Journal of Biomolecular NMR, 9 (1997) 245–258
ESCOM

J-Bio NMR 422

Floating stereospecific assignment revisited: Application to an 18 kDa protein and comparison with J-coupling data

Rutger H.A. Folmer[a], Cornelis W. Hilbers[a,*], Ruud N.H. Konings[a] and Michael Nilges[b]

[a]*Nijmegen SON Research Center, Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands*
[b]*European Molecular Biology Laboratory, Meyerhofstrasse 1, D-6900 Heidelberg, Germany*

Received 26 August 1996
Accepted 12 December 1996

Keywords: Structure calculation; Floating chirality; Stereospecific assignment; Simulated annealing; ssDNA binding protein

Summary

We report a floating chirality procedure to treat nonstereospecifically assigned methylene or isopropyl groups in the calculation of protein structures from NMR data using restrained molecular dynamics and simulated annealing. The protocol makes use of two strategies to induce the proper conformation of the prochiral centres: explicit atom 'swapping' following an evaluation of the NOE energy term, and atom 'floating' by reducing the angle and improper force constants that enforce a defined chirality at the prochiral centre. The individual contributions of both approaches have been investigated. In addition, the effects of accuracy and precision of the interproton distance restraints were studied. The model system employed is the 18 kDa single-stranded DNA binding protein encoded by *Pseudomonas* bacteriophage Pf3. Floating chirality was applied to all methylene and isopropyl groups that give rise to non-degenerate NMR signals, and the results for 34 of these groups were compared to J-coupling data. We conclude that floating stereospecific assignment is a reliable tool in protein structure calculation. Its use is beneficial because it allows the distance restraints to be extracted directly from the measured peak volumes without the need for averaging or adding pseudoatom corrections. As a result, the calculated structures are of a quality almost comparable to that obtained with stereospecific assignments. As floating chirality furthermore is the only approach treating prochiral centres that ensures a consistent assignment of the two proton frequencies in a single structure, it seems to be preferable over using pseudoatoms or (R^{-6}) averaging.

Introduction

Soon after the first NMR structures were reported more than a decade ago, it was recognized that the availability of stereospecific assignments for methylene protons and isopropyl groups improves both the accuracy and precision of a calculated ensemble (Driscoll et al., 1989; Güntert et al., 1989; Havel, 1991). If these assignments are missing usually pseudoatoms are introduced, replacing the methylene or methyl protons (Wüthrich et al., 1983). Consequently, the corresponding experimental distance constraints must be widened to correct for the position of the pseudoatom relative to those of the protons for which the NOE has been measured. These effects are significant because the pseudoatom corrections that are generally used (1 Å for methylene, 2.4 Å for isopropyl) are based on worst-case geometries. Recent improvements in the pseudoatom concept allow somewhat smaller corrections (Güntert et al., 1991; Fletcher et al., 1996), but a loss of information is inevitable when pseudoatoms are used to compensate for the lack of stereospecific assignments.

Before the widespread use of multidimensional heteronuclear NMR, stereo assignments were generally obtained by a careful analysis of intraresidue and sequential NOE patterns and $^3J_{\alpha\beta}$-couplings (Zuiderweg et al., 1985; Hyberts et al., 1987). Automated procedures have been introduced that analyse these NOEs and coupling constants by grid searches (Güntert et al., 1989) or searches in X-ray structure databases (Nilges et al., 1990). In contrast, the program GLOMSA analyses calculated structures for

*To whom correspondence should be addressed.

consistent positions of prochiral groups to obtain assignments for a further refinement of the structures (Güntert et al., 1991). The obvious aim of these procedures is to gather as many stereospecific assignments as possible so that in a final structure calculation run the number of (inefficient) pseudoatoms can be reduced. An alternative approach to achieve the same goal has been suggested by Blaney (Weber et al., 1988), which is referred to as the floating chirality method. Instead of using pseudoatoms to constrain experimental distances, NOEs are measured for both individual resonances of a methylene or isopropyl group, which are arbitrarily assigned, i.e. simply denoted $H^{\beta 2}$ and $H^{\beta 3}$ in the case of a β-methylene group. During the structure calculation (distance geometry (DG) or restrained molecular dynamics (RMD)) stereo-related atoms or methyls are then allowed to float between the *pro-R* and *pro-S* configurations. Of course, all energy terms that enforce a defined chirality at the prochiral centre have to be removed (Weber et al., 1988). In the case of a simulated annealing refinement also the bond-angle energy constants involving the two protons or methyl groups are reduced (Holak et al., 1989). In this way, the protons or methyl groups can choose (DG) or move to find (RMD) the energetically most favourable conformation. The advantage of such an approach is that no correction factors have to be introduced and therefore no information content of the NOEs is lost.

Nevertheless, floating chirality strategies have never gained much popularity in NMR structure calculations. In fact, critical papers have appeared demonstrating that wrong assignments may easily ensue which, if not recognized, produce misleading results (Beckman et al., 1993). Moreover, if wrong assignments obtained from a previous structure refinement are really used in subsequent cycles, they may result in a strong 'bias' away from the correct structure (Havel, 1991). In his study, Havel used simulated NMR data, obtained from the crystal structure of BPTI, to study the accuracy of floating chirality as a method to make prochiral assignments and to investigate the effect on the precision of the calculated structures. Beckman et al. (1993), on the other hand, used real NOESY data from oxidized horse cytochrome c to examine the floating chirality method, but in their study the true stereo assignments (e.g. from J-coupling experiments) are not available. Both these studies focus to a large degree on floating chirality as a method to obtain stereospecific NMR assignments as such. In particular, Beckman et al. describe quite an extensive mathematical analysis to statistically validate the obtained assignments.

Here, we use the floating chirality method in a simulated annealing protocol to calculate the structure of a mutant ($Phe^{36} \rightarrow His$) of the 18 kDa dimeric single-stranded DNA binding protein (ssDBP) encoded by *Pseudomonas* bacteriophage Pf3 (Folmer et al., 1994,1995b). We will present a description of the annealing protocol used, which contains several modifications compared to previously reported ideas (Nilges et al., 1988,1991). J-coupling experiments have revealed stereo assignments for 34 prochiral centres in the 78 amino acid protein, which thus can be used to verify the results obtained from the automatic assignment. We have tested three protocols to examine the effect of 'atom swapping' strategies (Williamson and Madison, 1990) versus truly 'floating' prochiral centres. Furthermore, using different sets of NOE restraints we studied in what respect the success of floating chirality depends on the NOE input, notably the tightness of upper and lower bounds. Nowadays, with the availability of isotopic labelling and heteronuclear NMR techniques, in many cases stereo assignments can be obtained directly from the NMR experiments. Clearly, this would be the preferred approach, allowing the assignments to be made prior to and independent of the structure calculation process. Nevertheless, isotope labelling is not always possible and especially γ- and δ-methylene prochiral centres are difficult to assign in larger proteins using present-day NMR techniques. However, the purpose of this paper is not to demonstrate that floating chirality is a computational equal of the often elegant J-coupling experiments. Instead, we want to show that it is a useful, reliable and easy-to-use *tool* in protein structure determination, its primary goal being to improve convergence and to increase accuracy and precision, rather than to yield a table of assigned chemical shifts.

Materials and Methods

NMR measurements

The ^{15}N-labelled and $^{13}C/^{15}N$ doubly labelled F36H Pf3 ssDBPs were isolated as described previously (Folmer et al., 1994,1995b). The concentration of the NMR samples was 2 mM for the ^{15}N-labelled protein and about 1.4 mM for the doubly labelled protein.

NMR experiments were performed at 500 and 600 MHz on a Varian Unity+ and Bruker AMX spectrometer, respectively, and were carried out at 300 K. Unless otherwise stated, States-TPPI was used for signal accumulation in the indirectly detected dimensions and low-power GARP (Shaka et al., 1985) was used to decouple ^{15}N or ^{13}C during acquisition. In experiments performed in D_2O solution the solvent resonance was suppressed by presaturation, which was not necessary in the experiments in H_2O as these involved coherence selection using gradients.

Quantitative distance constraints were obtained from a 3D gradient-enhanced ^{15}N NOESY-HSQC spectrum, recorded at 600 MHz with a mixing time of 40 ms and 12 transients per increment. Maximum evolution times were $20.2(t_1) \times 18.0(t_2,^{15}N) \times 67.6(t_3)$ ms. In addition, a 3D ^{13}C NOESY-HMQC (Ikura et al., 1990) was acquired at 600 MHz with a mixing time of 40 ms, 24 transients per increment and evolution times of $24.3(t_1) \times 7.1(t_2,^{13}C) \times$

69.6(t_3) ms. Carbon decoupling during t_1 was achieved by applying a 500 μs hyperbolic secant ($\mu = 5$, $\beta = 200$ Hz; see Silver et al. (1984)) with an rf field strength of 13 kHz and the carrier at 77 ppm (Folmer et al., 1995a).

