

CABM Symposium

Complete protein structure determination using backbone residual dipolar couplings and sidechain rotamer prediction

Michael Andrec¹, Yuichi Harano¹, Matthew P. Jacobson², Richard A. Friesner² & Ronald M. Levy^{1*}

¹*Department of Chemistry and Chemical Biology, Rutgers, The State University of New Jersey, 610 Taylor Road, Piscataway, NJ 08854-8087, USA;* ²*Department of Chemistry, Columbia University, New York, NY 10027, USA;* * *Author for correspondence (Tel: (732) 445-3947; Fax: (732) 445-5958; E-mail: ronlevy@lutece.rutgers.edu)*

Received 12 March 2002; accepted in revised form 7 May 2002

Key words: database, overlap, protein fragments, proteomics, sidechain prediction, solvation, structural genomics

Abstract

Residual dipolar couplings provide significant structural information for proteins in the solution state, which makes them attractive for the rapid determination of protein structures. While dipolar couplings contain inherent structural ambiguities, these can be reduced via an overlap similarity measure that insists that protein fragments assigned to overlapping regions of the sequence must have self-consistent structures. This allows us to determine a backbone fold (including the correct C^α–C^β bond orientations) using only residual dipolar coupling data from one ordering medium. The resulting backbone structures are of sufficient quality to allow for modeling of sidechain rotamer states using a rotamer prediction algorithm and a force field employing the Surface Generalized Born continuum solvation model. We demonstrate the applicability of the method using experimental data for ubiquitin. These results illustrate the synergies that are possible between protein structural database and molecular modeling methods and NMR spectroscopy, and we expect that the further development of these methods will lead to the extraction of high resolution structural information from minimal NMR data.

Introduction

There exists a need for methods that would allow the more rapid determination of protein structure using NMR than can currently be attained, both from the viewpoint of traditional structural biology, as well as from the ‘proteomics’ and ‘structural genomics’ perspective. One of the principle rate-limiting steps in NMR structure determination is the sequential assignment of sidechain proton resonances and the assignment of NOESY crosspeaks to particular sidechain resonances. These steps are considerably more difficult and time-consuming than the sequential assignment of chemical shifts along the peptide backbone, for which relatively robust automated methods al-

ready exist [1]. Therefore, it would be desirable to have a method for obtaining reliable structural information based on the smallest possible additional data collection beyond that needed for the backbone resonance assignments. It would be especially useful if such minimal data could yield all-atom models, rather than simply backbone structures. In this paper, we describe our recent work in the pursuit of this goal.

Backbone residual dipolar couplings are of particular interest for this purpose since they require relatively little data collection time and provide considerable structural information through their dependence on the orientation of internuclear vectors relative to an order frame [2]. The development of a variety of orienting media (such as lipid bicelles and

filamentous phage) [3–5] have increased the practicality of such measurements in recent years, and the use of residual dipolar couplings as a supplement to NOEs and scalar couplings in the refinement of high-resolution NMR solution structures is becoming more common [6].

Applications of residual dipolar couplings to the study of protein structure in the absence of NOEs have previously been reported [7–11]. All of these methods, however, are limited to different degrees by the orientational ambiguities arising from the many-to-one relationship between internuclear vector orientation and dipolar coupling. We have shown in earlier work [12] that the effect of these ambiguities can be greatly reduced by the use of filtering procedures based on measures of structural similarity between protein fragments with overlapping regions. In this paper, we summarize that work and give some new results on the use of sidechain rotamer prediction algorithms for the building of all-atom models from backbone structures determined from residual dipolar coupling data.

