doi:10.1006/jmbi.2000.3842 available online at http://www.idealibrary.com on IDEAL® J. Mol. Biol. (2000) 300, 197–212

JMB

Global Folds of Proteins with Low Densities of NOEs Using Residual Dipolar Couplings: Application to the 370-Residue Maltodextrin-binding Protein

Geoffrey A. Mueller[1,2], W.Y. Choy[1,2], Daiwen Yang[1], Julie D. Forman-Kay[2], Ronald A. Venters[3] and Lewis E. Kay[1]*

[1]*Protein Engineering Network Centers of Excellence and Departments of Medical Genetics and Microbiology Biochemistry, and Chemistry University of Toronto, Toronto Ontario, Canada M5S 1A8*

[2]*Structural Biology and Biochemistry, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario Canada M5G 1X8*

[3]*Duke University Medical Center, Box 3711, Duke University, Durham NC 27710, USA*

**Corresponding author*

The global fold of maltose-binding protein in complex with the substrate β-cyclodextrin was determined by solution NMR methods. The two-domain protein is comprised of a single polypeptide chain of 370 residues, with a molecular mass of 42 kDa. Distance information in the form of H^N-H^N, H^N-CH_3 and CH_3-CH_3 NOEs was recorded on ^{15}N, ^2H and ^{15}N, ^{13}C, ^2H-labeled proteins with methyl protonation in Val, Leu, and Ile ($C^{\delta 1}$ only) residues. Distances to methyl protons, critical for the structure determination, comprised 77 % of the long-range restraints. Initial structures were calculated on the basis of 1943 NOEs, 48 hydrogen bond and 555 dihedral angle restraints. A global pair-wise backbone rmsd of 5.5 Å was obtained for these initial structures with rmsd values for the N and C domains of 2.4 and 3.8 Å, respectively. Direct refinement against one-bond $^1H^N$-^{15}N, $^{13}C^\alpha$-^{13}CO, ^{15}N-^{13}CO, two-bond $^1H^N$-^{13}CO and three-bond $^1H^N$-$^{13}C^\alpha$ dipolar couplings resulted in structures with large numbers of dipolar restraint violations. As an alternative to direct refinement against measured dipolar couplings we have developed an approach where discrete orientations are calculated for each peptide plane on the basis of the dipolar couplings described above. The orientation which best matches that in initial NMR structures calculated from NOE and dihedral angle restraints exclusively is used to refine further the structures using a new module written for CNS. Modeling studies from four different proteins with diverse structural motifs establishes the utility of the methodology. When applied to experimental data recorded on MBP the precision of the family of structures generated improves from 5.5 to 2.2 Å, while the rmsd with respect to the X-ray structure (1dmb) is reduced from 5.1 to 3.3 Å.

Keywords: dipolar couplings; protein structure; labeling; maltodextrin binding protein; protein domains

Introduction

Recent methodological developments in solution state NMR spectroscopy of proteins have dramatically reduced the size limitations that have traditionally prohibited detailed NMR studies of many biologically important molecules (Bax, 1994; Wider & Wüthrich, 1999). Advances such as triple resonance NMR (Bax, 1994), multidimensional spectroscopy (Bax, 1994), deuteration (Farmer & Venters, 1998; Gardner & Kay, 1998) and more recently TROSY (Pervushin *et al.*, 1997, 1998) have led to significant gains in spectral sensitivity and resolution. These advances have, in turn, facilitated structural studies of proteins in the 20-30 kDa molecular mass range.

Despite these important advances, the number of structures determined by NMR techniques of proteins with molecular mass in excess of approximately 30 kDa is small. While chemical shift assignment of backbone resonances is now possible

Abbreviations used: MBP, maltose-binding protein; CT, constant time; TAD, torsion angle dynamics; NOESY, nuclear Overhauser enhancement spectroscopy; C′, carbonyl carbon.

E-mail address of the corresponding author: kay@pound.med.utoronto.ca

0022-2836/00/010197–16 $35.00/0

for many high molecular mass proteins using ^{2}H, ^{15}N, ^{13}C spectroscopy (Farmer & Venters, 1998; Gardner & Kay, 1998), subsequent steps in the structure determination process continue to be rate limiting. For example, overlap in ^{1}H-^{13}C correlation spectra of large proteins complicates the assignment of NOEs to unique sites in the molecule. Moreover, correct assignment of NOEs presupposes that chemical shifts are available for all or most of the side-chain positions in the first place, which is often difficult in the case of large proteins. With these limitations in mind, our laboratory has developed a labeling strategy where ^{15}N, ^{13}C, highly deuterated proteins are produced with protonation at methyl groups of Val, Leu and Ile ($C^{\delta 1}$ only) residues (Goto *et al*., 1999). This approach facilitates chemical shift assignment of backbone as well as methyl side-chain positions and allows distance restraints in the form of H^N-H^N, H^N-CH_3 and CH_3-CH_3 NOEs to be recorded. Because methyl groups are typically located in the core of the protein, these NOEs link residues that are often distal in primary sequence, providing key constraints for structure determination (Gardner *et al*., 1997; Rosen *et al*., 1996).

An additional important advance has been the emergence of the use of dipolar couplings as probes of molecular structure (Tjandra *et al*., 1997; Tolman *et al*., 1995). These couplings are measured on molecules with a small amount of residual alignment provided by oriented particles such as phage (Clore *et al*., 1998c; Hansen *et al*., 1998) or bicelles (Tjandra & Bax, 1997). Studies by Bax and co-workers (Tjandra & Bax, 1997) and Tolman *et al*. (1995) have demonstrated the utility of dipolar couplings for structure determination and, when a density of NOEs of the order of ten per residue or greater is available, direct refinement against measured dipolar couplings has resulted in improved structures (Bewley *et al*., 1998; Cai *et al*., 1998). Very recently, Delaglio *et al*. (2000) used molecular fragment replacement in concert with dipolar couplings to fold ubiquitin without using any NOE restraints. In cases where dipolar couplings are available for all or most residues, this approach appears very promising.

Maltose-binding protein, MBP, is a 370-residue molecule comprised of a single polypeptide chain which binds a variety of sugars. Detailed X-ray diffraction studies of the protein have established that the two domains of the molecule, connected by two β-strands and an α-helix, change in relative orientation in a manner dependent on the bound sugar (Sharff *et al*., 1993, 1992). The existence of an open conformation (ligand free) and a closed conformer (maltose bound) is crucial for MBP's role in the signal transduction cascade that regulates both maltodextrin uptake and chemotaxis (Sharff *et al*., 1993; Spurlino *et al*., 1991). For example, only the closed form of the protein is able to bind to the chemoreceptor Tar protein, necessary for maltose taxis (Zhang *et al*., 1999). MBP has been the focus of interest in our laboratory for a number of reasons. First, the protein is relatively large by NMR standards and thus serves as a test case for the development and application of new structural methodology. Second, the large numbers of crystal contacts in the β-cyclodextrin form of the molecule (Sharff *et al*., 1993; Spurlino *et al*., 1991) suggest that important differences in domain orientation may exist between solution and crystal environments, warranting further studies of this molecule. In this regard we have recently determined solution conformations of MBP in complex with β-cyclodextrin by using dipolar coupling data to adjust the relative orientation of domains in X-ray structures of MBP. The solution state conformation of the MBP/β-cyclodextrin complex obtained using this hybrid approach is significantly different than its X-ray-derived counterpart with the solution structure related to the crystal form (1 dmb) *via* an 11° domain closure (Skrynnikov *et al*., 2000).

Here, we describe the global fold of MBP in complex with the cyclic heptasacharride, β-cyclodextrin (42 kDa). Solution structures were obtained on the basis of 1943 NOEs, 48 hydrogen bonding, 555 dihedral angle and 940 dipolar coupling based restraints derived from an ^{15}N, ^{13}C, ^{2}H, Val, Leu, Ile (δ1 only) methyl protonated sample. A new protocol for the inclusion of dipolar coupling restraints is described which, for large molecules with limited numbers of NOEs per residue, has significantly better convergence properties than previous methods based on direct refinement strategies. All of the software discussed below is available from the authors upon request.

NOE analysis of MBP

Preliminary solution structures of the MBP/β-cyclodextrin complex were calculated on the basis of (ϕ,ψ) dihedral retraints provided by chemical shifts using the program TALOS (Cornilescu *et al*., 1999) and distance restraints measured from NOESY spectra recorded on ^{15}N,^{2}H or ^{15}N, ^{13}C, ^{2}H, Val, Leu, Ile (δ1 only) methyl protonated samples of the protein. Two 4D NOE experiments, the 4D ^{15}N, ^{15}N-edited NOESY (Grzesiek *et al*., 1995; Venters *et al*., 1995) and the 4D ^{13}C, ^{15}N-edited NOESY (Muhandiram *et al*., 1993), were recorded with mixing times of 175 ms. Figure 1(a) and (b) show planes from each of the two 4D data sets selected at the ^{15}N, H^N chemical shifts of Ile266, located in the middle strand of a three-stranded β-sheet. Correlations to residues on adjacent strands of the sheet, Ile60, Phe61 and Tyr106 are illustrated in Figure 1(a), along with an NOE to Ala264. Proximal methyl groups (Figure 1(b)) include Leu76, in a loop extending from Leu75 to Asp82, as well as Leu280 at the C-terminal end of an α-helix. As described, NOEs involving methyl groups are particularly significant for structure determination by NMR in that they very often bridge elements of diverse secondary structure. This example, involving methyl-H^N NOEs which serve to orient a β-sheet, an α-helix and a loop