The stereospecific assignments reported in this paper are based on the following experiments, all performed at 500 MHz: 3D CT-HNHB (Archer et al., 1991), 3D HACAHB-COSY (Grzesiek et al., 1995), 2D {^{15}N} spin-echo difference HSQC (Vuister et al., 1993) and 2D {^{13}CO} spin-echo difference HSQC (Grzesiek et al., 1993). The HNHB experiment was recorded with 32 scans per increment and evolution times of 20.0(t_1,^{15}N) × 19.1(t_2) × 77.8(t_3) ms. The original sequence of Archer et al. was modified to enable gradient coherence selection (Kay et al., 1992) and sensitivity enhancement (Cavanagh et al., 1991; Palmer et al., 1991). The HACAHB-COSY experiment was acquired on a 0.9 mM sample with 40 transients per increment and evolution times of 21.3(t_1,^{13}C) × 11.6(t_2) × 56.9(t_3) ms. We took the exact sequence of Grzesiek et al., using a separate channel for GARP decoupling of the carbonyl resonances (0.3 kHz field). The {^{13}CO} spin-echo difference HSQC was recorded with 64 transients per individual increment and evolution times of 50.0(t_1,^{13}C) × 77.6(t_2,^{1}H) ms, whereas the {^{15}N} analogue was recorded with 96 transients.

Annealing protocol

Structures were calculated using simulated annealing starting from conformations with random backbone torsion angles. The basic ideas have been published previously (Nilges et al., 1988,1991). Several modifications were introduced to make the protocol more efficient and to allow for the use of floating diastereospecific assignment. All calculations were performed with the program X-PLOR, v. 3.1 (Brünger, 1992), with an extension for floating chirality and with a modified version of the 'parallhdg' geometric force field which now more closely matches the geometric parameters reported by Engh and Huber (1991). The extension (commands for CPU-efficient atom swapping) and the force field will be available in the next version of X-PLOR.

The entire protocol is summarized in Table 1. In principle, the stages denoted '0' and 'I' can be merged into one high-temperature phase, but here stage 0 served to calculate a set of starting structures. These were accepted subject to the criterion that the backbone root-mean-square (rms) difference to a (the) correct structure, which we knew from a previous refinement (Folmer et al., 1995b), was within 6 Å.

The annealing protocol consists of four stages: a high-temperature search phase and three cooling phases. During the search phase a reduced representation for nonbonded interactions was used as described previously (Folmer et al., 1995b) to increase the convergence rate. The temperature is reduced from 2000 to 1000 K in the first cooling phase, and all weights on the different energy terms are brought to their final values (see Table 1). The second cooling phase comprises simple cooling from 1000 to 100 K. In principle, this yields well-converged, low-energy structures, but these may be further relaxed by a

TABLE 1
DETAILS OF THE SIMULATED ANNEALING PROTOCOLS

	Stage				
	0	I	II	III	IV
Temperature[a]	2000	2000	2000 → 1000	1000 → 100	1000 → 100
Number of steps[b]	5000	3000	3500	2000	3000
Parameters and force constants[c]					
K_{NOE}[d] (kcal mol^{-1} Å^{-2})	3/6/12/25/50	12/25/50	50	50	50
K_{vdW}[e] (kcal mol^{-1} Å^{-4})	$0.002 K_{NOE}$	$0.002 K_{NOE}$	0.003 → 4.0	4.0	0.5 → 4.0
$K_{dihedral}$ (kcal mol^{-1} rad^{-2})	10	10	20 → 200	200	200
K_{float}[f] (kcal mol^{-1} rad^{-2})	25	25	25 → 500	500	500
repel[e]	1.2	1.2	0.9 → 0.78	0.78	0.78
cutnb[g]	12	12	4.5	4.5	4.5

[a] The temperature was controlled by the heat bath coupling method of Berendsen et al. (1984). The friction coefficient was set to 10 ps^{-1}. The bath was cooled stepwise using 50 K decrements.

[b] The time step was 5 fs throughout the whole protocol. Masses were uniformly set to 100 amu.

[c] Force constants for bonds, angles and impropers were uniformly set to 1000 kcal mol^{-1} Å^{-2} and 500 kcal mol^{-1} rad^{-2}. The improper constants defining the chirality of the prochiral groups subjected to floating stereospecific assignment were obviously set to zero.

[d] We used a harmonic 'flat-bottom' potential (or square-well with harmonic walls) with a linear behaviour for upper bound violations larger than 1 Å (Nilges et al., 1988). The slope of the asymptote was set at 2. K_{NOE} was increased as indicated every 1000 steps during stages 0 and I.

[e] The repulsive nonbonded potential has the form $E_{vdW} = K_{vdW}\{\max[0,(\text{repel}^2 R_{min}^2 - R^2)]\}^2$, where R_{min} is the sum of the two van der Waals radii, repel is a scaling factor and K_{vdW} is the energy constant (Konnert and Hendrickson, 1980). During stages 0 and I, a reduced atom representation was used for the nonbonded interaction, evaluating only two atoms per amino acid (Folmer et al., 1995b). In stage II, the full nonbonded representation was introduced.

[f] This force constant refers to the three-atom bond angles of the prochiral centres subjected to floating assignment.

[g] This parameter specifies the interaction cutoff for the nonbonded list generation.

subsequent mild annealing phase consisting of cooling from 1000 to 100 K, while increasing the force constant of the nonbonded interactions from 1/8 of its standard value to the final 4 kcal mol^{-1} $Å^{-1}$. In our experience this does not significantly change the overall structure, but local reorientations allow the various energy terms to drop by about 10%. The protocol ends with a 500-step conjugate gradient minimization.

For NOE-derived distances we used a harmonic flat-bottom potential with a linear behaviour for upper bound violations larger than 1 Å (Nilges et al., 1988). This prevents large forces in the beginning of the calculation when restraints in the random starting conformations are violated by a large amount, and facilitates convergence especially for ambiguous distance restraints (Nilges, 1993). Large energy barriers between the random initial conformations and the correctly folded structures are reduced by using soft atoms (Nilges et al., 1988), here in the form of a quartic repulsive potential similar to that used in various distance geometry programs.

Floating diastereospecific assignment

In the most elegant approach, floating chirality is actually performed on the restraints themselves rather than on the involved atoms (Habazettl et al., 1990). Unfortunately, this approach is not easily combined with the use of ambiguous NOEs. In homodimeric proteins, NOE contacts of residues close to the symmetry axis will have both intra- and intermonomer contributions, and these can only be treated correctly with an ambiguous NOE approach (Nilges, 1993). Here, we have therefore chosen the more standard method, where the restraints are unchanged. We used floating chirality for all prochiral centres in Pf3 ssDBP for which the NMR frequencies are resolved. These include all valine and leucine isopropyl groups and a large fraction of the methylene groups. The method consists of removing all energy terms that enforce a defined chirality at the prochiral centre (improper torsion angles) (Weber et al., 1988), and reducing all relevant bond-angle energy constants (Holak et al., 1989). This allows the protons or methyl groups to move and find the energetically most favourable conformation. In our experience, the protocol as described by Holak et al. (1989) converges well only if rather tight limits on the distances are used, in particular on lower bounds, otherwise the protons very often do not 'flip'. In order to be able to use floating chirality also with more qualitative upper bounds, we have introduced systematic swapping of the prochiral groups (Williamson and Madison, 1990). At various stages in the refinement procedure, the atoms in a methylene or isopropyl group are explicitly swapped. The resulting conformation is accepted if it gives rise to a lower NOE energy, otherwise the original conformation is restored. This is done for all prochiral groups every 1000 steps of dynamics during the search phase, and for a small, randomly selected fraction each time the heat-bath temperature is reduced during cooling (i.e. every 50 K). No explicit swapping is performed during phase IV.

As can be deduced from Table 1, the prochiral groups are kept floating throughout the high-temperature search phase. During phase II the force constants of the nonbonded interactions and of the three-atom bond angles are slowly increased, which means that at the end of the first cooling phase (1000 K) true floating is no longer possible. As mentioned above, explicit swapping continues until the system has been cooled to 100 K during phase III.