Theory and methods

A residual dipolar coupling associated with a given internuclear vector is related to the orientation of that vector relative to an order tensor and is given by

$$D = D_a [(3 \cos^2 \theta - 1) + \frac{3}{2} R \cos 2\phi \sin^2 \theta] \quad (1)$$

where D_a is a constant which depends on the internuclear distance and the gyromagnetic ratios of the spins involved, R ($0 \leq R \leq 2/3$) is a measure of the asymmetry of the order tensor, and θ and ϕ are spherical coordinates which relate the internuclear vector to the principle axis system (PAS) of the order tensor [2]. Alternatively, one can rewrite equation 1 in the form

$$D = (x \ y \ z) \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{xy} & D_{yy} & D_{yz} \\ D_{xz} & D_{yz} & D_{zz} \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (2)$$

where D_{ij} are the elements of a symmetric and traceless matrix proportional to the Saupe order tensor [13, 14] in an arbitrary molecular frame defined by the direction cosines x , y , and z . The relationship between (θ, ϕ) or (x, y, z) and D is many-to-one, since there exist manifolds of (θ, ϕ) or (x, y, z) points which give

rise to the same dipolar coupling, e.g. circles of constant θ in the case of $R = 0$. These degeneracies lead to orientational ambiguities which can give rise to false positive hits when searching a database or to structural ambiguity when constructing a structural model *de novo*.

Since equation (2) is linear in the tensor elements D_{ij} , it is possible to solve for the optimal D_{ij} s which maximize the agreement between a set of internuclear vector orientations and the dipolar coupling data using a computationally efficient linear least squares procedure [14]. Given this fit, we can easily back-calculate the best-fit dipolar couplings and calculate a goodness-of-fit measure known as the ‘Q factor’ [15]:

$$Q = \frac{\sum_i (D_{\text{calc}, i} - D_{\text{obs}, i})^2}{\sum_i D_{\text{obs}, i}^2} \quad (3)$$

where $D_{\text{calc}, i}$ and $D_{\text{obs}, i}$ are the back-calculated best-fit and experimental values for the i -th residual dipolar coupling.

A database consisting of 191 696 seven-residue protein fragments from the SCOP40 database [16] was constructed. Seven-residue data windows, which included dipolar couplings for N-H^{N} , $\text{C}^{\alpha}\text{-H}^{\alpha}$, $\text{C}^{\alpha}\text{-C}$, C-N , and C-H^{N} internuclear vectors from the ‘charged bicelle’ data of Ottiger & Bax [17] (Supporting Information Table 2) were fit to each fragment in this database, and a Q factor calculated. Database fragments derived from domains having clear structural homology to ubiquitin with a CE z-score [18] greater than 4.0 (including ubiquitin itself) were excluded. For each seven-residue window in the protein (i.e. residues 1–7, 2–8, 3–9, ..., 70–76), the 15 fragments with the smallest Q factors were saved. Due to orientational ambiguities, these 15 hits in general contain both true positives (fragments which are similar in structure to that which generated the data) and false positives (fragments which are significantly different in structure from that which generated the data) (Figure 1). Because of this, a filtering procedure is required to locate the true positives.

Our filter is based on the idea that protein fragments assigned to overlapping regions of the sequence must have self-consistent structures: e.g. a selected hit for data window 1 (residues 1–7) must be structurally similar to the selected hit for data window

