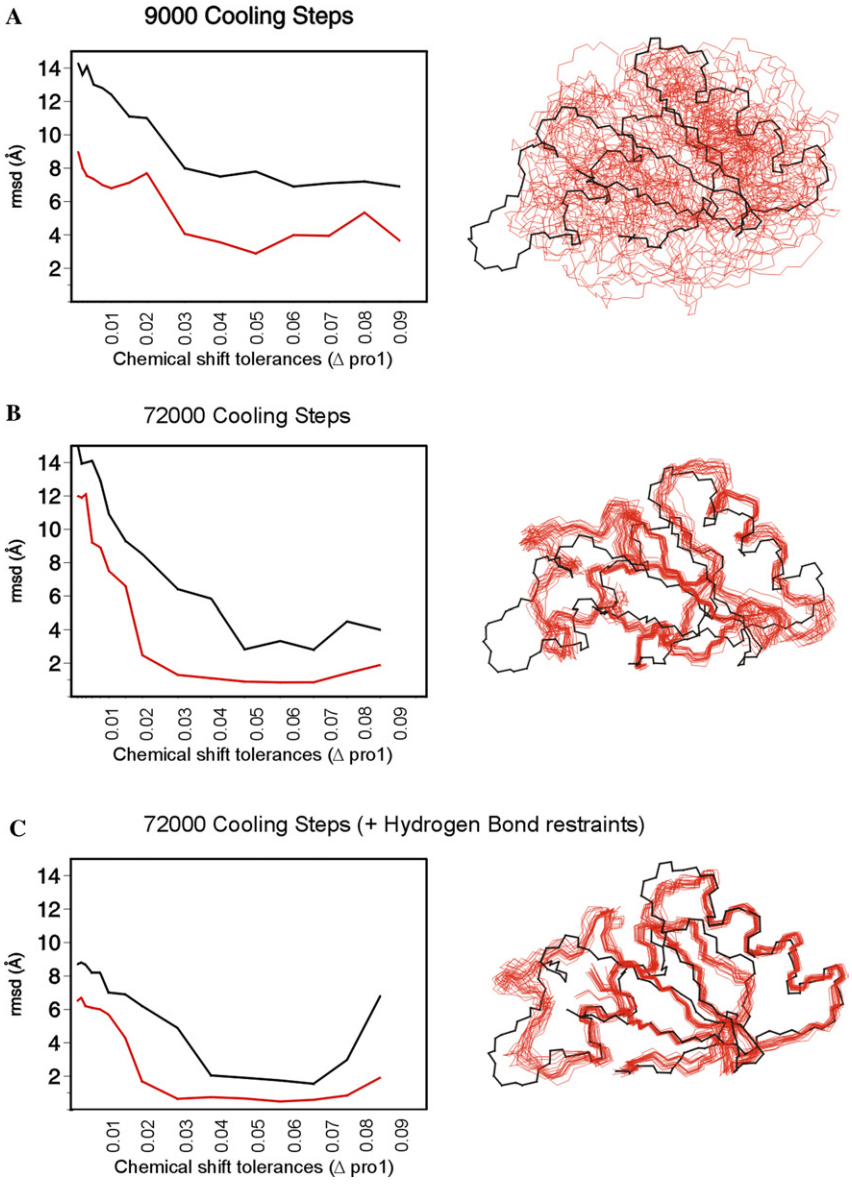

Fig. 7. De novo automatic structure determination of PB1 domain of CDC24p. Dependence of the accuracy (black line) and precision (red line) of the 20 lowest-energy structures on Δ. The structure ensembles calculated with $\Delta^{pro1} = 0.07$ are shown on the right. (A) Calculations with 9000 cooling steps. (B) Calculations with 72,000 cooling steps. When Δ is sufficiently large, the information contained in the dataset is very ambiguous but self-consistent. The slow-cooling yields convergent results and a correct overall fold. (C) Calculations with 72,000 cooling steps with hydrogen bond restraints added. The presence of these additional restraints restricts the conformational space to be searched and yielded very well converging and accurate results (accuracy = 1.55 for $\Delta^{pro1} = 0.07$).

windows led to a larger number of assignment possibilities per peak, i.e., a higher degree of ambiguity, which then requires much slower cooling to avoid the structures being trapped in local minima in the target function.