The main advantage of floating chirality as opposed to the pseudoatom approach is that it allows using distance restraints which are closest to the experimental data for resolved but unassigned prochiral groups. Similarly, in the case of NOE contacts to a group of equivalent spins two approaches can be distinguished. Wüthrich et al. (1983) proposed to have such NMR constraints refer to a single pseudoatom at the mean position of the atoms in the equivalent group. Again this involves adding pseudoatom corrections to the measured distance, resulting in a less accurate set of NOE constraints. In particular, introducing a pseudoatom which replaces the H^{δ} and H^{ε} protons in tyrosine and phenylalanine residues results in a loss of directionality for the aromatic ring. In contrast, in the so-called R^{-6} average method the interspin distances are measured separately for each proton in the equivalent group (r_i) and are appropriately averaged into a single 'effective' distance R; $R = \langle r_i^{-6} \rangle^{-1/6}$ (Brünger et al., 1986). This calculated average distance can only be compared to the experimental distance if the latter corresponds to an 'average' cross-peak volume, i.e. cross-peak integrals involving equivalent spins should be divided by the number of overlapping resonances (Brünger et al., 1987). Here, we used the method known in X-PLOR as 'sum averaging', originally introduced to treat ambiguous NOEs in symmetrical multimers (Nilges, 1993). With this method, the sum of the separate interactions is computed ($R = \{\sum r_i^{-6}\}^{-1/6}$) instead of the average, which means that the aforementioned divisions need not be performed. Hence, in combination with the floating chirality approach all cross-peak volumes integrated from the 2D or 3D spectra can be converted into distances without any kind of correction. Thus, after calibration we have converted all peak integrals from the ^{13}C- and ^{15}N-edited NOESY spectra directly into distances using a plain R^{-6} relation. We used sum averaging for all experimental distance restraints after separating unambiguous and ambiguous NOEs (Nilges, 1993). This means we used an R^{-6} sum also for the methyl groups, for which an R^{-3} sum would be more appropriate (Tropp, 1980; Koning et al., 1990). The difference, however, is not very large, the former being slightly more conservative. All nondegenerate methylene protons and isopropyl groups have been named according to their relative chemical shifts, e.g. $H^{\beta 2}$

and $H^{\beta 3}$ simply correspond to the low-field and high-field resonating lines, respectively.

In order to investigate the dependence of the floating chirality procedure on the quality of the NOE input and the efficiency of the molecular dynamics protocols, we calculated six different sets of structures. The set denoted A is the family of structures calculated from what is to be considered the optimal NOE input. A total of 1145 independent NOEs were used per protein monomer, 543 of which are intraresidue, 210 are sequential, 74 are medium range (two to five residues apart in the sequence) and 318 are long range (five residues or more apart in the sequence). Upper distance restraint limits were set by lengthening all determined distances r in an empirical fashion by r.max[0.15,(0.15 + (r − 2.6) 0.08)], and lower limits were set at 0.85r.

The NOE input for set B differs from that of A in that no lower limits are used, and that of set C uses the upper and lower bounds of A but now lengthened by 0.5 Å. Sets D and E were calculated from the same NOE input as A but the floating chirality was treated differently in the molecular dynamics protocols. While sets A–C were calculated with the combined floating–swapping approach described above, in set D only the floating aspect was incorporated and no explicit swaps were performed. Set E was calculated using only explicit swapping in the prochiral groups, while floating was prevented by keeping the three-bond angles close to their equilibrium values. Set F is the ensemble of structures calculated from the 75 ms ^{15}N- and ^{13}C-edited NOESY spectra reported earlier (Folmer et al., 1995b). Upper and lower restraint limits were set as for A, and the combined floating–swapping procedure was used.

Generally, in large systems not every random conformation will converge into a properly folded structure. In our case only approximately 40–50% of the conformations calculated in stage 0 with the optimal NOE input were sufficiently folded to converge into acceptable low-

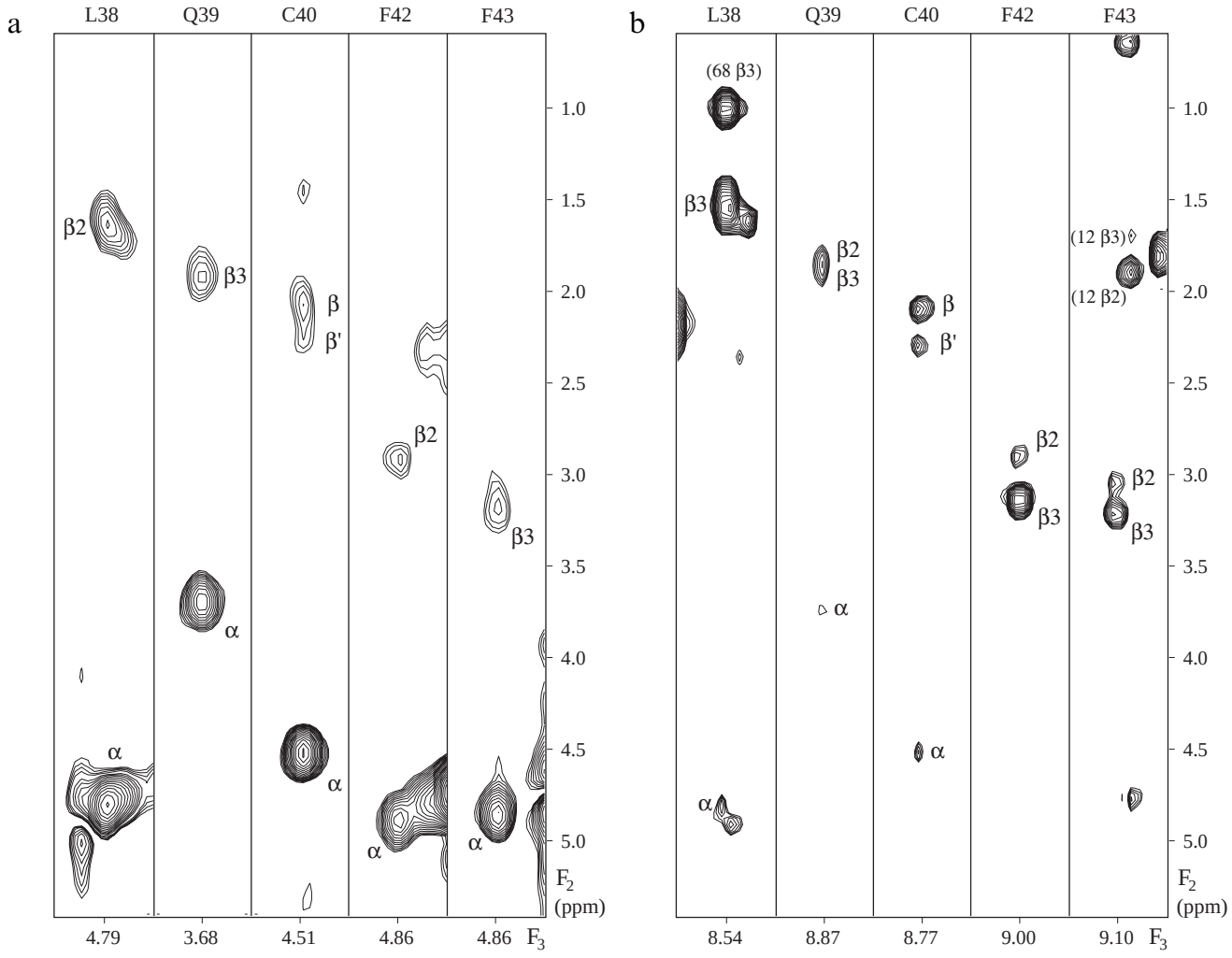


Fig. 1. (a) Strips from the 500 MHz 3D HACAHB-COSY spectrum of 0.9 mM doubly labelled Pf3 F36H ssDBP. The strips are taken at the $^{13}C^{\alpha}$ and $^{1}H^{\alpha}$ frequencies of residues 38–43, with the exception of Gln^{41} which has degenerate β-protons. (b) Strips from the 500 MHz 3D HNHB spectrum of 2.0 mM ^{15}N-labelled Pf3 F36H ssDBP. The strips are taken at the ^{15}N and $^{1}H^{N}$ chemical shifts of the residues shown in (a).

TABLE 2
STATISTICS OF THE FAMILY[a] OF STRUCTURES

Parameter	⟨A⟩	⟨B⟩	⟨C⟩	⟨D⟩	⟨E⟩	⟨F⟩
X-PLOR energies						
E_{tot} [b]	291 ± 35	132 ± 17	55 ± 6	474 ± 49	376 ± 67	350 ± 20
E_{bond}	14 ± 2	7.0 ± 1.2	1.9 ± 0.5	25 ± 4	20 ± 4	19 ± 2
E_{angle}	107 ± 13	57 ± 5	38 ± 2	140 ± 14	125 ± 20	124 ± 7
E_{impr}	18 ± 3	8.7 ± 0.9	4.8 ± 0.2	21 ± 3	21 ± 6	22 ± 2
E_{vdW}	27 ± 7	13 ± 4	4.9 ± 2.2	49 ± 8	40 ± 10	27 ± 4
E_{NOE}	124 ± 18	47 ± 9	4.5 ± 3.2	240 ± 37	169 ± 34	158 ± 12
Atomic rms differences[c]						
Backbone	0.69 ± 0.20	0.81 ± 0.27	1.08 ± 0.31	0.76 ± 0.13	0.76 ± 0.16	0.65 ± 0.18
All non-H	1.13 ± 0.16	1.22 ± 0.26	1.51 ± 0.29	1.22 ± 0.11	1.23 ± 0.15	1.10 ± 0.14

[a] The different ways in which the six ensembles of structures were generated are explained in the Materials and Methods section. In short they are: A: calculated from 'optimal' NOE input (see text), obtained from 40 ms NOESYs; B: as A, but no lower limits on NOE restraints; C: as A, but lower and upper limits relaxed by 0.5 Å; D: as A, but no explicit swaps in prochiral centres during dynamics, only floating; E: as A, but only explicit swaps, no floating; F: as A, but distance constraints extracted from 75 ms spectra.
[b] The force constants of the various energy terms are listed in Table 1.
[c] Rms differences of ⟨X⟩ from $\langle\overline{X}\rangle$ are listed for residues 1–11 and 25–78 (see text), where X denotes each of the six calculated families.

energy structures after annealing. This number will be lower when less accurate or precise NOE data are used. To remove these effects from our floating assignment study, we generated 80 starting conformations for each set (A–F), ensuring that each annealing calculation at least results in properly folded molecules.