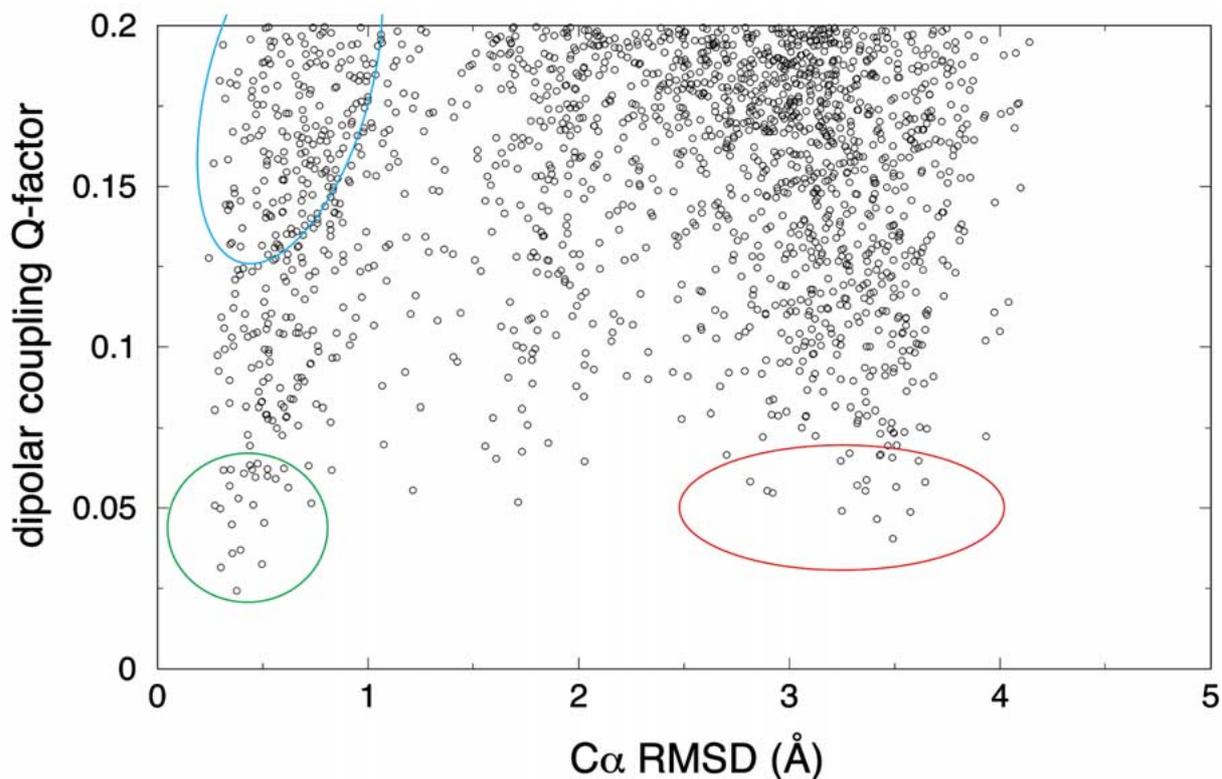


Figure 1. The results of the fragment database search for the seven-residue ubiquitin data window 52 corresponding to residues 52–58. Each small circle represents one fragment in the database, and its position is given by the Q factor for the fit to the dipolar coupling data and that fragment’s C α RMSD to the corresponding residues of the IUBQ crystal structure. Only those fragments with a Q factor less than or equal to 0.2 are shown. The green ellipses denote fragments which are true positives, the cyan ellipses denote false negative fragments, and the red ellipse denotes false positive fragments. If only the 15 fragments with the smallest Q factor values are saved, then only true and false positives will be observed.

2 (residues 2–8) over the overlapping region from residues 2 through 7. Finding the fragment for each data window such that the total overlap similarity is maximal leads to a combinatorial optimization problem that grows exponentially with the number of windows. However, we have developed a bounded tree search algorithm [12] which allows for the efficient search for optimal selections over a block of up to twenty windows.

Once we have found the optimal choice of fragment for each window, we construct a structural model by performing rigid body superpositions of the selected fragments. For each window i ($i > 1$), we translate and rotate the selected fragment for that window so as to minimize the C α RMSD with the selected fragment for window $i - 1$ over the six residues where they overlap [19]. When this has been done for all windows in the protein, each residue position will have atomic coordinates from up to seven

fragments. We construct a consensus structure by calculating the centroid of the atomic coordinates for each of the backbone atoms N, C α , and C, and use these to calculate backbone dihedral angles ϕ and ψ . Up to this point we have made no use of the ‘long range’ information inherent in residual dipolar couplings. This information is reintroduced by refining the backbone dihedral angles by direct minimization of the projected Q factor [20] as a function of the ϕ and ψ angles for the entire protein using an ideal peptide geometry and peptide bond torsion angle ω fixed at 180° as described in [12].