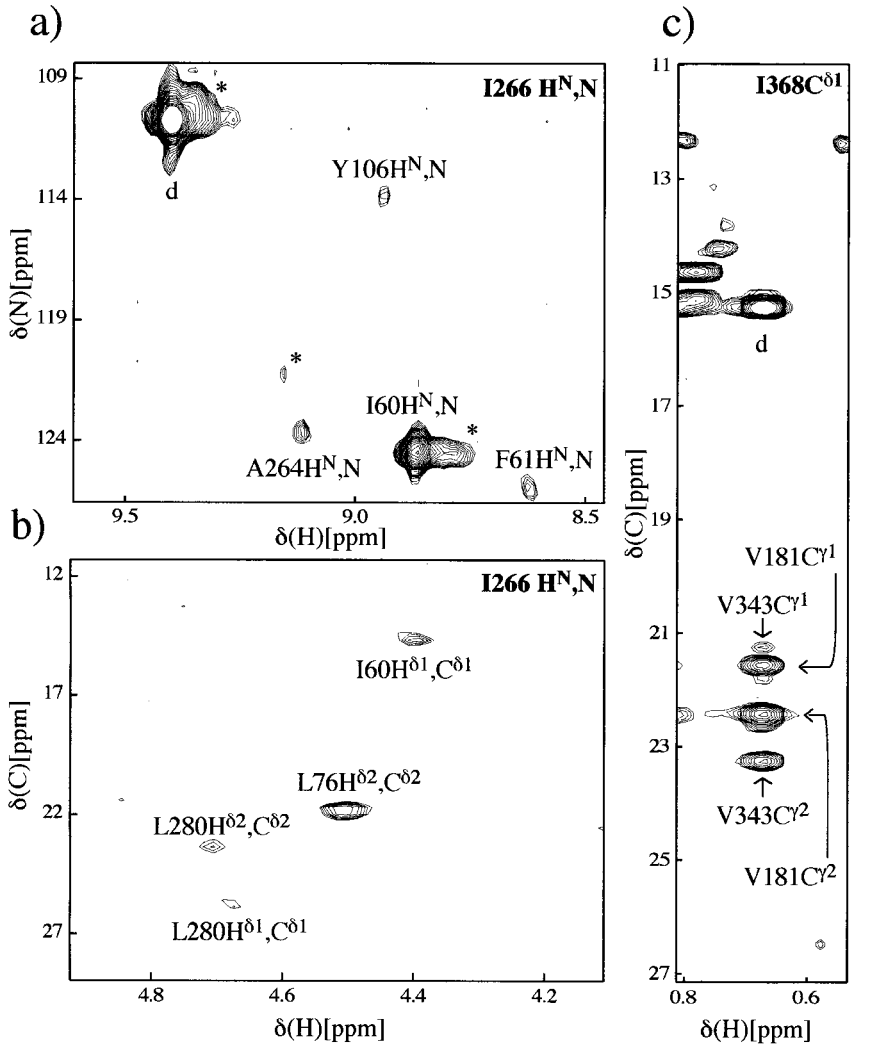

Figure 1. Planes from (a) 4D ^{15}N, ^{15}N-edited NOESY (Grzesiek *et al.*, 1995; Venters *et al.*, 1995), (b) 4D ^{13}C, ^{15}N-edited NOESY (Muhandiram *et al.*, 1993) and (c) 3D CT-^{13}C-edited NOESY (Zwahlen *et al.*, 1998a) data sets recorded on the MBP/β-cyclodextrin complex. Cross-peaks corresponding to NOEs between Ile266 H^N (a) and (b) or Ile368 $H^{\delta 1}$ (c) and proximal protons are labeled in each slice. Diagonal peaks are indicated by d, with cross-peaks that are more intense in different slices indicated by *. (a) and (b) are extracted from 4D data sets at (9.41 ppm, 110.68 ppm) and (c) is extracted from the 3D ^{13}C-edited NOESY at 15.28 ppm.

serves as a good illustration of the importance of this class of restraints.

In addition to the 4D data sets described above, two ^{13}C-edited 3D NOESY spectra (175 ms mixing times) were recorded that exploit the narrow methyl group carbon linewidths in proteins. The CT-^{13}C-edited NOESY (Zwahlen *et al.*, 1998a) provides correlations of the form $(\omega^i_{Cm},\omega^j_{Cm},\omega^j_{Hm})$, where methyls i and j are proximal. High resolution is provided in this experiment by recording the frequency ω^i_{Cm} in constant-time mode (Santoro & King, 1992; Vuister & Bax, 1992). In addition, as described by Zwahlen *et al.* (1998), correlations linking Cm^i with neighboring ^{15}N, H^N spin-pairs can also be observed in this experiment. Figure 1(c) shows NOE correlations from Val343 and Val181 to the δ-methyl group of Ile368, located near the end of the last helix in the C-terminal domain. These NOE correlations are particularly important in defining the orientation of this helix with respect to the rest of the molecule, since very few additional long range restraints have been observed. Because ''symmetry-related correlations'', $(\omega^i_{Cm},\omega^j_{Cm},\omega^j_{Hm})$ and $(\omega^j_{Cm},\omega^i_{Cm},\omega^i_{Hm})$ are obtained in this experiment, the assignment of cross-peaks to specific methyl pairs in the protein is often straight-forward.

A second 3D experiment, (HM)CMCB(CMHM)-NOESY (Zwahlen *et al.*, 1998b) was also recorded, providing NOE correlations of the form $(\omega^i_{Cm},\omega^i_{CB},\omega^j_{Hm})$, where CB is the carbon atom adjacent to methyl group carbon Cm^i. Because only a single chemical shift is obtained which identifies the destination site of magnetization, ω^j_{Hm}, this experiment is best used in combination with the CT-^{13}C edited NOESY described above. In a significant number of cases correlations from the (HM)CMCB(CMHM)-NOESY allowed unambiguous assignment of NOEs from the ^{13}C-edited NOESY that otherwise could not be uniquely resolved.

Initially, NOESY spectra were analyzed in a very conservative manner by assigning only symmetry-related pairs of NOEs from the 4D ^{15}N, ^{15}N-edited NOESY and the 3D CT-^{13}C edited NOESY data sets. In the case of the 4D ^{13}C, ^{15}N-edited NOESY where symmetry related correlations are not obtained, weak cross-peaks were only included in

preliminary structure calculations if they could also be found in the CT-^{13}C-edited NOESY. After the first round of structure calculations, the initial set of structures were used to help resolve ambiguities in NOE assignment. Despite the fact that the structures were of low quality at this stage (>8 Å pair-wise rmsd) they were nevertheless helpful in further assignment, since NOE ambiguities were in many cases limited to no more than four possibilities. Typically only one of the possibilities involved pairs of protons located within 7-8 Å in the preliminary set of structures, with the other distances usually 30 Å or more. Note that the small number of possibilities results directly from the labeling scheme employed where only methyl and amide groups are protonated; in fully protonated molecules the list of potential NOE partners in cases of ambiguous correlations would be more numerous. A second round of structure calculations was subsequently performed and many ambiguous NOEs that could not be assigned on the basis of the preliminary structures could now be assigned. At this stage 555 (ϕ,ψ) dihedral angle and 1943 distance retraints (826 H^N-H^N, 769 CH_3-H^N, 348 CH_3-CH_3) were available. In addition, 48 hydrogen bonding restraints were obtained from a comparison of H^N-^{15}N correlation spectra recorded on a sample of deuterated MBP prior to and after back exchange of protected amide protons, as described by Gardner *et al.* (1998).

Figure 2 provides a summary of the types of NOEs that were assigned from the spectra described above, with the NOEs subdivided into short range ($\leqslant$three residues apart in primary structure) and long range (>three residues) categories. Although significant numbers of long-range NOEs are observed between residues in either the N or C-domains and between the linker regions and one of the two domains, very few (ten) long-range NOEs were assigned between the two domains. The lack of NOEs between domains in multi-domain proteins where each domain is only loosely connected to the other is a significant limitation of structure-based studies involving only short range distance restraints and scalar couplings.

The importance of methyl NOEs is well illustrated in Figure 2(a), showing that 77 % of all long-range restraints derive from methyl correlations. It is interesting that a larger number of long range NOEs were obtained from the N-domain than from the C domain (410 *versus* 382), despite the fact that the C domain has 40 more residues. Figure 2(b) illustrates the fraction of possible NOEs which were actually observed, with the total possible estimated from those expected on the basis of the X-ray structure of MBP loaded with β-cyclodextrin, 1dmb (Sharff *et al.*, 1993). The number of H^N-H^N NOEs decreases steeply with distance, while a considerable fraction of the expected methyl-methyl correlations out to 8 Å are observed. In this regard it is noteworthy that under the conditions used to record the NOE data,

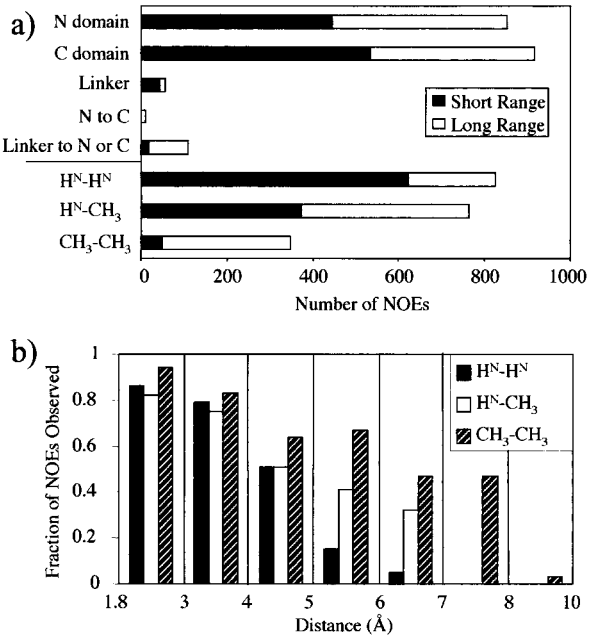

Figure 2. (a) Distribution of NOEs measured in MBP classified according to location in the protein and NOE type. The five categories corresponding to the location of NOEs include: N domain (residues 1 to 109 and 264 to 309), C domain (residues 114 to 258 and 316 to 370), linker (residues 110 to 113, 258 to 263, and 310 to 315), N to C (from the N domain to the C domain), linker to N or C (from the linker to the N or C domain). NOEs are divided according to H^N-H^N (826), H^N-CH_3 (769) and CH_3-CH_3 (348) and are classified as either long range (□, $>i,i+3$) or short range (■). (b) Fraction of possible NOEs (H^N-H^N, H^N-CH_3 and CH_3-CH_3) that were observed in MBP NOESY data sets *versus* distance in the X-ray structure, 1dmb (Sharff *et al.*, 1993).

pH 7.2, 37 °C, there is significant exchange with water, leading to attenuation of H^N-H^N NOEs.

It is of interest to compare the observed fraction of possible NOEs in deuterated MBP with methyl protonation *via* α-ketobutyrate and α-ketoisovalerate (Goto *et al.*, 1999) with the fraction of correlations obtained for the methyl protonated, ^{2}H-labeled C-terminal SH2 domain from phospholipase C$^{\gamma 1}$, PLCC SH2 (Gardner *et al.*, 1997). In the case of the SH2 domain, methyl protonation was achieved using ^{13}C protonated pyruvate as the sole carbon source (Rosen *et al.*, 1996). As described, this procedure produces significant levels of CH_2D, CHD_2 and CD_3 methyl isotopomers, which seriously degrade both spectral resolution and sensitivity. For example, in methyl correlation spectra of the PLCC SH2 domain, only 10 % of the expected methyl-methyl correlations between 6-7 Å were observed (Gardner *et al.*, 1997), compared to approximately 50 % for MBP.