Although the Pf3 ssDBP is a homodimer in solution, we did not enforce this symmetry in the calculations. Normally, one would use an extra energy term minimizing the rms difference between the two monomers, as well as a set of pseudo-NOEs ensuring twofold symmetry (Nilges, 1993). However, in order to test floating chirality in the most general application (i.e. nonsymmetric systems) we discarded all symmetry terms from the calculations. Effectively, this corresponds to calculating a 156-residue monomer, but it offers the additional advantage that better statistics are obtained as two equivalent prochiral centres can be independently evaluated per structure.

Finally, the amino protons of asparagine and glutamine residues, usually giving rise to well-resolved NMR signals, were also treated with a floating assignment approach. For those, a somewhat different protocol had to be used, aimed at keeping the H_2N-C=O moiety flat throughout the entire dynamics run. This was achieved by leaving all angle and improper force constants unaffected. The amino protons then can only change positions through explicit swapping; at various stages in the MD protocol (Table 1) the protons are swapped and the new conformation is accepted if its contribution to the NOE potential is lower than that of the original. The main difference to the methylene and isopropyl groups is that the improper angle defining the stereogeometry (C^{β}-C^{γ}-$N^{\delta 2}$-$H^{\delta 21}$ in asparagine, C^{γ}-C^{δ}-$N^{\varepsilon 2}$-$H^{\varepsilon 21}$ in glutamine) changes by 180° every time a new conformation is accepted, while the corresponding force constant remains unchanged (and nonzero). In contrast, the force constants of the tetrahedral centres are always set to zero, so that the actual values of their improper angles have become irrelevant.

Results

NMR analysis

Figure 1 shows strips taken from the 3D HACAHB-COSY (Grzesiek et al., 1995) and 3D HNHB (Archer et al., 1991) spectra of Pf3 ssDBP recorded at 500 MHz. Both spectra are of relatively high quality with good signal to noise ratios. The HNHB spectrum reveals nearly all β-protons, also when the three-bond ^{15}N-H^{β} coupling corresponds to a gauche conformation. In the HACAHB-COSY spectrum typically only one β-proton is visible, the one trans to the H^{α}. To rise above the noise level of this spectrum, an H^{α}-H^{β} cross peak should result from a $^3J_{\alpha\beta}$-coupling larger than about 4–5 Hz. The absence of such a cross peak, therefore, is diagnostic of the smaller gauche coupling. In total, the 78 amino acid Pf3 ssDBP contains 51 β-methylene groups, 43 of which give rise to resolved NMR signals. The combined analysis of the aforementioned experiments yielded stereospecific assignments for 30 of these centres. The remaining 13 prochiral β-pairs could not be unambiguously assigned due to spectral overlap (five residues) or because the J-couplings suggest significant degrees of rotamer averaging (eight residues).

Analysing the $^3J_{C\gamma N}$ and $^3J_{C\gamma C'}$ couplings in the $\{^{15}N\}$ and $\{^{13}CO\}$ spin-echo HSQC spectra (Grzesiek et al., 1993; Vuister et al., 1993), stereospecific assignments could be made for four out of the five valine isopropyl groups in Pf3 ssDBP. The various J-couplings measured for Val^{11} indicate the presence of rotamer averaging. The relative resonance positions of the β-methylene and valine isopropyl frequencies that could be assigned are presented in Fig. 2. A complete list of all the chemical shifts will be reported elsewhere.

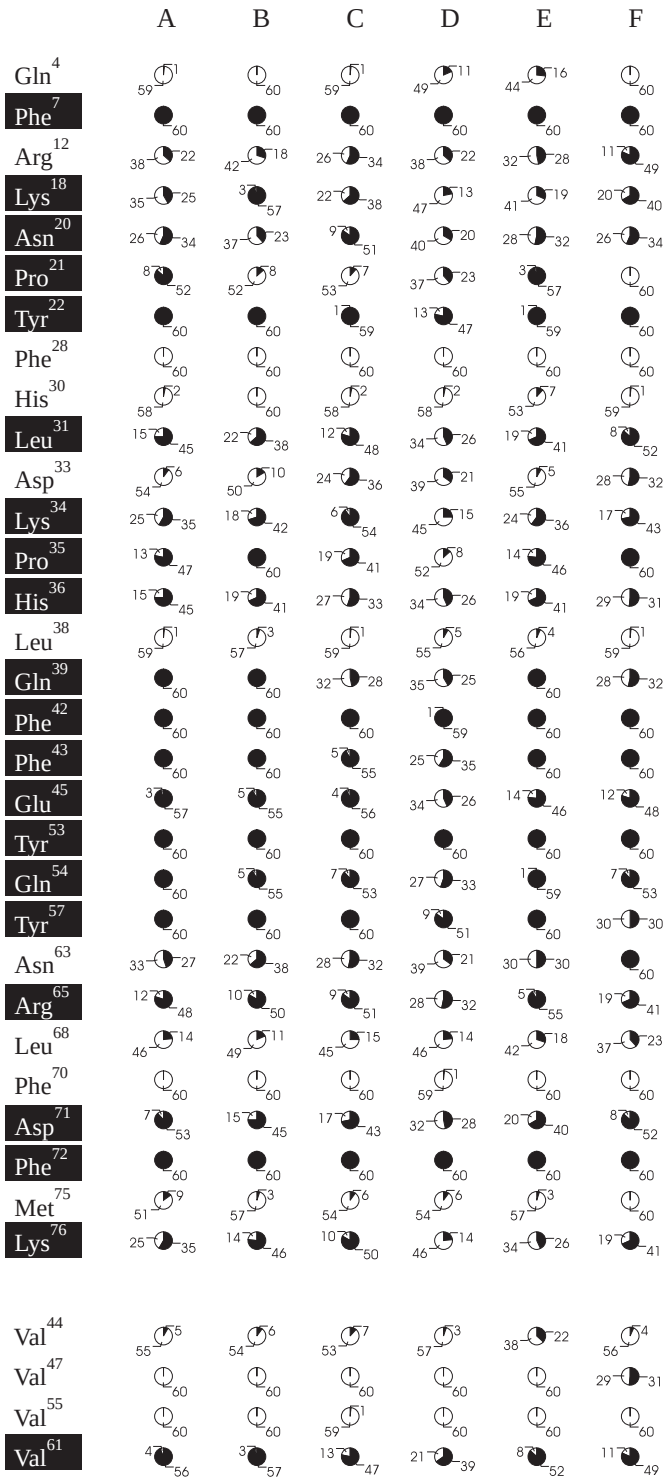


Fig. 2. Overview of the final prochiralities adopted by the β-methylene and valine isopropyl groups for which stereospecific assignments were available from J-coupling experiments. If the residue is printed black-on-white (e.g. Gln[4]) $H^{\beta 2}$ is the low-field and $H^{\beta 3}$ the high-field shifted resonance (as established experimentally). For these residues, the nomenclature used in the NOE input was (by 50% chance) already correct. Then, a correct floating assignment would be obtained if the two β-protons (methyls in valine) do not change positions during the dynamics, because the calculations started from molecules with the proper (IUPAC) prochirality. Conversely, if the residue is printed white-on-black $H^{\beta 2}$ should be the high-field shifted proton, while it is the low-field shifted proton in the NOE list (e.g. Phe[7]). Therefore, the atoms in the prochiral centre should swap positions during the dynamics to end up in the correct prochirality. The little pie charts show in white the number of centres that (effectively) did not swap, and in black the centres that did swap. As an example, in Phe[7] all 60 β-methylene protons switched positions in all six families. This corresponds to a correctly adopted prochirality (the residue is printed white-on-black), and hence to a successful floating assignment for Phe[7] in all six calculations. Arg[12], on the other hand, is an example of a residue whose prochiral centres could not be uniquely assigned in any of the calculations.