Sidechain rotamer modeling made use of the OPLS force field [21, 22] for protein intramolecular energetics along with the Surface Generalized Born (SGB) implicit model of solvation [23]. The SGB model can be understood as a relatively inexpensive, analytical approximation to the Poisson-Boltzmann description of continuum electrostatics, and the pa-

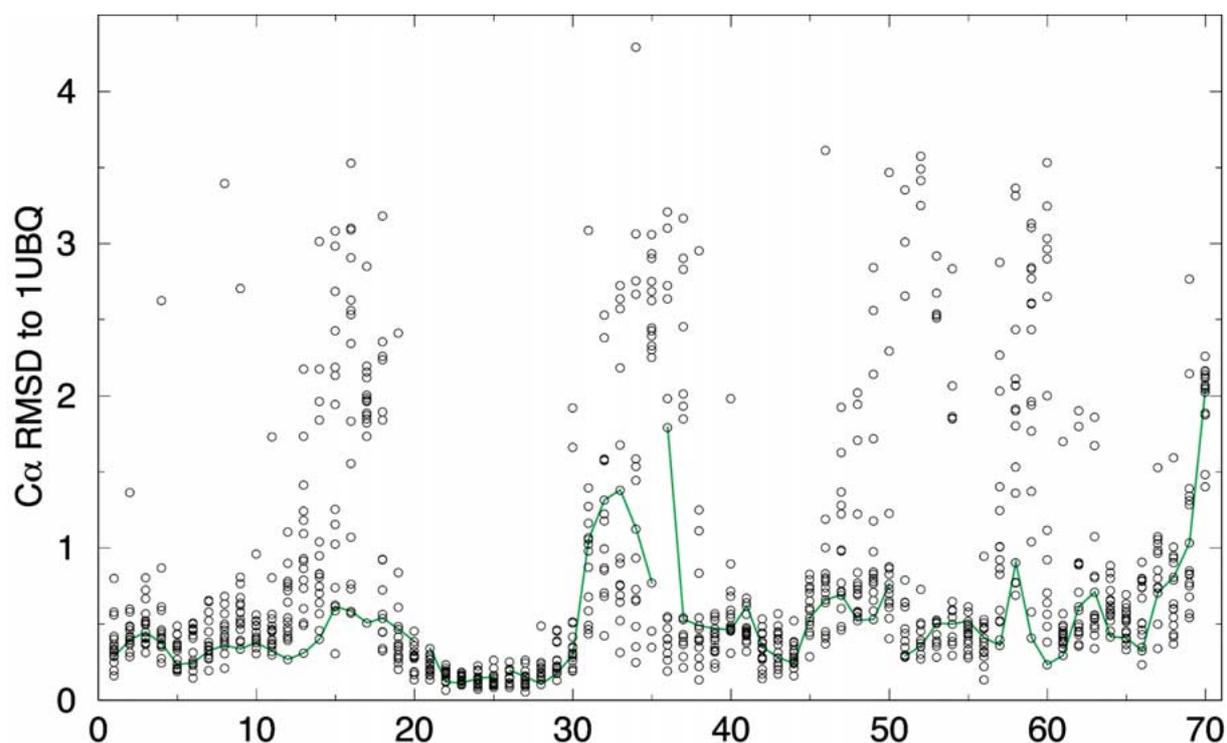


Figure 2. A summary of the fragment database search results for all data windows in ubiquitin using a window size of seven. The small circles represent (for each window) the 15 fragments in the database with the smallest Q factor, and the ordinate is the fragment's $C\alpha$ RMSD to the corresponding residues of the 1UBQ crystal structure. The green lines represent the fragment selections with the optimal overlap similarity. Selections were performed using a bounded tree search algorithm [12] in five blocks (corresponding to windows 1–20, 21–25, 26–35, 36–50, and 51–70) chosen to reduce the running time of the tree search.