Figure 3 presents the ten lowest energy solution structures of MBP calculated on the basis of the dihedral angle, hydrogen bonding and NOE restraints and results from the structure calculations are summarized in Table 1. The N and

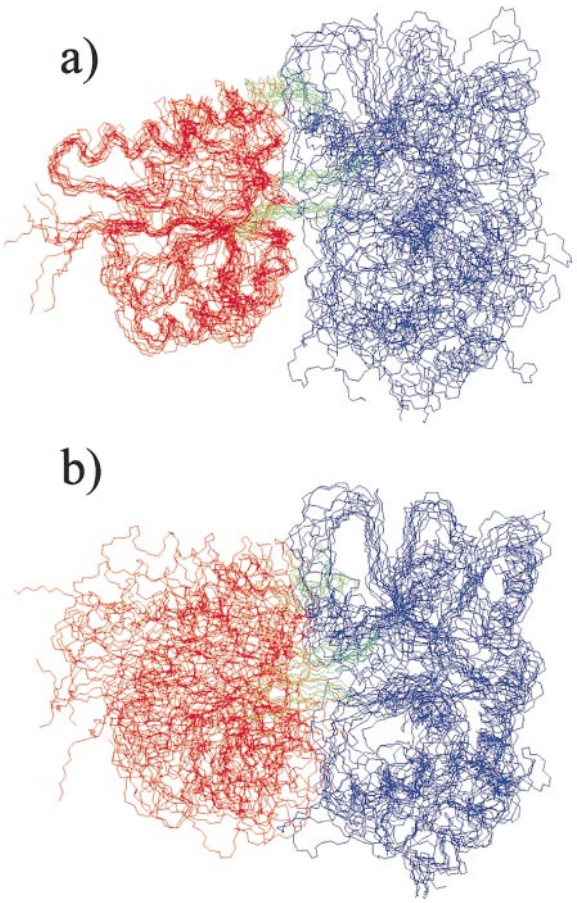

Figure 3. Solution structures of MBP/β-cyclodextrin calculated on the basis of 1943 NOE, 555 dihedral angle and 48 hydrogen bond restraints. Backbone traces of the ten lowest energy structures are displayed and superimposed by fitting to either (a) the N domain, alignment over residues 6 to 109 and 264 to 309 or (b) the C domain, alignment over residues 114 to 258 and 316 to 370. The domains are colored red (N domain, residues 1 to 109 and 264 to 309) and blue (C domain, residues 114 to 258 and 316 to 370) with the linker region in green (residues 110 to 113, 258 to 263, and 310 to 315).

C domains are aligned in Figure 3(a) and (b), respectively, with pair-wise rmsd values of 2.4 Å (N-domain) and 3.8 Å (C domain). It is clear that while the definition of the individual domains is reasonable, their relative orientation is somewhat ill defined, and a global pairwise rmsd value of 5.5 Å is thus obtained. Nevertheless, even at this level of precision it is possible to establish the overall topology of the protein as well as its composition of two domains connected *via* helical and β-sheet linkers. In addition, the relative orientation of individual elements of secondary structure within each domain can be ascertained. Although of a preliminary nature, these structures can subsequently be used in conjunction with dipolar couplings to generate significantly better structures, as described below.

Table 1. Statistics for the ten final structures of the MBP/β-cyclodextrin complex

	No dipolar couplings[a]	With dipolar couplings[b]
Average pairwise rmsd (Å)[c]		
Global	5.5 ± 1.4	2.2 ± 0.3
N domain	2.4 ± 0.3	1.7 ± 0.2
C domain	3.8 ± 1.1	1.8 ± 0.2
Average rmsd to 1dmb[d]		
Global	5.1 ± 0.7	3.3 ± 0.1
N domain	3.1 ± 0.2	2.7 ± 0.2
C domain	3.8 ± 0.7	2.8 ± 0.1
ϕ/ψ space: residues[e]		
Most favored region (%)	71.2 ± 2.0	74.7 ± 1.9
Additionally allowed region (%)	23.0 ± 1.8	21.3 ± 1.9
Generously allowed region (%)	3.7 ± 1.0	3.0 ± 0.7
Disallowed region (%)	1.4 ± 0.4	1.1 ± 0.4
rmsd from covalent geometry		
Bonds (Å)[f]	0.0001 ± 0.0000	0.0030 ± 0.0001
Angles (deg.)	0.2746 ± 0.0041	0.4514 ± 0.0101
Impropers	0.1255 ± 0.0064	0.4668 ± 0.0191
rmsd from experimental restraints		
NOEs (Å)	0.0044 ± 0.0032	0.0117 ± 0.0025
(ϕ,ψ) Dihedral angles (deg.)	0.3167 ± 0.0409	1.1126 ± 0.0716
Dipolar couplings (deg.)		1.7471 ± 2.1100

[a] Structures calculated on the basis of 1943 NOE, 555 dihedral angle and 48 hydrogen bond restraints.
[b] Structures calculated using the restraints in[a], above, and dipolar coupling-based restraints for 188 residues.
[c] Calculated with MOLMOL (Konradi *et al.*, 1996). In the rmsd calculations the following residues were used: Global: Lys6-Ile235, Asn241-Lys370 (Note that most of the assignments for residues 229 to 239 are not available); N-domain: Lys6-Ala109, Ala264-Glu309; C-domain: Ser114-Ile235, Asn241-Phe258, Arg316-Lys370.
[d] 1dmb, Sharff *et al.* (1993).
[e] Calculated with PROCHECK-NMR (Laskowski *et al.*, 1998).
[f] Evaluated by CNS (Brünger *et al.*, 1998).

Incorporating dipolar couplings into structure calculations

In addition to the restraints described above, supplementary restraints derived from residual dipolar coupling data can be included in structure calculation protocols. Direct refinement of the orientation of bond vectors using residual dipolar couplings has been shown to improve the accuracy and precision of structures when used in conjunction with nearly complete sets of NOE, coupling constant and chemical shift data (Clore *et al.*, 1998b). Furthermore, the use of dipolar couplings has led to improvements in the accuracy and precision of the structures of several small proteins with well-defined topologies determined from a limited set of NOEs (Clore *et al.*, 1999).

In the case of MBP we have used TROSY-based HNCO experiments to measure 280 ^{15}N-$^{1}H^{N}$, 262 ^{15}N-$^{13}C'$ (carbonyl) and 276 $^{13}C^{\alpha}$-$^{13}C'$ one-bond, 262 two-bond $^{13}C'$-$^{1}H^{N}$ and 276 three-bond $^{13}C^{\alpha}$-$^{1}H^{N}$ dipolar couplings (Yang *et al.*, 1999). A CNS (Brünger *et al.*, 1998) protocol which utilized the limited NOE, dihedral angle and hydrogen bond restraints described above in conjunction with

direct refinement against the dipolar couplings listed above produced structures with large numbers of violations specifically in the dipolar restraints (although the same protocol was much more successful in the case of ubiquitin). This is likely due to the fact that for each individual dipolar coupling there is an infinite number of orientations consistent with the data, leading to an extremely complex energy surface over which refinement occurs. In this sense it is far easier to alter the orientation of individual dipole vectors in an attempt to satisfy restraints than it is to converge to a global minimum in the energy landscape by changing the overall protein fold. Prestegard and co-workers have also observed related convergence problems in their study of a two-domain lectin molecule using dipolar couplings and a limited set of NOE restraints (Fischer *et al.*, 1999). In this case they were unable to determine the relative orientation of the domains unless a large number of distance restraints from the X-ray structure of the protein was included. Based on these observations, it is clear that a new approach is needed if dipolar couplings are to be used as restraints in the calculation of solution structures of large proteins where there is a limited quantity of other structural data. The approach presented here involves the use of dipolar couplings to orient peptide planes as discrete structural units.

The dipolar coupling, δ_{AB}, between two spin 1/2 nuclei A and B is given by:

$$\delta_{AB} = \delta^{o}_{AB} A_a \left[(3\cos^2\theta_{AB} - 1) + \frac{3}{2} R \sin^2\theta_{AB} \cos 2\phi_{AB} \right] \quad (1)$$

where θ_{AB} and ϕ_{AB} are the polar angles that describe the orientation of the vector connecting A and B with respect to the principal alignment frame, A_a and R are the axial and the rhombic components of the molecular alignment tensor, respectively, and

$$\delta^{o}_{AB} = -\left(\frac{1}{2\pi}\right)\left(\frac{\mu_o \hbar}{4\pi}\right) \hbar \gamma_A \gamma_B \langle r^{-3}_{AB} \rangle$$

is the dipolar interaction constant (Tjandra & Bax, 1997). In principle, there are five degrees of freedom associated with the choice of an alignment frame including A_a, R and three Euler angles (α,β,γ) describing the transformation from the coordinate frame of the molecule to the alignment frame. The values of A_a and R can, however, be estimated from the distribution of measured dipolar couplings for the entire protein (Clore *et al.*, 1998a) leaving (α,β,γ) to be determined. Accordingly, at least three dipolar coupling measurements are required within a rigid system, such as the peptide plane, in order to solve for the transformation. In general, however, we prefer to consider residues only if five dipolar couplings are available in order to minimize the effects of experimental error.

In MBP the five dipolar couplings listed above have been measured for 240 of the 370 peptide planes in the molecule (Yang *et al.*, 1999). These dipolar couplings are illustrated using red arrows superimposed on the peptide plane drawn in Figure 4(a). The transformation from an initial set of molecular coordinates into the fragment (peptide plane) alignment frame is determined by a grid search in the space of (α,β,γ) which minimizes the difference between the measured couplings and those predicted from a trial alignment frame *via*:

$$\chi^2 = \sum_{j=1}^{5} (\delta^{j}_{\text{Predicted}} - \delta^{j}_{\text{Measured}})^2 \quad (2)$$

where the sum is over the five measured dipolar couplings at a particular site.

Unfortunately, dipolar coupling data alone does not define a unique orientation of the peptide within its alignment frame. In fact, there are at least eight possible orientations that are consistent with the dipolar data. Rotation of the plane by 180° about any of the three alignment axes gives four possible structures. Further rotation of each of these structures by an additional 180° about the plane normal leads to four more structures that satisfy the dipolar coupling data. Starting from a peptide plane lying in the X-Y plane of a coordinate axis system it can be shown that if (α,β,γ) are Euler angles that transform the peptide into its alignment frame then equally valid transformations are given by (α,β,180° + γ), (180° + α,180° − β,180° − γ), (180° + α,180° − β,360° − γ), (180° + α,β,γ), (180° + α,β,180° + γ), (α,180° − β,180° − γ) and (α,180° − β,360° − γ). Illustrated in Figure 4(c) are the eight different orientations of the peptide plane spanning residues Phe149 and Asn150 in MBP consistent with the measured dipolar coupling data. In order to choose the proper orientation we use the preliminary NMR structures calculated exclusively from the NOE, dihedral angle and hydrogen bonding restraints described above. This is accomplished by calculating an average structure from the lowest energy preliminary structures and determining the molecular alignment frame using all of the measured dipolar coupling values in concert with this structure. Subsequently, the average structure is rotated into its alignment frame and the orientation of a given peptide plane extracted from the average structure is compared with the eight possible peptide orientations obtained by considering only dipolar coupling data. Assuming that the structure is static the orientation of the peptide in the peptide alignment frame (Figure 4(c)) and in the molecular alignment frame (Figure 4(b)) is the same and the correct orientation in Figure 4 can therefore be obtained.