4.3. Time costs of the method

Use of a slow cooling phase during simulated annealing has the obvious disadvantage of increasing the CPU time needed for structure calculations. The time needed to calculate a single structure in each of the successive ARIA iterations, when using different numbers of cooling steps, is shown in Fig. 8A (with the example of calculation L of the HRDC domain). The initial iterations, in which the average number of assignment options per peak n_{av} are still large, are clearly the slowest, but the time needed to calculate a single structure decreases as the iterations progress. This is illustrated by the three examples shown in Fig. 8B. In addition, convergence to the correct overall fold is already achieved in the earlier iterations (in iteration 4 in the case of HRDC and EVH1 and in iteration 6 in the case of the highly ambiguous dataset of ArgR), as shown in Fig. 8C. These trends suggest that, although slow-cooling is of benefit

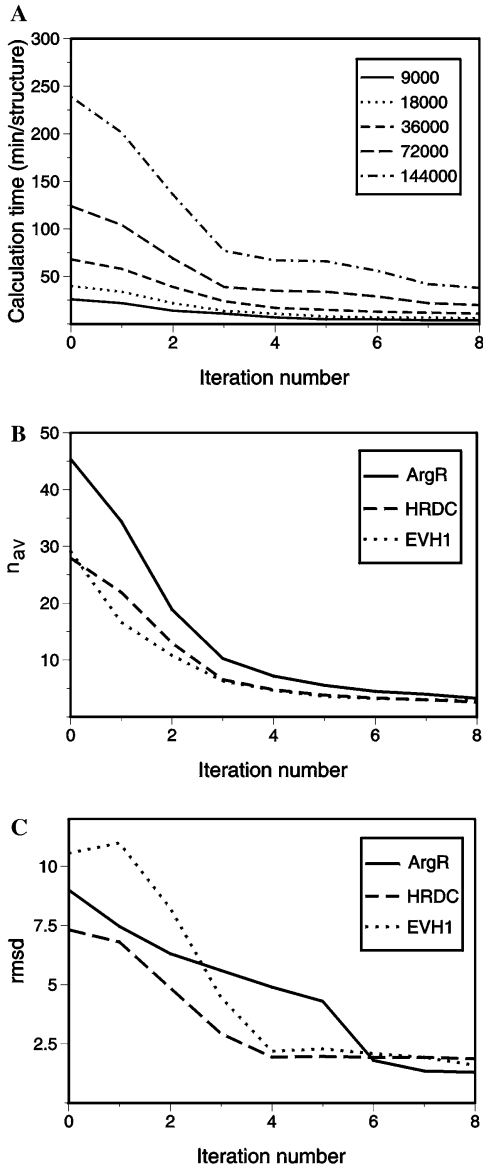

Fig. 8. Variations of structure calculation time, n_{av} and accuracy during the ARIA iterations. (A) Computational time required per structure at each iteration, using calculation L of the HRDC domain as an example. The differently dashed and dotted lines represent different numbers of cooling steps ranging from 9000 to 144,000. (B) Reduction in n_{av} with successive iterations (shown for three calculations: F (ArgR), L (HRDC), and S (EVH1), corresponding to F, L, and S in Fig. 2). (C) Reduction in backbone rmsd to the reference structure with successive iterations. Rmsds were calculated using the eight lowest-energy structures (out of 20) for iterations 0–7 and the 20 lowest-energy structures (out of 100) for the final iteration 8. The correct fold is already obtained by iteration 3 for calculation L, by iteration 4 for calculation S and by iteration 6 for calculation F, due to the higher initial value of n_{av} in the latter case.

in the early iterations, it is probably unnecessary to cool slowly in all ARIA iterations.

We therefore repeated the calculations shown in Fig. 8B using 72,000 cooling steps up until iteration 4 for HRDC and EVH1 domains, and until iteration 6 for ArgR (i.e., the points at which good folds were obtained) and then 9000 cooling steps for the remaining iterations, during which the structures were refined. The structures showed no notable difference in quality to those obtained when using 72,000 cooling steps for all iterations (Table 2). This modified strategy allowed substantial savings in CPU time, because fivefold more structures are calculated in the final iteration (iteration 8) than in all earlier iterations.

Fig. 9 shows the total time required to accomplish 9 iterations of calculation L of the HRDC domain using

Table 2
Comparison of accuracy and precision for three different calculation strategies

Protein		Method		
		A	B	C
ArgR (calculation F)	Accuracy	1.51	1.51	1.56
	Precision	0.61	0.61	0.62
HRDC (calculation L)	Accuracy	1.92	1.92	1.96
	Precision	0.67	0.67	0.73
EVH1 (calculation S)	Accuracy	1.62	1.62	1.62
	Precision	0.57	0.57	0.58

Once the correct fold is obtained, slow-cooling can be replaced by faster cooling. Using three test calculations corresponding to each NMR dataset, the quality of the structures in terms of accuracy and precision are compared for three different calculation strategies: (A) 72,000 cooling steps used in each iteration of ARIA; (B) 72,000 cooling steps used in iterations 1–7 and 9000 in iteration 8; (C) 72,000 cooling steps used in the early iterations until the calculation converges to a consistent global fold, then 9000 in the following iterations. Good folds were observed already in iteration 3 for HRDC, in iteration 4 for EVH1, and in iteration 6 for ArgR. The structures obtained with the faster strategies B and C were only marginally less accurate and precise than those obtained with the slower strategy A (see Fig. 9).