rameters have been calibrated against both Poisson-Boltzmann calculations and experimental solvation free energies for a wide range of small organic molecules. The combinatorial optimization of the sidechain rotamer states was performed using the algorithm of Jacobson and Friesner (manuscript submitted), which is an adaptation of the method described by Xiang and Honig [24] and is similar to that used previously by Brucoleri and Karplus [25]. For cases where the backbone structure is assumed to be known (e.g. when testing the algorithm and energy function on high-resolution crystal structures), the algorithm first chooses a rotamer state for each sidechain at random from the highly detailed rotamer library of Xiang and Honig [24]. This library is considerably larger than that used in other rotamer prediction methods, with a resolution of 10° and approximately 2000 rotamer states for Lys and Arg. This library is first pre-screened using hard-sphere overlap as a rejection criterion to prevent the selection of rotamers which lead to steric clash with the protein backbone. Next, the rotamer state of each sidechain

in the protein is optimized one at a time while holding the rotamer states of all the other sidechains constant. This is done by searching the pre-screened rotamer library for the lowest energy rotamer, which was then energy minimized with respect to all of its degrees of freedom in Cartesian space using a novel multiscale minimization algorithm (Jacobson and Friesner, manuscript submitted) based on the Truncated Newton method of Xie and Schlick [26]. This procedure is iterated until no sidechains change rotamer states, and then all sidechains are energy minimized simultaneously in Cartesian coordinates to remove any remaining clashes.

For cases where the backbone is known to be approximate, such as backbones modeled using residual dipolar couplings, a modified version of the above algorithm was used. The rotamer library was first pre-screened to avoid rotamer states which lead to steric clash with the protein backbone. For residues which gave rise to no acceptable rotamer states, the steric screening criterion was relaxed until at least one acceptable rotamer state was found. The iterative

sidechain optimization procedure then proceeded as above, followed by a minimization of the whole protein with respect to the OPLS/SGB energy. The sidechain prediction was then repeated based on this 'relaxed' backbone.

Results and discussion

The results of the dipolar coupling-based database search for ubiquitin are shown in Figure 2, where we plot the RMSD to the ubiquitin crystal structure 1UBQ [27] for each of the 15 smallest Q factor hits in each data window. It is clear that some regions (e.g. windows 13–18, 31–39, and 47–63) have many more false positives than others (e.g. windows 1–12 and 20–30). Some of this variability is easily explained. For example, windows 20–30 correspond to the alpha-helical region of ubiquitin (residues 23–34). Since helices are common and tend to have relatively similar structures on a seven-residue level, there will be many structurally similar fragments in the database, many of which will also have a small Q factor when fit with data derived from a helical structure. Other regions with a large number of false positives correspond to regions in which there is a large amount of missing data, such as windows 31–39, which overlap the data-poor region in the vicinity of proline residues 37 and 38. However, despite the false positives, there is still a sufficient number of true positives for us to expect to be able to construct a reasonably accurate structure.

The use of the bounded tree search algorithm to select one hit for each window which maximizes the overlap similarity with the neighboring windows greatly reduces the impact of the false positives. This can be seen in the green line in Figure 2, which indicates the selections made by the tree search procedure. In general, the selections are quite good, though usually not the best in terms of the RMSD of the fragment to 1UBQ. The only substantial deviation is in the area of windows 31–36, which could be another consequence of the low data density for residues 36–38 mentioned above. The local structure corresponding to each block is generally quite good, as can be seen in the remarkably tight bundle for the fragments corresponding to windows 1–20 (residues 1–26) (Figure 3). The overall structure before refinement is topologically correct but of somewhat poorer quality than the fit in the individual blocks: the RMSD of the mean C^α positions to 1UBQ is 5.94 Å.