The method of incorporating dipolar couplings into structure calculations described above is illustrated using a flow chart in Figure 5. In steps 1 and

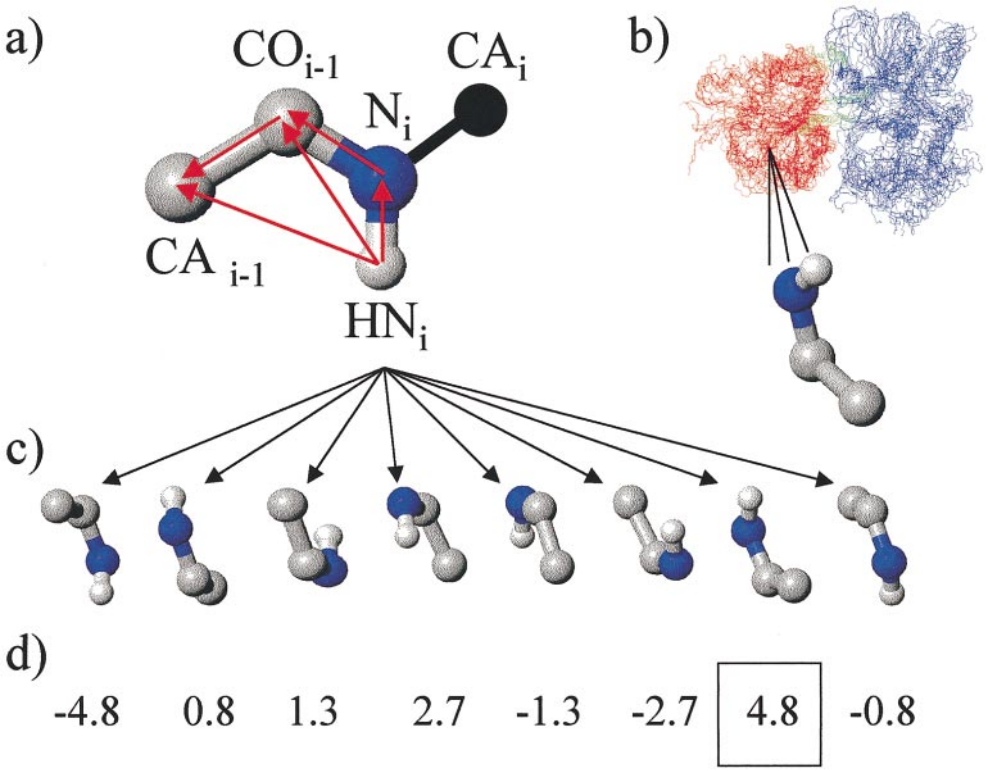

Figure 4. Summary of the protocol used to choose between the eight possible orientations of peptide planes established on the basis of dipolar coupling data. Dipolar couplings are measured from the five dipole vectors illustrated in (a), resulting in eight possible peptide plane orientations (example shown for the plane bridging residues Phe149 and Asn150 of MBP). The eight orientations (c) are compared with the corresponding plane from the average structure (in the alignment frame) derived on the basis of NOE, dihedral and hydrogen bonding data (b) in order to select a set of restraints. The sum of the dot products of the five dipole vectors from each plane with the corresponding vectors from the average structure are shown in (d), illustrating that structure 7 is the best match.

2 NOE and dihedral angle restraints along with dipolar couplings are collected. Subsequently, structures based solely on the NOE, hydrogen bonding and dihedral angle data are calculated as described above (step 3) and an average structure obtained from a set of lowest energy structures with no restraint violations (step 4). A molecular alignment frame (magnitude and orientation) is calculated using all of the dipolar coupling data and the average structure obtained in step 4 by minimizing the difference between the dipolar couplings predicted from the structure and those measured (step 5). The average structure is subsequently rotated into the alignment frame (step 6). Note that although there are four copies of the structure in this alignment frame which satisfy the dipolar couplings (Skrynnikov *et al.*, 2000) any one can be chosen since the choice does not affect the relative orientation of planes selected using the procedure shown in Figure 4.

Figure 4(b) illustrates the ensemble of the ten lowest energy structures of MBP calculated on the basis of NOE, dihedral angle and hydrogen bond restraints only. All of the structures are oriented in the molecular alignment frame of the average structure with the peptide plane spanning residues Phe149-Asn150 illustrated. Independently, in steps 7 and 8, the dipolar coupling data are used to predict the eight orientations for each peptide plane that satisfy the dipolar coupling data, illustrated for the Phe149-Asn150 plane in Figure 4(c). At this stage, the eight predicted peptide plane orientations are compared with the orientation determined from the average structure (step 9). From a simple visual inspection of Figure 4(b) and (c) it is apparent that the seventh orientation is the best match in this particular case. More generally, the level of agreement between the orientation of a given peptide plane in the average structure and the eight orientations generated from dipolar couplings exclusively can be obtained by taking the sum of the dot products, S_k, of the normalized vectors spanning the one-bond N-H^N, N-C′, C′-C^α, the two-bond C′-H^N and the three-bond C^α-H^N coupled atoms in the average structure (vectors A_{A-B}) with the corresponding vectors in each of the k ($1 \leqslant k \leqslant 8$) predicted orientations (vectors P^k_{A-B}). The score, S_k,

$$S_k = \sum_{A,B} \vec{A}_{A-B} \times \vec{P}^k_{A-B} \qquad (3)$$

for a given peptide plane k ranges from −5 to 5. Values of S_k are displayed below the corresponding

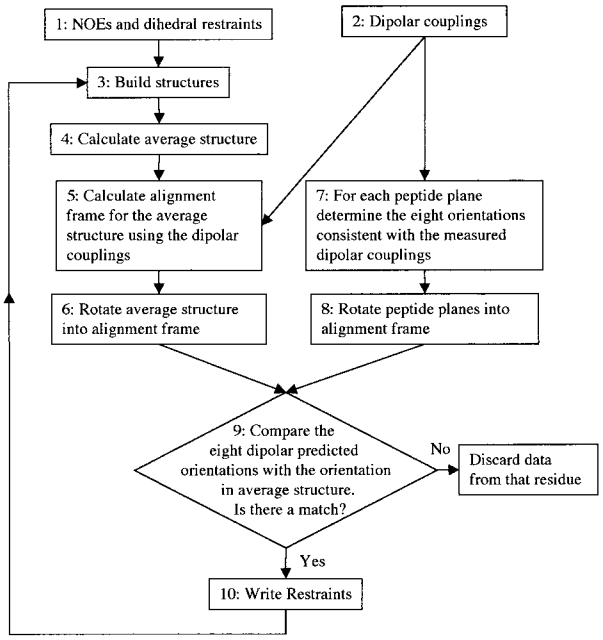

Figure 5. Flow chart illustrating how dipolar coupling-based restraints are incorporated into the structure refinement scheme. The numbers next to the boxes refer to specific steps which are described in the text.

peptide orientations in Figure 4(d), with the best match, $S_{k,max} = S_7 = 4.8$. This procedure is repeated for each peptide plane for which five dipolar couplings are available and in each case the plane corresponding to $S_{k,max}$ is used to obtain orientational restraints. It is noteworthy that restraints are not employed in cases where $S_{k,max} < 3.5$ (step 9 in Figure 5). The cut-off value of 3.5 was chosen as a good compromise between minimizing the number of incorrect restraints and ensuring that a high number of restraints are available for structure refinement. Poor agreement between a peptide plane orientation predicted from NOE/dihedral data and an orientation obtained on the basis of dipolar couplings may be the result of a number of factors. These include errors in measured dipolar couplings, errors in the values of A_a and R estimated on the basis of the distribution of couplings, poorly defined regions in NOE generated structures, deviations from assumed ideal peptide plane geometry and internal dynamics at a peptide site. There is an additional source of error. In the above discussion we noted that in the general case there are eight possible copies of a peptide plane that are consistent with the five dipolar couplings that are measured. However, if the z-axis of the peptide alignment frame lies in the peptide plane (with the plane making an angle of ϕ_a with respect to the x-axis of the alignment frame, for example) there is an additional twofold degeneracy. In this case all the dipolar vectors are oriented with $\phi = \phi_a$ or $\phi = \phi_a + \pi$ in the alignment frame. From equation (1) it is clear that if the plane were rotated so that the angle with respect to the x-axis of the alignment frame becomes $2\pi - \phi_a$ the dipolar couplings would still be satisfied. In principle, this extra twofold degeneracy can be taken into account, although in the present implementation we have not done so.

After establishing the best peptide plane orientation, restraints are written for subsequent structure refinement utilizing a new CNS module (Brünger *et al.*, 1998) that has been coded for this purpose. In this case, a restraint is comprised of the polar angles (θ,ϕ) defining the orientation of a vector connecting pairs of atoms listed above (red lines in Figure 4(a)) with respect to the peptide alignment frame. For each peptide plane where restraints are to be written, five statements orient each of the five atom pairs whose connecting vectors are defined by the dipolar coupling data (step 10 of Figure 5). The new module for CNS is a straightforward modification of the susceptibility anisotropy module (SANI) described by Clore and co-workers (1998b) and details are given in Materials and Methods. Finally, the new dipolar restraints are added to the NOE and dihedral restraints to build new structures, indicated by the loop from step 10 to 3 in Figure 5. This procedure can be iterated to further refine the structures, although we have not done so here.

It is important to emphasize the differences between direct refinement against dipolar couplings and the procedure developed here. In the case of direct refinement, the orientations of individual bond vectors are changed to satisfy dipolar couplings; in principle, for any given dipolar coupling value there are sets of solutions, which in the case of an axially symmetric alignment frame are described by a pair of cones. In our approach the orientation of each of the dipolar vectors in a peptide plane is determined from five dipolar couplings and the preliminary NOE-based structure. The orientations of these planar vectors are self-consistent and are restrained (directly) to well defined values, as described above.

Simulations using model protein systems

In order to test the general applicability of the approach outlined above, structure calculations were performed using simulated data derived from the X-ray crystal structures of four large protein domains (256 to 347 residues). The proteins used for this study were chosen in order to include a variety of structural motifs as defined by the structural classification of proteins database, SCOP (Murzin *et al.*, 1995). The following protein domains belonging to each of the four structural classifications were used: all β (1bgl (Jacobson *et al.*, 1994), residues 731 to 1023 from chain F of *Escherichia coli* β-galactosidase), all α (1fps (Tarshis *et al.*, 1994), residues 20-367 of avian farnesyl diphosphate synthase), separated α + β (1mua (Tweedy *et al.*, 1993), residues 4 to 260 of human carbonic anhydrase II), and intermixed α/β (1rla (Kanyo *et al.*, 1996), residues 6 to 319 of rat arginase). For

the purpose of calculating rmsd values between structures, the less well defined ''tails'' of 1mua (residues 4 to 25) and 1rla (residues 304 to 319) were not considered. Note that MBP belongs to the mixed α/β category of structural motifs.