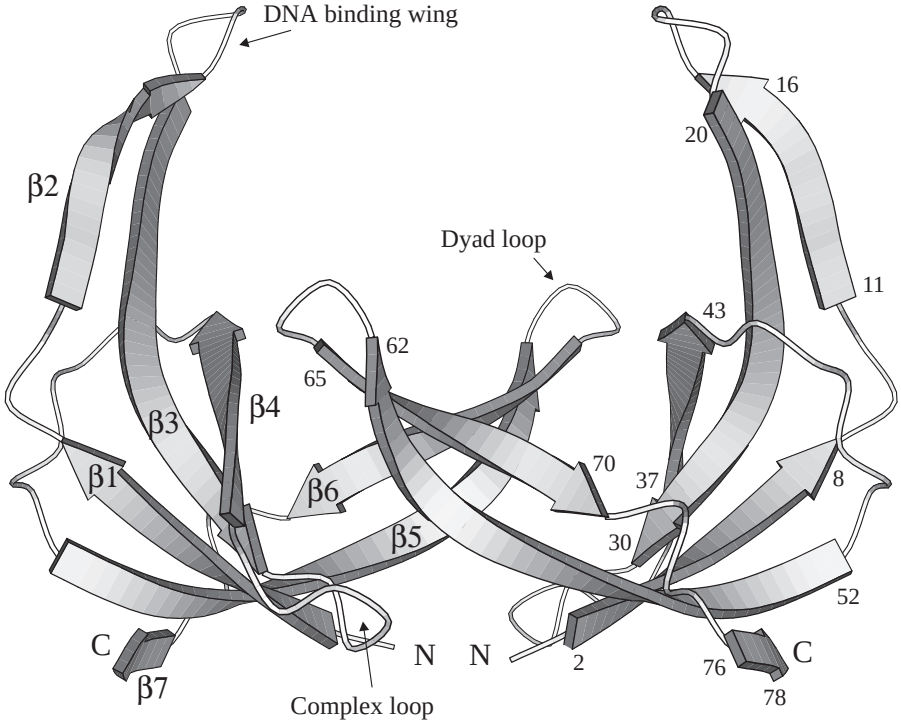


Fig. 3. Schematic drawing of $\langle\overline{A}\rangle_r$. The three major loops are denoted on the left-handed monomer. The seven β-strands are indicated with their strand numbers (left-handed monomer) and first and last residue numbers (right-handed monomer). The molecule was drawn with MOLSCRIPT (Kraulis, 1991).

Structures and floating chirality

Six ensembles of 80 structures were calculated using simulated annealing from six sets of 80 approximately correctly folded conformations (see the Materials and Methods section). In each ensemble, the 30 structures with the lowest total energy were selected, which will be denoted ⟨A⟩ to ⟨F⟩. No distance restraint was violated by more than 0.5 Å in any of these final structures. The 30 structures in each family were superimposed for the backbone atoms (N,C^{α} C') of residues 1–11 and 25–78, i.e. excluding the rather flexible DNA-binding wing (Folmer et al., 1995b), and the coordinates were averaged to create the respective average structures, $\langle\overline{A}\rangle$ to $\langle\overline{F}\rangle$. These were energy minimized through 1500 steps of conjugate gradient minimization to give $\langle\overline{A}\rangle_r$ to $\langle\overline{F}\rangle_r$. Table 2 affords some statistics for the six ensembles of 30 ssDBP structures, and Fig. 3 shows a schematic representation of $\langle\overline{A}\rangle_r$.

⟨A⟩ to ⟨F⟩ were analysed for the final prochirality adopted by the centres subjected to floating chirality (all isopropyl and nondegenerate methylene groups). As all symmetry terms were excluded from the calculation, 60 conformations were evaluated per prochiral centre in these ensembles. Figure 2 shows how the two possibilities (*pro-R* and *pro-S*) are distributed over the 60 monomers for all prochiral centres for which unambiguous stereospecific assignments could be obtained from the J-coupling experiments.

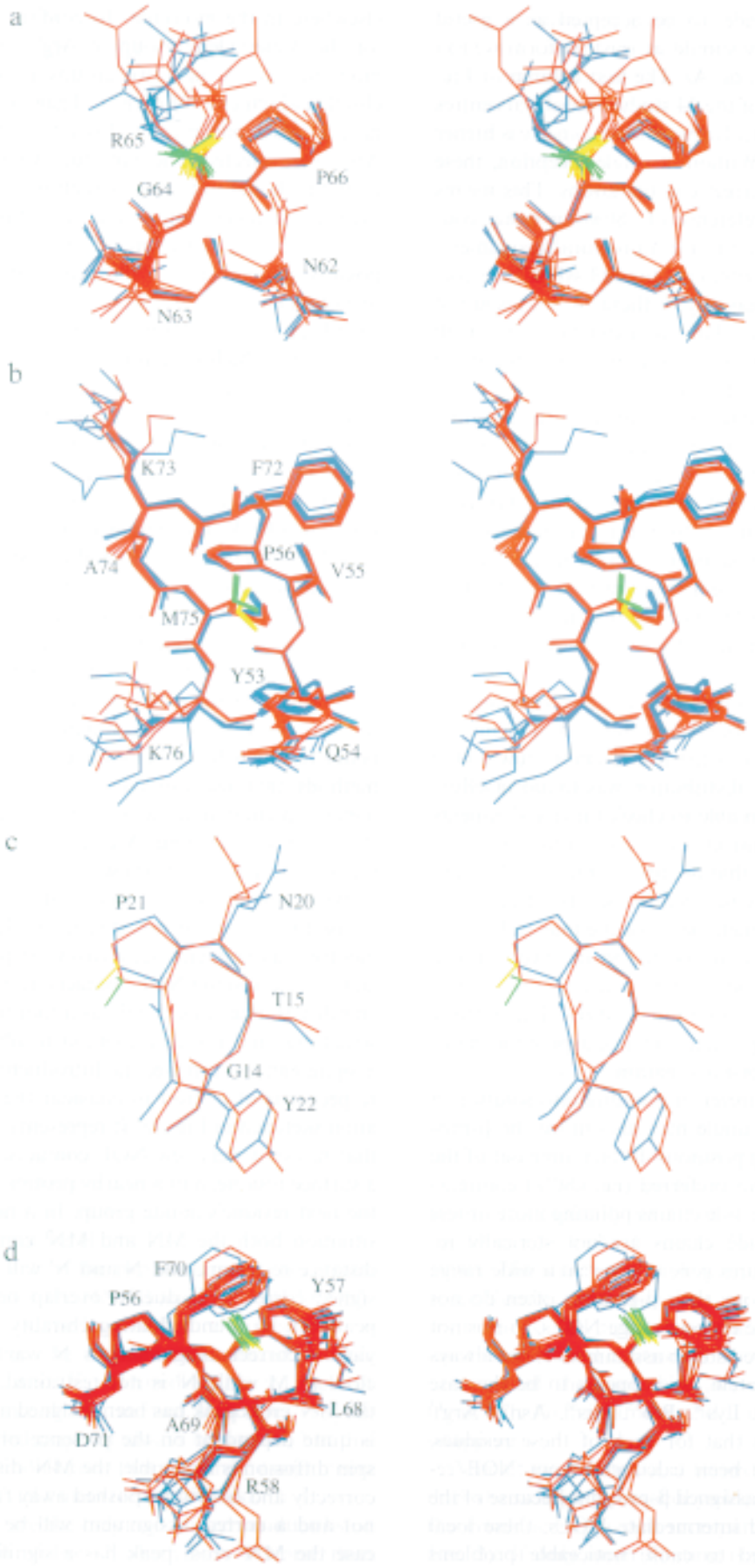
Discussion

General features

The structures in family A were calculated from completely assigned 40 ms NOESY spectra, using rather tight upper and lower distance bounds (±15% of the distance). To current standards, this would be considered fairly optimal experimental NOE input for biomolecules of the

→

Fig. 4. Stereoviews of substructures centred around a prochiral group. Structures with the correct and incorrect prochiral assignments are shown in blue and red, respectively. Only non-hydrogen atoms are shown, except for the methylene protons of interest, which are coloured yellow and green when assigned correctly and incorrectly, respectively. For clarity, only about 15 structures have been plotted, arbitrarily taken from the ensembles. (a) Substructures taken from ⟨A⟩ centred around the β-CH_2 of Arg^{65}. Note that the two assignments result in clearly different rotamer states of χ_1 in Arg^{65} (except for one 'red' structure). (b) Substructures taken from ⟨A⟩ centred around the β-CH_2 of Met^{75}. Particularly the position of the side chain of Tyr^{53} appears to be sensitive to the assignment of the Met^{75} methylene group. This is basically the result of one (weak) NOE between Tyr^{53}-H^{δ} and Met^{75}-$H^{\beta 2}$. A similar but less pronounced effect is observed for the ring of Phe^{72}. (c) Comparison of $\langle A\rangle_r$ (blue and yellow) and $\langle C\rangle_r$ (red and green) around Pro^{21}. (d) Substructures taken from ⟨F⟩ centred around the β-CH_2 of Tyr^{57}.


a
R65
G64
P66
N62
N63
b
K73
F72
P56
V55
A74
M75
Y53
K76
Q54
c
P21
N20
T15
G14
Y22
d
F70
P56
Y57
L68
A69
D71
R58

size of Pf3 ssDBP. Hence, to be accepted as a useful method floating chirality should at least perform well in the structure calculation of ⟨A⟩. The first column of Fig. 2 shows that for 24 out of the 34 studied prochiral centres the preference for a particular prochiral geometry is higher than 48/60 monomers. Without a single exception, these turned out to be the correct configurations. This means that if one accepts a preference of ≥80% for either configuration to be the criterion for a prochiral assignment, then all 24 assignments obtained for Pf3 ssDBP are correct. Interestingly, the majority of these 24 centres are of buried residues. In particular, the aromatic rings of all tyrosines and phenylalanines listed in Fig. 2 (except Phe^{43}) are very much buried in the protein, displaying a large amount of long-range NOEs. Consequently, there is little play in the χ_1 angle of these residues, and exchanging positions has become the only degree of freedom along which the methylene protons can adapt to the NOE constraints. Therefore, flipping of β-protons in (buried) tyrosines and phenylalanines appears to be a true yes-or-no situation which is likely to enhance the success rate. This suggestion is supported by the fact that in ⟨A⟩ the correct configuration for each of these residues is found in all 60 monomers.