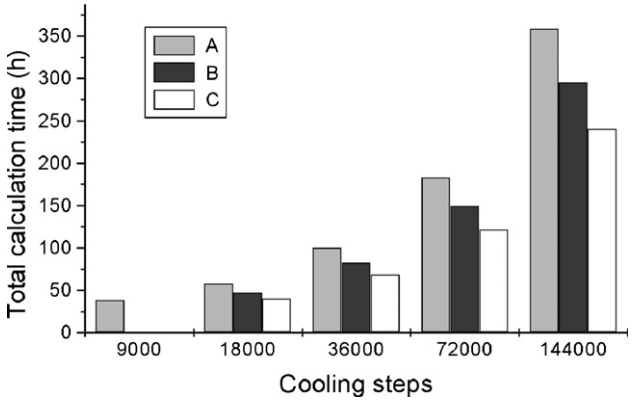

Fig. 9. Total computational time for different cooling schemes using calculation L of HRDC as an example (see Fig. 2): (A) constant slow cooling (in all iterations); (B) slow cooling for the first seven iterations, then reducing the number of cooling steps to 9000 in iteration 8; (C) slow cooling in early iterations until reasonable convergence was observed, then reducing the number of cooling steps to 9000 in iterations 5–8. The latter allows considerable time savings without noticeably affecting the quality of the results (results for the case of 72,000 cooling steps, for all NMR datasets used in these studies, are summarised in Table 2).

fixed-rate versus variable cooling schemes, starting from different initial numbers of cooling steps (9000 to 144,000). The general trend shows a decrease in the total computational time taken when the number of cooling steps is reduced to 9000 for: (i) the final iteration (iteration 8) only (black bars; B); and (ii) iterations 5–8 (white bars; C). The reduced-time protocols allow time savings of around 18 and 33%, respectively. For example, on four 1.8 GHz processors, the CPU time required with 72,000 cooling steps in all iterations was 37 h and the reduced-time version of this, using 9000 steps for iterations 5–8 was 31 h. Although the slow-cooling protocols are approximately 3–4 times slower than the fast (9 h) protocol which uses 9000 steps throughout, the former consistently yields high quality, accurate structures, while the latter frequently fails when applied to imperfect datasets. We therefore conclude that the investment of additional CPU time for slow-cooling is clearly time well spent. Hence, if computational time available is limited (rarely a problem with the recent generations of computers), protocols C or B are recommended. However, if time is not limiting, we would suggest employing the slower, most robust, protocol A.

5. Conclusions

We have shown elsewhere [3] that over-restricting the chemical shift tolerance windows, Δ, during automated NOE assignment can result in the exclusion of the correct assignment from the available options for a given peak and consequently induce distortions in the structures calculated. Here, we have shown that even highly ambiguous and incomplete NMR datasets can be handled by the automatic NOE assignment software ARIA if sufficient time for thermal equilibration is allowed during the SA cooling phase. For the ArgR, HDRC, and EVH1 datasets tested, using larger values of Δ produced accurate and precise structures if the number of cooling steps was increased (from the default 9000 to 72,000 steps). The combined use of large Δ values and slow-cooling dramatically increased the robustness of the SA protocol towards assignment ambiguity. Using this approach, we successfully calculated reliable structures from data which suffered from poor peak alignment, uncertainty in chemical shift positions, and extensive resonance overlap; problems that commonly plague the spectra of large proteins and complexes. The method described in this work could thus extend the use of NMR to in solving the structures of larger proteins, and enhance the role of NMR in structural genomics.

In addition, we show clearly that, whenever the ambiguity or the incompleteness of the peak-list prevents calculations from converging, a slow-cooling strategy is far more productive than previous suggestions simply to calculate many more structures in the computational time available [11]. When the input NMR data are largely complete and unambiguous, fast cooling may indeed produce accurate structures in a short time. However, with incomplete or highly ambiguous input data, the percentage success level of the calculations drops dramatically, and increases only when more generous chemical shift tolerances are coupled with slower SA cooling protocols. The higher demand on CPU time of slow-cooling protocols is well compensated by a dramatic rise of the probability of success in finding the global minimum in the overall potential during simulated annealing. As demonstrated by the PB1 domain, CPU time well spent on slow-cooling can save months of painstaking manual assignment.