NOE and dihedral angle data for preliminary structure calculations were predicted from the crystal structures of each of these proteins. Since these simulations are intended to reflect the situation typically encountered in the NMR study of large proteins, only NOEs that are available from ^{2}H, ^{15}N, ^{13}C methyl-protonated samples were employed. More specifically, these included H^N-H^N, CH_3-CH_3, and H^N-CH_3 NOEs where only backbone amides and methyls of Val, Leu, and Ile ($C^{\delta 1}$ only) residues were included. Lower and upper bounds for the restraints were set as specified for MBP (see Materials and Methods). Typically, not all of the possible NOEs are observable and/or assignable in experimental NMR data. Therefore, restraints from the list of possible NOEs were discarded at random in a type and distance-dependent manner consistent with our experience in the examination of the experimental MBP data (Figure 2); for each protein five different NOE data sets were generated by randomly eliminating potential NOEs in this manner. Dihedral angles were restrained to the crystal structure values $\pm 40^\circ$ only for those residues found in regular secondary structural elements. Preliminary NMR structures were obtained using exactly the same simulated annealing protocol and parameters used in the calculation of initial MBP structures, described in Materials and Methods. A summary of the restraints used in each of the calculations is provided in Table 2A.

Table 2B presents, for each of the four proteins considered, the resulting average pair-wise RMSD values for the ten lowest energy structures based solely on NOE and dihedral angle restraints. The backbone atom precision and accuracy of the preliminary structures are given in the first and second columns, respectively. A range of rmsd values is found, extending from approximately 2.5 to 6 Å, correlating roughly with the type of fold.

Dipolar coupling data were also simulated from the X-ray crystal structures. In all simulations the calculated dipolar couplings were obtained assuming A_a and R values of 0.0017 and 0.26, respectively, very similar to the values calculated from the distribution of dipolar couplings in MBP, and the alignment frame for each molecule was taken (arbitrarily) to coincide with the x,y,z coordinate frame of the X-ray structure (referred to as AF_{PDB}). Random errors, within the precision of the measured values for MBP (Yang *et al.*, 1999), were added to the simulated dipolar couplings according to: $\delta N_i - H_i^N (\pm 0.69)$ Hz, $\delta N_i - C'_{i-1} (\pm 0.18)$ Hz, $\delta C_{i-1}^{\alpha} - C'_{i-1} (\pm 0.75)$ Hz, $\delta H_i^N - C'_{i-1} (\pm 0.64)$ Hz, and $\delta H_i^N - C_{i-1}^{\alpha} (\pm 0.78)$ Hz. In the case of MBP, 69 % of the expected dipolar coupling data for non-Pro residues was obtained, corresponding to 240 of 348 non-Pro residues. With this in mind, 31 % of the calculated dipolar couplings for each of the test proteins were randomly removed from the calculations. It is interesting to note that alignment frames calculated from the X-ray structures of each of the proteins using these simulated dipolar couplings with random errors (referred to as AF_{XTAL}) were in all cases within 1.4° of the corresponding AF_{PDB} (no errors). These errors introduce, therefore, only a very small uncertainty in the calculation of the overall alignment.

Central to our approach for integrating dipolar couplings into the structure calculation process is the establishment of the molecular alignment frame using the preliminary NOE/dihedral angle derived average structures and the dipolar couplings, referred to as AF_{AVG} (Figure 5, step 5). The main source of error in this step derives from the uncertainty in these preliminary structures. For all structural classes listed in Table 2 the z-axes of AF_{XTAL} and AF_{AVG} were within a few degrees (on average less than 5°; worst case of 11°), while the angle between the corresponding x and y-axes was 11° on average, with a worst case of 27°. The poor definition of the x,y-axes relative to z is due to the fact that their position is determined by the magnitude of R, which in the present calculations was set to 0.26; if R is set to 0.5, the largest deviation between the x and y-axis of AF_{AVG} and the corresponding axes in AF_{XTAL} is 17°. Conversely if $R = 0.1$ a maximum difference of 84° is observed for one of the five NOE/dihedral-based structures of 1fps (the least well defined of the four structural classes considered), although the other four alignment frames were within 20° of the crystal-defined x-y positions. Thus, despite the relatively poor accuracy of the starting structures, it is nevertheless possible to establish the orientation of the alignment frame (AF_{AVG}) reasonably well (except for one or two cases for $R = 0.1$). It is not surprising that the accuracy of AF_{AVG} is significantly improved when the average structure from the family of preliminary structures is used rather than any one of the individual structures.

Once the molecular alignment frame has been obtained (Figure 5, step 5), the correct orientation for individual peptide planes aligned on the basis of dipolar data exclusively must be established using the average preliminary structure as a guide (Figure 4(c); Figure 5, step 9). In contrast to determining the molecular alignment frame where all of the dipolar coupling data is employed, the position of individual peptide plane alignment frames (and hence the orientation of each peptide plane) is based on a maximum of five measurements. Therefore, errors in measurement become more critical at this stage. In order to assess the effects of experimental error on the position of the peptide planes, all orientations which predict couplings that are within the error of measurement have been visualized using a program which places the dipole vectors on a sphere. For each peptide eight ''patches'' corresponding to the eight predicted orientations are generally observed, with the spread of vectors

Table 2.
A. *Model protein systems*

Protein class	PDB code[a]	Residues	NOEs restraints	Dihedral restraints	Residue with dipolar restraints
all β	1bgl	293	1162	315	146
all α	1fps	348	1902	471	180
α + β	1mua	256	1290	225	128
α/β	1rla	323	1811	349	164

B. *Influence of dipolar coupling based restraints on precision and accuracy*

		No dipolar couplings		With dipolar couplings		Control 1[f]		Control 2[g]	
Protein	Simulation	Precision[b]	Accuracy[c]	Precision[d]	Accuracy[e]	Precision	Accuracy	Precision	Accuracy
all β	1	3.7	4.1	2.6	3.5	2.6	3.4	1.7	2.3
	2	3.1	3.7	2.3	3.0	2.3	2.7	1.5	2.1
	3	3.6	3.8	2.3	2.9	2.5	2.8	1.5	2.1
	4	3.3	3.5	2.5	3.1	2.6	3.0	1.5	2.3
	5	3.2	3.7	2.4	3.3	2.4	2.9	1.5	2.1
all α	1	6.7	6.2	3.1	3.9	2.8	3.1	2.2	2.5
	2	5.0	5.2	2.7	3.5	2.8	3.4	2.1	2.4
	3	6.1	5.2	2.9	3.8	3.1	3.2	2.1	2.3
	4	4.6	4.5	2.7	3.1	2.6	2.8	2.0	2.3
	5	6.0	5.5	2.6	3.3	2.9	3.0	2.2	2.4
α + β	1	2.5	3.0	1.9	2.5	1.9	2.4	1.3	2.0
	2	2.9	3.1	2.0	2.4	2.0	2.4	1.3	1.9
	3	2.8	3.1	1.8	2.3	1.9	2.2	1.3	1.8
	4	2.8	3.0	1.9	2.5	2.1	2.4	1.3	1.9
	5	2.9	3.2	2.0	2.5	2.1	2.4	1.2	1.9
α/β	1	2.3	2.5	2.0	2.3	1.9	2.2	1.2	1.8
	2	2.6	2.7	1.9	2.3	2.0	2.2	1.3	1.8
	3	2.7	2.8	1.9	2.4	1.9	2.3	1.4	1.8
	4	2.5	2.8	1.9	2.3	1.8	2.2	1.3	1.9
	5	2.3	2.5	1.8	2.3	1.9	2.1	1.3	1.9

[a] 1bgl, Jacobson *et al*. (1994); 1fps, Tarshis *et al*. (1994); 1mua, Tweedy *et al*. (1993); 1rla, Kanyo *et al*. (1996).
[b] Average pair-wise rmsd between the ten lowest energy structures calculated without the addition of dipolar coupling based restraints.
[c] Average rmsd between the X-ray structure and the ten lowest energy structures calculated without the addition of dipolar coupling-based restraints.
[d] Average pair-wise rmsd between the ten lowest energy structures calculated using dipolar coupling based restraints.
[e] Average rmsd between the X-ray structure and the ten lowest energy structures calculated using dipolar coupling based restraints.
[f] Orientational restraints calculated from the X-ray structure substituted for dipolar coupling based restraints.
[g] Orientational restraints calculated from the X-ray structure included for all residues in the protein.

as large as 30° from the mean position in some cases. Therefore, the structures refined herein (described below) were all calculated using errors of 30° in the orientation of each dipole vector.

The experimental errors discussed above affect the position of the peptide planes (Figure 4(c)) and hence can influence which of the eight possible orientations is ultimately selected (see Figures 4 and 5). As described above, selection is achieved by comparing the planes with the corresponding plane from the average preliminary structure. The low resolution of the preliminary structure therefore ultimately affects this choice. A comparison has been made between orientations of peptide planes chosen using the preliminary structure (oriented in AF_{AVG}) with those planes selected on the basis of the crystal structure (oriented in AF_{XTAL}) (step 9 of Figure 5). The orientations selected by the two approaches were compared by summing the dot products of the corresponding five dipolar vectors (Figure 4(a)) in each of the selected planes. If the average angle between dipole vectors based on this sum is within 30°, corresponding to the error bounds used in structure calculations (see below), then the peptide plane orientations selected by the preliminary average structure and the X-ray are scored as similar. This was the case for 80.1(±3.7)% of the peptide planes in the four model proteins considered. In order to establish what level of improvement might be expected if the molecular alignment frame were known exactly we have repeated the calculations described above using a single alignment frame, AF_{XTAL}, for both the preliminary and X-ray structures. In this case, 82% of the orientations selected are similar, suggesting that the majority of the differences, at least in the case of $R = 0.26$, result from errors in the average structure. In order to minimize errors in peptide plane orientations preliminary structures are, of course, rejected from consideration if they contain any violations of the NOE or dihedral angle restraints.

As described above, for each protein considered in Table 2 five sets of NOE restraints were generated, along with dihedral restraints and dipolar couplings. Starting from dipolar couplings for 69 % of the residues and following the protocol outlined in Figure 5, orientational restraints were obtained for approximately 50 % of the residues (i.e. $S_{k,max} \geqslant 3.5$ for ~50 % of the residues in each protein or ~75 % of the residues for which dipolar couplings were available). Structures were calculated as described in the previous section and in Materials and Methods. A comparison of the precision (average pair-wise rmsd between structures) and accuracy (average pair-wise rmsd between structures and the X-ray) of the structures without and with the inclusion of dipolar coupling-based restraints is provided in Table 2, along with the number and type of restraints used for each protein. In all cases the results reported are averages based on the ten lowest energy structures for which no NOE or dihedral angle violations were obtained. A dramatic improvement in the quality of structures is noted in the case of 1fps, where initial structures were poor, with smaller but significant gains observed for the other motifs as well. Moreover, in every case the percentage of residues in the most favored region of Ramachandran space increased, on average by 4.2(±2.2) %, as calculated by PROCHECK-NMR (Laskowski *et al.*, 1998). Finally, the procedures described above were repeated with the same families of structures using $A_a = 0.0017$, $R = 0.1$ and $R = 0.5$ and very similar levels of precision and accuracy to those reported in Table 2 for $R = 0.26$ were obtained.