To the same extent, this is true for the α-protons of glycines, and we indeed found that their assignments are highly decisive. Four out of the five glycines show a 60/0 preference, while a 58/2 distribution was found in Gly^{51}. So far, we have not been able to check these assignments experimentally, but intuitively one may feel that these are likely to be correct, and that floating chirality is also very much suited to treat glycine residues. Similar effects were also observed for the γ-methylenes of the four isoleucines in Pf3 ssDBP. These are all located in the hydrophobic core of the molecule, and their γ- and δ-methyl groups display a vast number of long-range NOEs. These firmly define both the χ_1 and χ_2 angles, again clasping the methylene protons in a yes-or-no situation.

If the side chain is shorter or pointing into solution, a rotation around the χ_1 angle may also move the β-protons into NOE-fulfilling positions. In fact, nine out of the 10 residues for which no preferred (i.e. <80%) configuration is found have their side chains pointing more or less into solution. These side chains are not sterically restrained and the structures generally reveal a wide range of χ_1 angles. Furthermore, their β-protons often do not display any intermediate or long-range NOEs, so it is not surprising that a preferred stereo assignment is not always induced. In the Pf3 protein, this appears to be the case for Arg^{12}, Lys^{18}, Asn^{20}, Lys^{34}, Pro^{35}, His^{36}, Asn^{63}, Arg^{65} and Lys^{76}. This means that for each of these residues, several structures have been calculated from NOE restraints referring to misassigned β-protons. Because of the lack of long-range and intermediate NOEs, these local errors are not expected to cause noticeable problems elsewhere in the molecule. To confirm this, for instance, for the β-methylene group of Arg^{65}, we divided the 60 monomers of ⟨A⟩ into two groups according to the prochiral assignment (of Arg^{65}). Then, the residues which have at least one proton closer than 5 Å to the C^{β} of Arg^{65} were selected. These residues were superimposed for their backbones to the first structure of the ensemble, and their coordinates were averaged within each group to yield two average structures. These correspond to the two possible assignments of the prochiral centre (of Arg^{65}), one of which of course is the wrong one. These two averaged local structures were compared *without* being superpositioned, which should reveal any possible interresidue effects originating from the incorrect prochiral assignment. These were, however, not observed as can also be deduced from Fig. 4a; the two groups of substructures are highly similar. The rms difference for the backbone atoms (N,C,C^{α}) between these two averaged structures is as small as 0.10 Å. This number varies from 0.05 (His^{36}) to 0.14 Å (Lys^{18}) for the nine residues listed above, generally well below significance in an ensemble of NMR structures. Only for Lys^{34} a backbone rms difference of 0.27 Å was measured between the two average structures, which might be explained by this residue being located at the tip of a rather flexible and somewhat disordered loop (Folmer et al., 1995b). One should, however, realize that better results will not be obtained for these residues with methods that use averaging to treat prochiral groups, namely, pseudoatoms with appropriate corrections, and R^{-6} or 'sum' averaging. A comparison of these different methods will be given below.

We conclude therefore that floating chirality can safely be used for solvent-exposed residues, despite the fact that incorrect assignments may easily be produced due to a lack of interresidue NOE contacts. In these cases floating chirality is indeed used only as a tool to calculate protein structures, and not as a method to obtain stereospecific assignments as such (see the Introduction section). In this respect it is of interest to consider the hypothetical situation sketched in Fig. 5a. It represents a methylene group that has only very few NOE contacts, e.g. the β-CH_2 of a surface residue, and a nearby proton, which could be of the next residue's amide group. In a nonspectral overlap situation both the MN and MN' contacts will become distance restraints, and N and N' will most likely be assigned correctly. If due to overlap only the MN cross peak can be found, floating chirality will probably still yield a correct assignment as N wants to be relatively close to M while N' is not restrained. However, if only the MN' cross peak has been assigned and used, the result is quite dependent on the presence of spin diffusion. If spin diffusion is negligible, the MN' distance is measured correctly and N' will be pushed away from M, while N is not and a correct assignment will be produced. But in case the MN' cross peak has a significant contribution

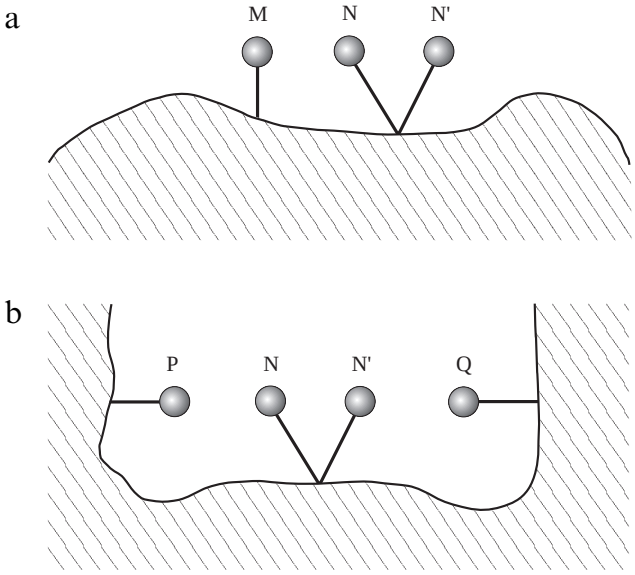


Fig. 5. (a) Schematic representation of a methylene group on the surface of a protein and a nearby proton M. Both the M–N and M–N' distances are short enough to produce an NOE cross peak, that of M–N of course being more intense. (b) Schematic picture of a methylene group in the interior of the protein and two protons fixed in the protein matrix. Both the P–N and Q–N distances are short enough to produce NOEs.

from spin diffusion through proton N, the MN' distance will be underestimated, which can easily result in proton N' taking up the position of N. Thus, N and N' are swapped during the MD simulation, yielding a wrong stereospecific assignment. We emphasize that the resulting structure is nevertheless essentially the same as the correct structure, with one proton close and the other more distant to M. In the end only the proton names are incorrect, which is of interest to chemical shift libraries but not really to the protein structure itself. Of course, Fig. 5a illustrates an oversimplified situation, but the effects described here are likely to play a role in protein structure determination through floating chirality. In Pf3 ssDBP this was observed, for instance, for Gln^{13}, which is located at the surface of the protein pointing away from the molecule. In the structures calculated from the 75 ms NOESYs, a 59/1 distribution was found for $H^{\beta 2}$ being the low-field shifted proton, while the 40 ms NOESYs produce a 60/0 preference for $H^{\beta 2}$ being the high-field proton (it is not known which of the two assignments is correct, due to overlap in both the HNHB and HACAHB-COSY spectra). Nevertheless, the local structure around Gln^{13} is basically identical in the respective minimized average structures (i.e. $\langle \overline{A} \rangle_r$ and $\langle \overline{F} \rangle_r$). Moreover, the χ_1 angle of Gln^{13} differs by only 4° in the two structures. So the assignments obtained may not always be correct, but this example clearly demonstrates that incorrect assignments do not necessarily yield incorrect structures.

Leu^{68} and Met^{75} are the only two buried residues listed in the first column of Fig. 2 which have a worse than 58/2 preference for either of the prochiral assignments. Leu^{68} is located very close to the dyad axis of the molecule and, as a consequence, all its NOEs had to be incorporated as ambiguous (Nilges, 1993). Considering furthermore that its δ-methyl groups are treated with floating chirality as well, it is not too surprising that the convergence rate for this particular residue is not optimal. To check whether the misassignment gives rise to serious structure errors, we performed a similar evaluation as described above for Arg^{65}. The rms difference between the two averaged substructures (residues 57–59 and 67–70) is 0.24 Å for the backbone and 0.72 Å for all atoms (excluding the distorted side chain of Arg^{58}). These numbers indicate that the misassignment does have some interresidue effects, albeit rather small. It is less clear what causes the wrong prochiral assignments in Met^{75}; there are many NOEs to this residue, in particular to the ε-methyl group and the $H^{\beta 3}$ proton. No long-range NOEs, however, were observed for $H^{\beta 2}$, which may help to explain why it ends up at the wrong side of $H^{\beta 3}$ in a number of structures. Figure 4b shows how the local structure around Met^{75} is affected by the misassignment. In fact, it is the only example in Pf3 ssDBP where a wrong assignment caused noticeable long-range effects. Still, the overall differences are relatively small, although Met^{75} is one of the most buried residues in the protein.