Acknowledgments

M.F. acknowledges financial support from Protein Struktur Fabrik (Grant FKZ 01GG9812/4). Andrea Steuer is kindly acknowledged for her secretarial help.

References

[1] K. Wuthrich, NMR of Proteins and Nucleic Acids, Wiley, New York, 1986.
[2] P. Guntert, Automated NMR structure calculations, Progr. Nuc. Mag. Res. Spectrsc. 43 (2003) 105–125.
[3] M. Fossi, J. Linge, D. Labudde, D. Leitner, M. Nilges, H. Oschkinat, Influence of chemical shift tolerances on NMR structure calculations using ARIA protocols for assigning NOE data, J. Biomol. NMR 31 (2005) 21–34.
[4] J. Jee, P. Guntert, Influence of the completeness of chemical shift assignments on NMR structures obtained with automated NOE assignment, J. Struct. Funct. Genom. 4 (2003) 179–189.
[5] J.C. Lougheed, P.J. Domaille, T.M. Handel, Solution structure and dynamics of melanoma inhibitory activity protein, J. Biomol. NMR 22 (2002) 211–223.
[6] A. Kharrat, M.J. Macias, T.J. Gibson, M. Nilges, A. Pastore, Structure of the dsRNA binding domain of *E. coli* RNase III, EMBO J. 14 (1995) 3572–3584.
[7] J.P. Linge, S.I. O'Donoghue, M. Nilges, Automated assignment of ambiguous nuclear overhauser effects with ARIA, Nuc. Mag. Res. Biol. Macr. 339 (Pt. B) (2001) 71–90.
[8] J.P. Linge, M. Habeck, W. Rieping, M. Nilges, ARIA: automated NOE assignment and NMR structure calculation, Bioinformatics 19 (2003) 315–316.
[9] M. Nilges, M.J. Macias, S.I. ODonoghue, H. Oschkinat, Automated NOESY interpretation with ambiguous distance restraints: the refined NMR solution structure of the pleckstrin homology domain from beta-spectrin, J. Mol. Biol. 269 (1997) 408–422.
[10] M. Nilges, S.I. O'Donoghue, Ambiguous NOEs and automated NOE assignment, Prog. Nuc. Mag. Res. Spectrosc. 32 (1998) 107–139.
[11] A.T. Brunger, P.D. Adams, L.M. Rice, New applications of simulated annealing in X-ray crystallography and solution NMR, Structure. 5 (1997) 325–336.
[12] A.T. Brunger, M. Nilges, Computational challenges for macromolecular structure determination by X-ray crystallography and solution NMR-spectroscopy, Q. Rev. Biophys. 26 (1993) 49–125.

[13] S. Kirkpatrick, C.D. Gelatt, M.P. Vecchi, Optimization by simulated annealing, Science 220 (1983) 671–680.
[14] A. Jain, N. Vaidehi, G. Rodriguez, A fast recursive algorithm for molecular dynamics simulation, J. Comput. Phys. (1993) 258–268.
[15] J.P. Ni, V. Sakanyan, D. Charlier, N. Glansdorff, G.D. Van Duyne, Structure of the arginine repressor from *Bacillus stearothermophilus*, Nat. Struct. Biol. 6 (1999) 427–432.
[16] M. Sunnerhagen, M. Nilges, G. Otting, J. Carey, Solution structure of the DNA-binding domain and model for the complex of multifunctional hexameric arginine repressor with DNA, Nat. Struct. Biol. 4 (1997) 819–826.
[17] Z. Liu, M.J. Macias, M.J. Bottomley, G. Stier, J.P. Linge, M. Nilges, P. Bork, M. Sattler, The three-dimensional structure of the HRDC domain and implications for the Werner and Bloom syndrome proteins, Struct. Fold. Des. 7 (1999) 1557–1566.
[18] L.J. Ball, R. Kuhne, B. Hoffmann, A. Hafner, P. Schmieder, R. Volkmer-Engert, M. Hof, M. Wahl, J. Schneider-Mergener, U. Walter, H. Oschkinat, T. Jarchau, EMBO J. 19 (2000) 4903–4914.
[19] K.E. Prehoda, D.J. Lee, W.A. Lim, Structure of the enabled/VASP homology 1 domain-peptide complex: a key component in the spatial control of actin assembly, Cell 97 (1999) 471–480.