The effect of incorporating incorrect peptide plane orientations on the precision and accuracy of the resultant structures was evaluated by replacing all of the dipolar based restraints obtained using the protocol outlined in Figure 5 with restraints determined directly from the crystal structure oriented in AF_{XTAL}. In columns five and six of Table 2, only the peptide planes that were restrained by the preliminary NMR structures were used (control 1), providing a picture of the quality of structures that could be expected if all of the available dipolar data were used correctly. The accuracy of the structures improve, on average, from 3.2 to 3.0 Å (1bgl), 3.5 to 3.1 Å (1fps), 2.4 to 2.4 Å (1mua) and 2.3 to 2.2 Å (1rla), with somewhat smaller gains in precision. Thus, an upper bound in accuracy that can be expected in cases where approximately 50 % of the peptide planes are restrained varies from 2 to 3 Å, depending on the quality of the initial structures. In order to determine the limit of the approach used here, all of the peptide planes of the protein were restrained to the orientation found in the crystal structure (columns seven and eight of Table 2) and structures calculated (control 2). Again, the structures all improve in accuracy (accuracy between 2.4 Å for 1fps and 1.8 Å for 1rla) and, in this case, in precision as well. There are very few violations in these structures (<0.5 % of the dipolar based restraints are violated, on average); therefore, the precision and accuracy in this case result from the error limits (30 °) placed on the dipolar generated restraints.

MBP structures calculated with residual dipolar couplings

The goal of the present work has been to develop a simple and efficient approach for incorporation of dipolar coupling-based restraints into a structural refinement scheme for large proteins where only a limited number of NOEs and an incomplete dipolar coupling set are available. Although high-resolution structures cannot be generated using the methodology described, the results illustrated in Table 2 suggest that, in many cases, significant improvements relative to structures calculated from NOEs and dihedral angles exclusively can be obtained, with an average coordinate accuracy on the order of ~3 Å. We have applied this methodology to refine the NOE-based solution structures of MBP, shown in Figure 3.

Figure 6 illustrates the improvements in MBP structures resulting from inclusion of dipolar coupling based restraints for 188 residues using the protocol outlined in Figure 5 and described above. In Figure 6(a) a superposition of the ten lowest energy structures generated from 1943 NOE, 555 (ϕ,ψ) dihedral angle and 48 hydrogen bonding restraints is shown for reference, while in (b) the ten lowest energy structures obtained by including dipolar coupling derived restraints are illustrated. The X-ray structure of MBP with β-cyclodextrin, 1dmb (Sharff *et al.*, 1993), is superimposed on the ten lowest energy NMR structures in Figure 6(c). It is clear that including dipolar restraints has significantly improved the precision of the structures. The largest difference between the NMR and X-ray derived structures lies in a region extending from Pro229 to Lys239 (indicated by * in Figure 6). These residues are part of a helix which lies in the cleft between the two domains that interacts with bound sugar (Sharff *et al.*, 1993) and because of conformational heterogeneity, many of the expected cross-peaks were not observed in any spectra (Gardner *et al.*, 1998).

Although high-resolution X-ray structures exist for MBP both in the absence of sugar and in the presence of a number of different carbohydrate ligands (Sharff *et al.*, 1992), the relative orientation of domains in the molecule can be influenced by crystal packing interactions which are quite extensive in the β-cyclodextrin bound form (Sharff *et al.*, 1993; Spurlino *et al.*, 1991). It is of significant interest, therefore, to compare the position of domains as established by X-ray (1dmb) and NMR methods. In order to quantify differences between 1dmb and the average, energy minimized NMR structure, the C domains of both molecules have been superimposed and the transformation resulting in the superposition of the N domains of 1dmb and the NMR structure calculated. The Euler angles

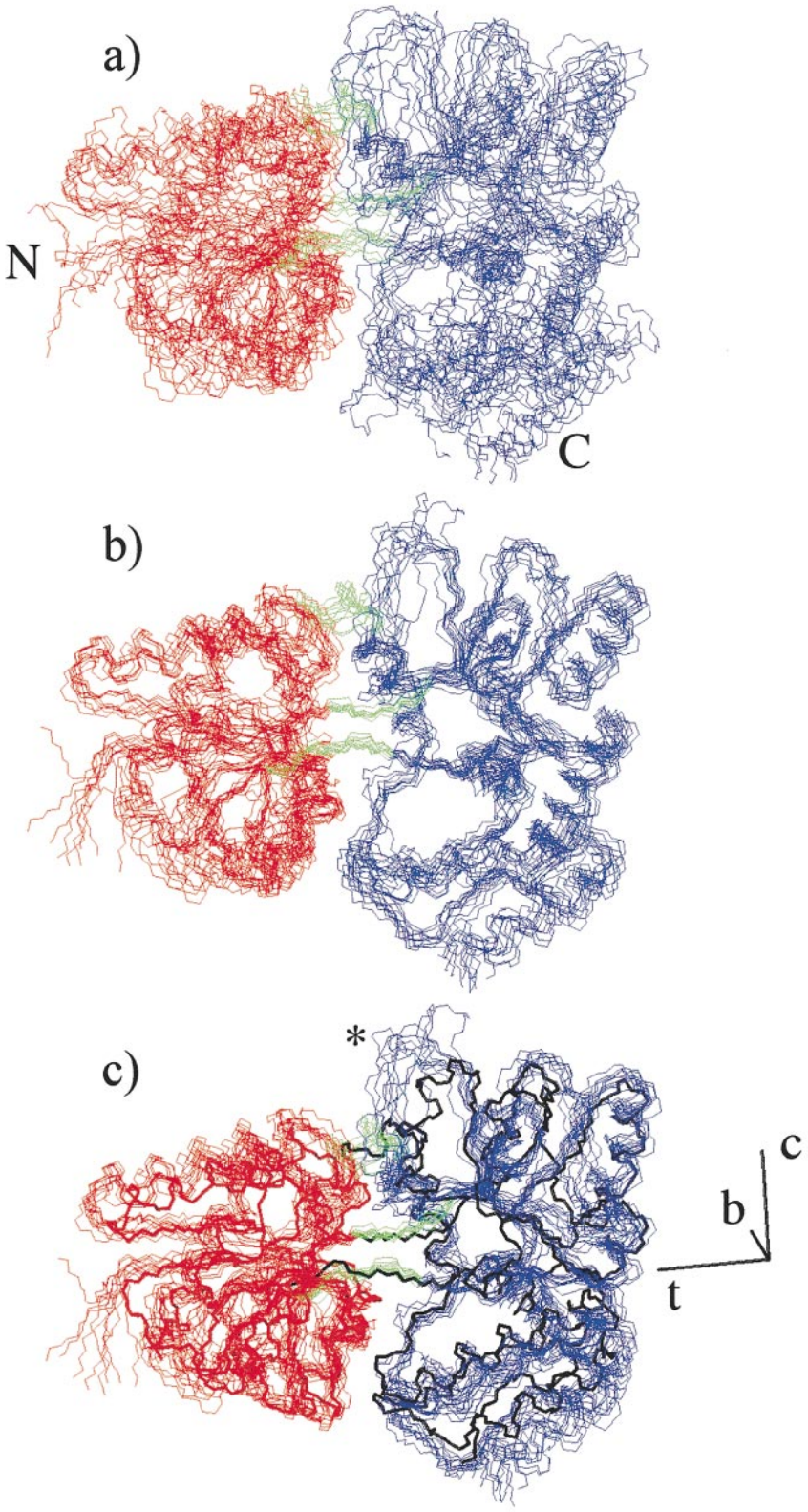

Figure 6. Comparison of solution structures of the MBP/β-cyclodextrin complex obtained with and without dipolar coupling-based restraints. Best-fit (residues 6 to 370) superposition of the ten lowest energy structures generated from (a) NOE, dihedral angle and hydrogen bond restraints and (b) NOE, dihedral angle, hydrogen bond and dipolar restraints. The domains are colored red (N domain), blue (C domain), and green (Linker). In (c) the structures from (b) are aligned with the crystal structure of the MBP/β-cyclodextrin complex, 1dmb (black heavy line). The closure, twist, bend (c,t,b) coordinate frame is illustrated, with the bend axis extending out of the plane of the paper.

required for this transformation, in turn, define a hinge axis about which rotation from one N domain to the other occurs, as well as the amplitude of this rotation. It is convenient, however, to recast this transformation in terms of another set of angles which correspond to rotations about a defined molecular frame, providing more insight into the changes. Specifically, we define a closure, bend, twist axis frame where the twist axis is colinear with a vector connecting the centers of mass

of the two domains, the closure axis lies in a plane comprising the twist and hinge axes and the bend axis is perpendicular to the first two (Skrynnikov *et al*., 2000). This frame is illustrated in Figure 6. In the absence of dipolar-based restraints, the extent of domain closure is ill defined, although it is clear that the domains are more closed in solution than in the corresponding crystal form. In contrast, the addition of the dipolar restraints results in an average closure angle of 12(±3)° relative to the X-ray structure. It is noteworthy that the values for closure (12(±3)°), bend (−2(±2)°) and twist (−3(±4)°) obtained from the solution structure are very similar to values calculated from a combined NMR/X-ray study where the dipolar couplings were used to ''guide'' the relative orientation of the domains in a number of different X-ray structures to their solution positions (Skrynnikov *et al*., 2000). However, we feel that this level of agreement is somewhat serendipitous, given the accuracy of the NMR structures (3 Å). For example, if the dipolar coupling data are reinterpreted using A_a and R-values obtained from a fit of the measured dipolar couplings with those predicted from equation (1) using the crystal structure of the molecule ($A_a = 0.00155$, $R = 0.18$), an average closure angle of 4(±3)° is obtained. In this case structures with precision and accuracy levels similar to what was calculated with $R = 0.26$ have been generated.

Consistent with the results of the simulations, the accuracy of both of the domains have improved, illustrated in Table 1. In addition, the percentage of residues in the most favored region of Ramachandran space has also increased. It is also noteworthy that although the ten lowest energy structures calculated with dipolar restraints did not show NOE violations, on average six of the 555 dihedral angle and 51 of the 940 dipolar restraints from 41 peptide planes were violated (recall that each of the five dipole vectors in Figure 4(a) is restrained/peptide plane). Table 1 summarizes the violations; the average violation of dipolar restraints is 1.7° and a maximum violation of 10° was obtained over all ten lowest energy structures. The level of convergence obtained with the present approach is to be contrasted with the situation when direct refinement against dipolar couplings was employed when roughly the same number of dihedral angles were violated, but the great majority of dipolar restraints could not be satisfied to within the measured errors in dipolar couplings. In the latter case structures were generated with average rmsd values to 1dmb of 3.4 , 4.8 and 6.2 Å for the N domain, C domain and the overall molecule. These values are significantly larger than those reported in Table 1 using the methodology described in the text.

Summary

The global fold of MBP has been derived using a set of distance and dipolar coupling restraints measured on an ^{15}N, ^{13}C, ^{2}H, Val, Leu, Ile (δ1) methyl-protonated sample. A new protocol for incorporation of orientational restraints has been presented which has better convergence properties than previous methods in cases where structures are defined by only a limited number of NOE restraints. Although the methodology offers benefits relative to previous approaches, significant improvements can, nevertheless, be envisioned. Specifically, protocols which are less sensitive to the precise value of rhombicity, R, and are less dependant on initial structures would represent an important advance and are currently under development.