Finally, in Pf3 ssDBP many asparagines and glutamines are solvent-exposed and their amino protons frequently do not show very different NOE contacts. Clear preferences nevertheless were obtained for Asn^{20}, Gln^{25}, Gln^{41}, Asn^{60} (all 60/60) and Asn^{62} (58/60), of which only Gln^{25} is a buried residue. The amino protons of the remaining two asparagines and four glutamines showed nondecisive distributions of the two possible geometries.

Floating or swapping?

When floating chirality was first introduced in MD calculations (Holak et al., 1989), the prochirality of a methylene group could only be reversed if the two protons moved (floated) into each other's positions. Alternatively, atom swapping protocols can be used in which the atoms are explicitly swapped if the opposite geometry results in a lower NOE energy (Williamson and Madison, 1990). The structures of A were calculated with a combined floating–swapping strategy, the floating part being restricted to the high-temperature (>1000 K) stages. To test the individual contributions of these two tools to the outcome of the procedure, we calculated ensemble D using only floating atoms, and E using only explicit atom swapping.

Much to our surprise, it turned out that for most residues the swapping protocol yielded very similar results as the combined approach (compare columns A and E in Fig. 2). If only atom floating is employed (column D),

much less correct assignments are made. In particular, residues Pro[21], Asp[33], Gln[39], Phe[43], Glu[45], Gln[54] and Val[61] show clear preferences in ⟨A⟩ and ⟨E⟩ while random-like *pro-R*/*pro-S* distributions are found in ⟨D⟩. The opposite result occurs only once; Val[44] is correctly assigned in ⟨A⟩ and ⟨D⟩, while without the floating atoms a random-like distribution is found. Furthermore, sometimes only the combined approach gives rise to clear assignments, whereas both separate methods yield unsatisfactory results (e.g. Gln[4], Glu[45], Asp[71]). These different observations infer that both atom floating and swapping should be included in floating chirality protocols. This is further supported by the finding that there are no residues for which omitting either floating or swapping yields significantly better assignments (with the possible exception of Pro[21] and Arg[65]). In other words, the results in ⟨D⟩ or ⟨E⟩ are never better than those in ⟨A⟩.

Analysing the results for the 'rigid' methylenes (vide supra), such as those of glycines, prolines and isoleucines, we conclude that here explicit swapping is particularly necessary to get proper assignments. In ⟨D⟩ many of these show random distributions while 59/1 and 60/0 ratios are found in ⟨E⟩ (data not shown). This can be explained from the discussion above relating to the yes-or-no situation for these methylene groups. Explicit swapping is of course perfectly suited to handle these 'rigid' centres.

Sensitivity towards accuracy and precision of NOE restraints

So far, the results described were from structures calculated from a rather precise and accurate NOE data set (for definitions of 'precise' and 'accurate', see Zhao and Jardetzky (1994)). Upper and lower bounds were set at ±15% of the distance, and the NOE mixing time was as short as 40 ms. Often one will use longer mixing times and, to compensate for concomitant spin diffusion, less precise bounds. To investigate whether floating chirality still performs well with less optimal NOE input, we calculated three more families of structures.

First, to individually study the need for lower bounds, family B was generated from NOE input in which all the lower bounds were set to zero (i.e. not enforcing a minimal distance). Interestingly, the distribution of assignments in ⟨B⟩ is very similar to that in ⟨A⟩; only four out of the prochiral centres listed in Fig. 2 show significantly different *pro-R*/*pro-S* ratios. Pro[21] and Asp[71] are more successfully assigned in ⟨A⟩ while Lys[18] and Pro[35] in fact show better results when lower bounds are absent. These results suggest that the presence of lower bounds is not essential and hence that the upper bounds strongly dominate the floating chirality process.

Still the upper bounds used are relatively tight. To test whether such precise restraints are necessary, family C was calculated from distances whose upper and lower bounds had been relaxed by an additional 0.5 Å. Despite this considerable distance of 1.0 Å added to the zero-potential bottom of the NOE term, the success rate of the automatic assignment did not drop dramatically. Compared to ⟨A⟩, selectivity is only lost for Asp[33], Gln[39] and Asp[71], and to a lesser extent also for the isopropyl group of Val[61]. A few more residues (His[36], Phe[43], Val[44], Glu[54]) show slightly less conclusive *pro-R*/*pro-S* ratios, but the differences are rather small taking into account the loose NOE restraints. Like in ⟨B⟩, the distribution in ⟨C⟩ of configurations for the β-CH_2 of Pro[21] is almost opposite to that in ⟨A⟩, the latter corresponding to the correct assignment. Apparently, there is a crucial lower bound that, when tight enough, enforces the correct positioning of the two β-protons. The overall effect of the other NOE restraints, however, seems to be more compatible with the opposite assignment. This may be explained by the fact that spin diffusion is quite effective in proline residues, even during 40 ms in an 18 kDa particle. Figure 4c compares the local structures around Pro[21] of $\langle\overline{A}\rangle_r$ and $\langle\overline{C}\rangle_r$, in which the proline β-protons have adopted opposite positions. Although the two structures are difficult to compare because they were derived from different NOE data, the figure shows that the misassignment of Pro[21] in $\langle C\rangle_r$ does not significantly affect the overall structure. As Pro[21] is on the surface of the molecule, possibly a situation applies here similar to that sketched in Fig. 5a, for which it was indeed argued that wrong assignments do not necessarily cause structural errors.

Thirdly, to investigate the effect of a less accurate NOE data set, family F was generated from the 75 ms NOESY spectra. Cross peaks in these spectra will contain a significant contribution from spin diffusion, but in most cases the two protons in a methylene group still show slightly different NOE intensities. The same setting of upper and lower bounds was used as for A, to which F therefore must be compared. From an inspection of Fig. 2 it is apparent that for many residues the assignment procedure has yielded the proper results in ⟨F⟩; in particular, the aromatic residues (except Tyr[57]) show 60/0 preferences for the correct assignment. Rather similar *pro-R*/*pro-S* distributions are found in ⟨A⟩ and ⟨F⟩, the ratios in the latter being slightly smaller. Some residues, however, could not be assigned from the 75 ms spectra, while conclusive preferences were found in ⟨A⟩ (Gln[39], Val[47], Tyr[57]). Figure 4d shows the local structures of ⟨F⟩ around Tyr[57] for both assignments of its methylenes. Clearly, these structures are virtually indistinguishable, indicating that the NOE input simply cannot differentiate between the two assignments. Again, if wrong assignments easily appear, these do not seem to have significant effects on the overall structure. Other residues (Pro[21], Pro[35], Asn[63], Met[75]) show in fact conclusive *pro-R*/*pro-S* ratios in ⟨F⟩, while less so in ⟨A⟩. Of these, Pro[21] and Asn[63] are incorrectly assigned; Asn[63] is a surface residue whose

assignment has no consequences for nearby residues, and problems with Pro[21] have been described above (see Fig. 4c).

An interesting observation during the structure elucidation of Pf3 ssDBP (Folmer et al., 1995b), using floating chirality and 75 ms NOESY spectra, was that for many pairs of NOEs connected to methylene groups (particularly of aromatic residues and glycines) often the longer distance was consistently violated while the shorter was not. This is very indicative of the presence of spin diffusion, and in the early stages of the structure calculation process these longer restraints were usually removed or relaxed. These violations nicely illustrate that floating chirality can be quite advantageous also when some amount of spin diffusion is present. Apparently, relatively small differences in intensities between the two NOEs of a methylene group to a third atom are sufficient for the prochiral protons to be distinguished. Because no pseudoatom corrections were applied, the closest proton did properly end up close to this third atom. Spin diffusion obviously caused the distant proton to experience a too tight NOE restraint, resulting in a distance violation.

Table 2 shows that quite large differences occur in the statistics of the six sets of structures. These can be explained directly from the differences in NOE input or calculation protocol. The energies of ⟨B⟩ are lower than those of ⟨A⟩ simply because the distance restraints are less restrictive, which is even more so in ⟨C⟩. Concomitantly, the atomic rms differences increase upon going from ⟨A⟩ to ⟨C⟩. The fact that the energies in ⟨D⟩ and ⟨E⟩ are higher than those of ⟨A⟩ is a more interesting observation, indicating that the combined swapping–floating protocol indeed yields the best structures. Apparently, the higher energies in ⟨D⟩ and ⟨E⟩ are due to incorrectly assigned prochiral centres, causing geometric stress. The energies of ⟨F⟩ cannot be compared to the others because these structures were calculated from different NOE spectra.

Relation to pseudoatom approach and R^{-6} averaging

In cases where the floating chirality method does not produce a unique result, some of the structures have incorrect assignments at some prochiral centres. It is illustrative to compare the behaviour of floating assignment to that of methods that use averaging to treat prochiral groups, namely, pseudoatoms with appropriate corrections, and R^{-6} or 'sum' averaging.

The distinctive difference between floating stereospecific assignment and averaging methods is that only floating assignment assures that the assignment of prochiral groups is consistent in one structure (Habazettl et al., 1990). This is illustrated by the hypothetical case sketched in Fig. 5b, where two protons that are fixed in the protein matrix (P and Q) have an NOE to the *same* proton of a methylene group. With floating assignment, both NOEs would be assigned either to proton N or proton N'. R^{-6}-type averaging, on the other hand, could assign the NOE from proton P to proton N, and that from proton Q to proton N'. Since this relaxes effectively the restraint by 1.8 Å in the situation of Fig. 5b (i.e. the N–N' distance), this might in fact be the more likely outcome. A pseudoatom between N and N' with the worst-case correction of 0.9 Å would have a very similar effect, i.e. de facto one might end up with structures with effectively inconsistent and partially incorrect assignments. Both methods could conceivably lead to a rotation of the methylene group or local distortions of the structure. It seems therefore unlikely that the convergence problems observed for the residues discussed above when using floating chirality could be alleviated by resorting to an averaging technique.

As discussed before, any of the methods should really be seen as a way to sample the space of conformations and possible assignments of the prochiral groups for low-energy conformations. Since floating assignment restricts the sampling to assignments that are internally consistent, it seems the preferable method. Another important advantage is the easy automation of floating chirality, as described in the Materials and Methods section.

Conclusions

We have demonstrated that floating stereospecific assignment is a reliable tool in protein structure calculation employing restrained MD, allowing one to stay as close to the NOE data as possible. The calculated structures are of a quality almost comparable to that obtained with experimental stereospecific assignments. Furthermore, floating chirality is a very easy method to use since no data manipulations have to be performed before the calculation (e.g. averaging of the intensities to use R^{-6} averages or choice of the larger intensity). Also, in the form described here it lends itself especially well to cases where ambiguous NOEs are treated with sum averaging (as is necessary in homodimeric proteins). Optimal results were obtained only when explicit atom swapping was included in the protocol, and the combined floating–swapping approach appears to be the most successful. Floating assignment has a very high probability to produce the correct result for 'rigid' methylene and isopropyl groups and for buried residues, also when less precise or accurate NOE input is used. Solvent-exposed residues may not always be correctly assigned, even when using fairly good NOE data. In Pf3 ssDBP, however, this never resulted in noticeable interresidue effects, indicating that floating chirality can safely be used for all prochiral groups in a protein.

An important observation in this study is that the correct prochiral assignment is often found with a very high preference when this is essential to the protein's structure (e.g. in core residues). Conversely, when both

prochiralities are equally produced, these do not seem to have significant interresidual consequences for the three-dimensional structure.

Finally, as was already pointed out in the Introduction section, we believe that floating chirality should be used as a tool in protein structure calculation, rather than as a method to produce stereospecific assignments. In other words, one should refrain from converting the *pro-R*/*pro-S* distributions found in an ensemble of structures into true assignments. In particular, we would discourage using such assignments in subsequent structure refinement cycles. Floating chirality should be considered as an alternative to pseudoatom correction or R^{-6} averaging, and not to NMR experiments that directly yield stereospecific assignments.

Acknowledgements

The NMR experiments were performed at the SON Hf-NMR facility (Nijmegen, The Netherlands). Jos Joordens and Jan van Os are acknowledged for expert technical assistance. We thank Tineke Papavoine for a critical reading of the manuscript. This research was supported by the Netherlands Organization of Scientific Research (NWO), and in part by a European Union Access to Large Scales Facilities grant (ERBCHGECT940062) to EMBL.

References

Archer, S.J., Ikura, M., Torchia, D.A. and Bax, A. (1991) *J. Magn. Reson.*, **95**, 636–641.

Beckman, R.A., Litwin, S. and Wand, A.J. (1993) *J. Biomol. NMR*, **3**, 675–700.

Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., DiNola, A. and Haak, J.R. (1984) *J. Chem. Phys.*, **81**, 3684–3690.

Brünger, A.T., Clore, G.M., Gronenborn, A.M. and Karplus, M. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 3801–3805.

Brünger, A.T., Clore, G.M., Gronenborn, A.M. and Karplus, M. (1987) *Protein Eng.*, **1**, 399–406.

Brünger, A.T. (1992) *X-PLOR. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT, U.S.A.

Cavanagh, J., Palmer, A.G., Wright, P.E. and Rance, M. (1991) *J. Magn. Reson.*, **91**, 429–436.

Driscoll, P.C., Gronenborn, A.M. and Clore, G.M. (1989) *FEBS Lett.*, **243**, 223–233.

Engh, R. and Huber, R. (1991) *Acta Crystallogr.*, **A47**, 392–400.

Fletcher, C.M., Jones, D.N.M., Diamond, R. and Neuhaus, D. (1996) *J. Biomol. NMR*, **8**, 292–310.

Folmer, R.H.A., Folkers, P.J.M., Kaan, A., Jonker, A.J., Aelen, J.M.A., Konings, R.N.H. and Hilbers, C.W. (1994) *Eur. J. Biochem.*, **224**, 663–676.

Folmer, R.H.A., Hilbers, C.W., Konings, R.N.H. and Hallenga, K. (1995a) *J. Biomol. NMR*, **5**, 427–432.

Folmer, R.H.A., Nilges, M., Konings, R.N.H. and Hilbers, C.W. (1995b) *EMBO J.*, **14**, 4132–4142.

Grzesiek, S., Vuister, G.W. and Bax, A. (1993) *J. Biomol. NMR*, **3**, 487–493.

Grzesiek, S., Kuboniwa, H., Hinck, A.P. and Bax, A. (1995) *J. Am. Chem. Soc.*, **117**, 5312–5315.

Güntert, P., Braun, W., Billeter, M. and Wüthrich, K. (1989) *J. Am. Chem. Soc.*, **111**, 3997–4004.

Güntert, P., Braun, W. and Wüthrich, K. (1991) *J. Mol. Biol.*, **217**, 517–530.

Habazettl, J., Cieslar, C., Oschkinat, H. and Holak, T.A. (1990) *FEBS Lett.*, **268**, 141–145.

Havel, T.F. (1991) *Prog. Biophys. Mol. Biol.*, **56**, 43–78.

Holak, T.A., Nilges, M. and Oschkinat, H. (1989) *FEBS Lett.*, **242**, 649–654.

Hyberts, S.G., Märki, W. and Wagner, G. (1987) *Eur. J. Biochem.*, **164**, 625–635.

Ikura, M., Kay, L.E., Tschudin, R. and Bax, A. (1990) *J. Magn. Reson.*, **86**, 204–209.

Kay, L.E., Keifer, P. and Saarinen, T. (1992) *J. Am. Chem. Soc.*, **114**, 10663–10665.

Koning, T.M.G., Boelens, R. and Kaptein, R. (1990) *J. Magn. Reson.*, **90**, 111–123.

Konnert, J.H. and Hendrickson, W.A. (1980) *Acta Crystallogr.*, **A36**, 344–350.

Kraulis, P.J. (1991) *J. Appl. Crystallogr.*, **24**, 946–950.

Nilges, M., Gronenborn, A.M., Brünger, A.T. and Clore, G.M. (1988) *Protein Eng.*, **2**, 27–38.

Nilges, M., Clore, G.M. and Gronenborn, A.M. (1990) *Biopolymers*, **29**, 813–822.

Nilges, M., Kuszewski, J. and Brünger, A.T. (1991) In *Computational Aspects of the Study of Biological Macromolecules by Nuclear Magnetic Resonance Spectroscopy* (Eds., Hoch, J.C., Poulsen, F.M. and Redfield, C.), Plenum, New York, NY, U.S.A., pp. 451–455.

Nilges, M. (1993) *Proteins*, **17**, 297–309.

Palmer, A.G., Cavanagh, J., Wright, P.E. and Rance, M. (1991) *J. Magn. Reson.*, **93**, 151–170.

Shaka, A.J., Barker, P.B. and Freeman, R. (1985) *J. Magn. Reson.*, **64**, 547–553.

Silver, M.S., Joseph, R.I. and Hoult, D.I. (1984) *J. Magn. Reson.*, **59**, 347–353.

Tropp, J. (1980) *J. Chem. Phys.*, **72**, 6035–6043.

Vuister, G.W., Wang, A.C. and Bax, A. (1993) *J. Am. Chem. Soc.*, **115**, 5334–5335.

Weber, P.L., Morrison, R. and Hare, D. (1988) *J. Mol. Biol.*, **204**, 483–487.

Williamson, M.P. and Madison, V.S. (1990) *Biochemistry*, **29**, 2895–2905.

Wüthrich, K., Billeter, M. and Braun, W. (1983) *J. Mol. Biol.*, **169**, 949–961.

Zhao, D. and Jardetzky, O. (1994) *J. Mol. Biol.*, **239**, 601–607.

Zuiderweg, E.R.P., Boelens, R. and Kaptein, R. (1985) *Biopolymers*, **24**, 601–611.