Materials and Methods

Sample preparation

Two MBP samples were prepared for data collection, including $^{15}N,^{2}H$ and $^{15}N,^{13}C,^{2}H$-Val, Leu, Ile (δ1 only) methyl protonated molecules. ^{15}N-labeled MBP was generated with $^{2}H,^{12}C$-glycerol as the sole carbon source (Mok *et al*., 1999), while the ^{13}C-labeled sample was prepared using ^{13}C, ^{2}H-glucose as the carbon source, supplemented by the precursors [3,3 $^{2}H_2$] ^{13}C α-ketobutyrate and [3-^{2}H] ^{13}C α-ketoisovalerate (Goto *et al*., 1999). The samples were purified as described (Gardner *et al*., 1998). To fully protonate slowly exchanging backbone amide sites, the protein was partially unfolded for three hours at room temperature using a guanidinium hydrochloride buffer, as described (Gardner *et al*., 1998). Samples consisted of 1.0-1.4 mM protein, 2 mM β-cyclodextrin, 20 mM phosphate buffer (pH 7.2), 3 mM NaN_3, 100 μM EDTA and 5-10 % $^{2}H_2O$. Dipolar coupling measurements were made with Pf1 phage liquid-crystalline solvent at a concentration of 19 mg/ml (^{2}H splitting of 19 Hz), as described by Yang *et al*. (1999).

Data collection and analysis

All spectra were recorded on a Varian Inova 600 MHz spectrometer at 37°C. 4D ^{15}N, ^{15}N-edited NOESY (Grzesiek *et al*., 1995; Venters *et al*., 1995), 4D ^{13}C, ^{15}N-edited NOESY (Muhandiram *et al*., 1993) and the 3D CT-^{13}C-edited NOESY (Zwahlen *et al*., 1998a) data sets were recorded using mixing times of 175 ms with pulse schemes and parameters that have been described. Note that the latter experiment was obtained with constant-time (CT) acquisition only in the t_1 dimension and NOEs to both ^{15}N and ^{13}C-coupled protons were recorded during acquisition. The (HM)CMCB(CMHM)-NOESY scheme used in the present study was modified slightly from the sequence reported (Zwahlen *et al*., 1998b), in that the second CT period (t_2 acquisition) was set to a duration of $1/J_{CC}$, the RE-BURP pulses at positions b and e in Figure 1 by Zwahlen *et al*. (1998b) were omitted and the RE-BURP pulse (Geen & Freeman, 1991) of phase ϕ5 covered the whole aliphatic bandwidth. In this way NOEs originating on Leu residues can be observed with a high level of intensity. Data sets were processed using NMRPipe software (Delaglio *et al*., 1995) and analyzed with NMRView version 3 (Johnson & Blevins, 1994).

Distances between amide protons were estimated from peak intensities in the 4D ^{15}N, ^{15}N-edited NOESY by calibrating assigned NOEs from elements of regular

secondary structure with distances measured in high resolution X-ray structures of proteins. Specifically, sequential amide protons in α-helices are less than 3 Å, $i,i+2$ NOEs in α-helices and sequential NOEs in β-sheets correspond to distances ranging from 3.5-4.5 Å, while $i,i+2$ distances in β-sheets and $i,i+4$ distances in α-helices are 6.4(±0.8) Å and 6.2 Å(±0.3) Å, respectively (Wüthrich, 1986). The intensities of NOE correlations were classified as strong, medium and weak, corresponding to distance limits of 1.8-3.5 Å, 1.8-5.0 Å, and 1.8-7.0 Å, respectively. NOEs measured in methyl-based experiments were not quantitated since intensities of correlations in these experiments can be influenced significantly by relaxation during the CT intervals, as well as by strong coupling effects in the case of some Leu residues. Thus, distance bounds of 1.8 to 8.0 Å for the amide-methyl NOEs and 1.8 to 8.5 Å for the methyl-methyl NOEs were employed.

Dihedral angle predictions for backbone angles ϕ and ψ were made from the backbone chemical shifts of MBP after correction for the ^{2}H (^{1}H-methyl Val, Leu, Ile $C^{\delta 1}$) labeling scheme employed, as described by Venters *et al.* (1996). Restraints were generated for the most part using the program TALOS (Cornilescu *et al.*, 1999), after the chemical shifts for MBP had been removed from the TALOS database. Restraints consisting of the average ϕ,ψ values ±2 standard deviations or at least ±15° from the average predicted value were employed for 219 residues. In cases where angles found by TALOS did not satisfy acceptance criteria, chemical shift index (Wishart & Sykes, 1994) predictions were used, resulting in constraints for an additional 59 residues. Values of $\phi = -70(\pm 50)^\circ$ and $\psi = -50(\pm 50)^\circ$ and $\phi = -140(\pm 60)^\circ$ and $\psi = 130(\pm 90)^\circ$ were employed for residues in α-helices and β-sheets, respectively, when the chemical shift index was used (Luginbühl *et al.*, 1995).

In addition to the 1943 NOE and the 555 dihedral angle restraints employed in structure calculations, 48 hydrogen bond restraints were also used for residues in regular elements of secondary structure. Residues with hydrogen bonding were identified by comparing H^N-^{15}N correlation spectra of samples of MBP prepared prior to back-exchange of slowly exchanging amide protons (see above) with spectra recorded after the exchange process was allowed to occur. Note that hydrogen bonding restraints were included only for residues in regular elements of secondary structure where a clear donor and acceptor could be identified.

Dipolar couplings ($\delta N_i - H_i^N$, $\delta N_i - C'_{i-1}$, $\delta C_{i-1}^{\alpha} - C'_{i-1}$, $\delta H_i^N - C'_{i-1}$, and $\delta H_i^N - C_{i-1}^{\alpha}$) were measured using the methyl-protonated, ^{15}N, ^{13}C, ^{2}H-labeled MBP sample with TROSY-based HNCO pulse sequences described by Yang *et al.* (1999).

Structure calculations

Dipolar coupling based restraints have been incorporated into a new module based on the original SANI module (Clore *et al.*, 1998b) for performing direct refinement against measured dipolar couplings and written for CNS version 0.5 (Brünger *et al.*, 1998). Here, the orientation of a peptide plane is restrained by defining the polar angles describing the orientation of the five normalized planar dipolar vectors ($\vec{P}_{A-B}^{k,best}$) illustrated in Figure 4(a). These polar angles define the orientation of dipolar vectors with respect to an alignment frame, described by a pseudo-residue, arbitrarily numbered 500, with the atoms OO, X, Y, and Z corresponding to the origin and the *x*, *y*, and *z*-axes respectively.

The new restraints are of the form:

assign (residue 500 and name OO)
(residue 500 and name Z)
(residue 500 and name X)
(residue 500 and name Y)
(residue i_A and name A)
(residue i_B and name B) θ_{A-B} ϕ_{A-B} ψ_{error},

where θ_{A-B} and ϕ_{A-B} are the polar angles that $\vec{P}_{A-B}^{k,best}$ makes with the defined-axis system, $\vec{P}_{A-B}^{k,best}$ is a normalized dipole vector from the peptide plane which gives $S_{k,max}$ (see above) and ψ_{error} is the allowed error in the angle between the target vector and the actual vector in the calculated structures. Note that a separate restraint is needed for each of the 5 *A-B* dipole vectors ($P_{A-B}^{k,best}$) per peptide plane. The square-well potential function:

$$E_{DIP} = k_{DIP}(1 - \vec{P}_{A-B}^{k,best} \times \vec{C}_{A-B})^2 \text{ if } \vec{P}_{A-B}^{k,best} \times \vec{C}_{A-B} < \cos\psi_{error}$$
$$= 0 \text{ if } \vec{P}_{A-B}^{k,best} \times \vec{C}_{A-B} > \cos\psi_{error} \quad (4)$$

is used to restrain each of the *A-B* dipolar vectors during simulated annealing, where E_{DIP} is the energy, k_{DIP} is the force constant, and $\vec{C}_{A-B}$ is the normalized vector connecting atoms A and B in the structure.

Structure calculations were performed starting from extended structures (Nilges *et al.*, 1988) and using a combination of torsion angle dynamics, TAD (Stein *et al.*, 1997), and Cartesian dynamics. In most cases default CNS parameters were employed; therefore only a brief description is provided which emphasizes the modifications. An initial TAD (hot) phase was performed at a temperature of 50,000 K consisting of 6000 molecular dynamics steps each of 15 fs. During this stage all of the force constants (set to standard values, $k_{DIP} = 50$) were kept constant. Subsequently, a TAD cooling phase comprised of 30,000 steps each of 2 fs was employed with the temperature decreasing from 50,000 K to 0 K during this interval. All parameters are scaled using default values and k_{DIP} increases from 50 to 10,000 over the course of this period. Finally, a second cooling phase using Cartesian dynamics (5000 steps with a 1fs time step, starting from a temperature of 300 K) was employed with parameters starting from their values during the hot TAD phase and increasing to their final values over the duration of this interval, as in the second TAD period. All structures were calculated using the same basic protocol, whether or not dipolar coupling based restraints were incorporated.

Protein Data Bank accession codes

Structures have been deposited in the RCSB Protein Data Bank, accession codes lezo and lezp.

Acknowledgments

G.A.M. would like to dedicate this paper to a friend and fellow scientist who passed away, Dr Jill Hungerford. The authors also wish to thank Drs Kevin Gardner, Catherine Zwahlen, Natalie Goto and especially Nikolai Skrynnikov for many helpful discussions. G.A.M. is a

recipient of a Human Frontiers Science Program Postdoctoral fellowship and L.E.K. is an International Howard Hughes Scholar. This work was supported by a grant from the Medical Research Council of Canada to L.E.K. All software developed and all of the data used in the structure calculations will be made available from the authors upon request.

References

Bax, A. (1994). Multidimensional nuclear magnetic resonance methods for protein studies. *Curr. Opin. Struct. Biol.* **4**, 738-744.

Bewley, C. A., Gustafson, K. R., Boyd, M. R., Covell, D. G., Bax, A., Clore, G. M. & Gronenborn, A. M. (1998). Solution structure of cyanovirin-N, a potent HIV-inactivating protein. *Nature Struct. Biol.* **5**, 571-578.

Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallog. sect. D,* **54**, 905-921.

Cai, M., Huang, Y., Zheng, R., Wei, S. Q., Ghirlando, R., Lee, M. S., Craigie, R., Gronenborn, A. M. & Clore, G. M. (1998). Solution structure of the cellular factor BAF responsible for protecting retroviral DNA from autointegration. *Nature Struct. Biol.* **5**, 903-909.

Clore, G. M., Gronenborn, A. M. & Bax, A. (1998a). A robust method for determining the magnitude of the fully asymmetric alignment tensor of oriented macromolecules in the absence of structural information. *J. Magn. Reson.* **113**, 216-221.

Clore, G. M., Gronenborn, A. M. & Tjandra, N. (1998b). Direct structure refinement against residual dipolar couplings in the presence of rhombicity of unknown magnitude. *J. Magn. Reson.* **131**, 159-162.

Clore, G. M., Starich, M. R. & Gronenborn, A. M. (1998c). Measurement of residual dipolar couplings of macromolecules aligned in the nematic phase of a colloidal suspension of rod-shaped viruses. *J. Am. Chem. Soc.* **120**, 10571-10572.

Clore, G. M., Starich, M. R., Bewley, C. A., Cai, M. & Kuszewski, J. (1999). Impact of residual dipolar couplings on the accuracy of NMR structures determined from a minimal number of NOE restraints. *J. Am. Chem. Soc.* **121**, 6513-6514.

Cornilescu, G., Delaglio, F. & Bax, A. (1999). Protein backbone angle restraints from searching a database for chemical shift and sequence homolgy. *J. Biomol. NMR,* **13**, 289-302.

Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J. & Bax, A. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR,* **6**, 277-293.

Delaglio, F., Kontaxis, G. & Bax, A. (2000). Protein structure determination using molecular fragment replacement and NMR dipolar couplings. *J. Am. Chem. Soc.* **122**, 2142-2143.

Farmer, B. T. & Venters, R. A. (1998). NMR of perdeuterated large proteins. In *Biological Magnetic Resonance*. (Krishna, N. R. & Berliner, L. J., eds), vol. 16, pp. 75-120, Kluwer Academic/Plenum Publishers, New York.

Fischer, M. W., Losonczi, J. A., Weaver, J. L. & Prestegard, J. H. (1999). Domain orientation and dynamics in multidomain proteins from residual dipolar couplings. *Biochemistry,* **38**, 9013-9022.

Gardner, K. H., Rosen, M. K. & Kay, L. E. (1997). Global folds of highly deuterated, methyl protonated proteins by multidimensional NMR. *Biochemistry,* **36**, 1389-1401.

Gardner, K. H. & Kay, L. E. (1998). The use of ^{2}H, ^{13}C, ^{15}N multidimensional NMR to study the structure and dynamics of proteins. *Annu. Rev. Biophys. Biomol. Struct.* **27**, 357-406.

Gardner, K. H., Zhang, X., Gehring, K. & Kay, L. E. (1998). Solution NMR studies of a 42 kDa *E. coli* maltose binding protein/β cyclodextrin complex: chemical shift assignments and analysis. *J. Am. Chem. Soc.* **120**, 11738-11748.

Geen, H. & Freeman, R. (1991). Band-selective radiofrequency pulses. *J. Magn. Reson.* **93**, 93-141.

Goto, N. K., Gardner, K. H., Mueller, G. A., Willis, R. C. & Kay, L. E. (1999). A robust and cost-effective method for the production of Val, Leu, Ile (δ1) methyl-protonated ^{15}N-,^{13}C-,^{2}H-labeled proteins. *J. Biomol. NMR,* **13**, 369-374.

Grzesiek, S., Wingfield, P., Stahl, S., Kaufman, J. & Bax, A. (1995). Four-dimensional ^{15}N separated NOESY of slowly tumbling perdeuterated ^{15}N-enriched proteins. Application to HIV-1 Nef. *J. Am. Chem. Soc.* **117**, 9594-9595.

Hansen, M. R., Mueller, L. & Pardi, A. (1998). Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. *Nature Struct. Biol.* **5**, 1065-1074.

Jacobson, R. H., Zhang, X. J., DuBose, R. F. & Matthews, B. W. (1994). Three-dimensional structure of β-galactosidase from *E. coli*. *Nature,* **369**, 761-766.

Johnson, B. A. & Blevins, R. A. (1994). NMRView: a computer program for the visualization and analysis of NMR data. *J. Biomol. NMR,* **4**, 603-614.

Kanyo, Z. F., Scolnick, L. R., Ash, D. E. & Christianson, D. W. (1996). Structure of a unique binuclear manganese cluster in arginase. *Nature,* **383**, 554-557.

Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graphics,* **14**, 51-55.

Laskowski, R. A., Rullman, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1998). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR,* **8**, 477-486.

Luginbühl, P., Szyperski, T. & Wüthrich, K. (1995). Statistical basis for the use of $^{13}C^{\alpha}$ chemical shifts in protein structure refinement. *J. Magn. Reson. ser. B,* **109**, 229-233.

Mok, Y. K., Kay, C. M., Kay, L. E. & Forman-Kay, J. D. (1999). NOE data demonstrating a compact unfolded state for an SH3 domain under non-denaturing conditions. *J. Mol. Biol.* **289**, 619-638.

Muhandiram, D. R., Xu, G. Y. & Kay, L. E. (1993). An enhanced-sensitivity pure absorption gradient 4D ^{15}N, ^{13}C-edited NOESY experiment. *J. Biomol. NMR,* **3**, 463-470.

Murzin, A. G., Brenner, S. E., Hubbard, T. & Chothia, C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536-540.

Nilges, M., Gronenborn, A. M., Brünger, A. T. & Clore, G. M. (1988). Determination of three-dimensional structures of proteins by simulated annealing with interproton distance restraints. Application to cram-

bin, potato carboxypeptidase inhibitor and barley serine proteinase inhibitor 2. *Protein Eng.* **2**, 27-38.
Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1997). Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl Acad. Sci. USA,* **94**, 12366-12371.
Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. (1998). Transverse relaxation-optimized spectroscopy (TROSY) for NMR studies of aromatic spin systems in ^{13}C-labeled proteins. *J. Am. Chem. Soc.* **120**, 6394-6400.
Rosen, M. K., Gardner, K. H., Willis, R. C., Parris, W. E., Pawson, T. & Kay, L. E. (1996). Selective methyl group protonation of perdeuterated proteins. *J. Mol. Biol.* **263**, 627-636.
Santoro, J. & King, G. C. (1992). A constant-time 2D overbodenhausen experiment for inverse correlation of isotopically enriched species. *J. Magn. Reson.* **97**, 202-207.
Sharff, A. J., Rodseth, L. E., Spurlino, J. C. & Quiocho, F. A. (1992). Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis. *Biochemistry*, **31**, 10657-10663.
Sharff, A. J., Rodseth, L. E. & Quiocho, F. A. (1993). Refined 1.8-Å structure reveals the mode of binding of β-cyclodextrin to the maltodextrin binding protein. *Biochemistry,* **32**, 10553-10559.
Skrynnikov, N. R., Goto, N. K., Yang, D., Choy, W. Y., Tolman, J. R., Mueller, G. A. & Kay, L. E. (2000). Orienting domains in proteins using dipolar couplings measured by liquid-state NMR: Differences in solution and crystal forms of maltodextrin binding protein loaded with β-cyclodextrin. *J. Mol. Biol.* **295**, 1265-1273.
Spurlino, J. C., Lu, G. Y. & Quiocho, F. A. (1991). The 2.3-A resolution structure of the maltose or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. *J. Biol. Chem.* **266**, 5202-5219.
Stein, E. G., Rice, L. M. & Brunger, A. T. (1997). Torsion-angle molecular dynamics as a new efficient tool for NMR structure calculation. *J. Magn. Reson.* **124**, 154-164.
Tarshis, L. C., Yan, M., Poulter, C. D. & Sacchettini, J. C. (1994). Crystal structure of recombinant farnesyl diphosphate synthase at 2.6-Å resolution. *Biochemistry,* **33**, 10871-10877.
Tjandra, N. & Bax, A. (1997). Direct measurement of distances and angles in biomolecules by NMR in a dilute liquid crystalline medium. *Science,* **278**, 1111-1114.
Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M. & Bax, A. (1997). Using dipolar ^{1}H-^{15}N and ^{1}H-^{13}C couplings in the structure determination of magnetically oriented macromolecules in solution. *Nature Struct. Biol.* **4**, 732-738.
Tolman, J. R., Flanagan, J. M., Kennedy, M. A. & Prestegard, J. H. (1995). Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. *Proc. Natl Acad. Sci. USA,* **92**, 9279-9283.
Tweedy, N. B., Nair, S. K., Paterno, S. A., Fierke, C. A. & Christianson, D. W. (1993). Structure and energetics of a non-proline *cis*-peptidyl linkage in a proline-202 → alanine carbonic anhydrase II variant. *Biochemistry,* **32**, 10944-10949.
Venters, R. A., Farmer, B. T., Fierke, C. A. & Spicer, L. D. (1996). Characterizing the use of perdeuteration in NMR Studies of large proteins: ^{13}C, ^{15}N and ^{1}H assignments of human carbonic anhydrase II. *J. Mol. Biol.* **264**, 1101-1116.
Venters, R. A., Metzler, W. J., Spicer, L. D., Mueller, L. & Farmer, B. T. (1995). Use of ^{1}HN-^{1}HN NOEs to determine protein global folds in perdeuterated proteins. *J. Am. Chem. Soc.* **117**, 9592-9593.
Vuister, G. W. & Bax, A. (1992). Resolution enhancement and spectral editing of uniformly ^{13}C-enriched proteins by homonuclear broadband ^{13}C decoupling. *J. Magn. Reson.* **98**, 428-435.
Wider, G. & Wüthrich, K. (1999). NMR spectroscopy of large molecules and multimolecular assemblies in solution. *Curr. Opin. Struct. Biol.* **9**, 594-601.
Wishart, D. S. & Sykes, B. D. (1994). The ^{13}C chemical-shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical-shift data. *J. Biomol. NMR,* **4**, 171-180.
Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids,* John Wiley & Sons, New York.
Yang, D., Venters, R. A., Mueller, G. A., Choy, W. Y. & Kay, L. E. (1999). Trosy-based HNCO pulse sequences for the measurement of ^{1}HN-^{15}N, ^{15}N-^{13}CO, ^{1}HN-^{13}CO, ^{13}CO-^{13}Cα dipolar couplings in ^{15}N, ^{13}C, ^{2}H-labeled proteins. *J. Biomol. NMR,* **14**, 333-343.
Zhang, Y., Gardina, P. J., Kuebler, A. S., Kang, H. S., Christopher, J. A. & Manson, M. D. (1999). Model of maltose-binding protein/chemoreceptor complex supports intrasubunit signaling mechanism. *Proc. Natl Acad. Sci. USA,* **96**, 939-944.
Zwahlen, C., Gardner, K. H., Sarma, S. P., Horita, D. A., Byrd, R. A. & Kay, L. E. (1998a). An NMR experiment for measuring methyl-methyl NOEs in ^{13}C labeled proteins with high resolution. *J. Am. Chem. Soc.* **120**, 7617-7625.
Zwahlen, C., Vincent, S. J. F., Gardner, K. H. & Kay, L. E. (1998b). Significantly improved resolution for NOE correlations from valine and isoleucine ($C^{\gamma 2}$) methyl groups in ^{15}N,^{13}C and ^{15}N,^{13}C,^{2}H-labeled proteins. *J. Am. Chem. Soc.* **120**, 4825-4831.

Edited by P. E. Wright

(Received 29 March 2000; received in revised form 4 April 2000; accepted 9 April 2000)