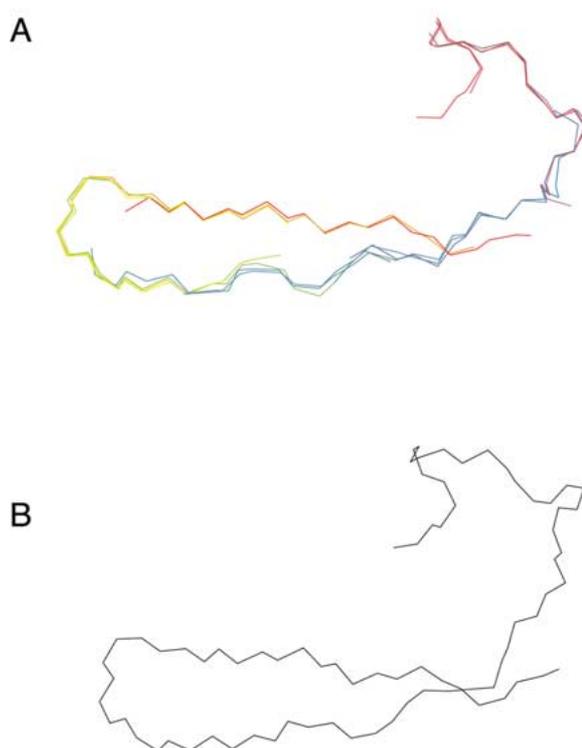


Figure 3. (A) The result of the superposition of the selected fragments for the window block 1–20 (residues 1–26) as described in *Theory and methods*. The 20 fragments chosen are 1SLU:A(69–75), 1BTN(75–81), 1BTN(76–82), 1BTN(77–83), 1TSS(35–41), 1AGQ:A(115–121), 1AGQ:A(116–122), 1DNP:A(67–73), 1DNP:A(68–74), 1RIE(84–90), 1RIE(85–91), 1AGQ:A(44–50), 1AGQ:A(45–51), 1AGQ:A(46–52), 1ALO(15–21), 1ALO(16–22), 1ALO(17–23), 1EXN:A(90–96), 1EXN:A(91–97), and 1HVD(116–122), respectively. Each fragment is shown in a different color. (B) The corresponding region of 1UBQ. The C^α RMSD between it and the mean C^α positions of the fragments shown in (A) is 1.56 Å.

This is not surprising, however, as absolutely no 'long-range' information has been used in the construction of this initial model, i.e. the fragments have been fit only to each other and not to a common ordering frame.

In order to reintroduce long-range dipolar coupling information into the structure, the backbone torsion angles were adjusted so as to minimize the overall Q factor using the structure obtained from the fragment search and overlap filtering as a starting point. During the iterative minimization procedure [12] the overall RMSD relative to 1UBQ decreased from 6.77 to 2.69 Å, while the RMSD for the core of the protein (neglecting the unstructured N and C termini) decreased from 6.20 to 2.40 Å. The resulting final model is shown in Figure 4 along with the best-fit su-

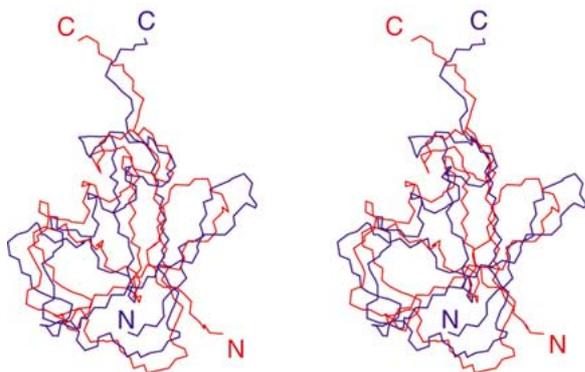


Figure 4. Stereo diagram of the backbone traces of the C^α superposition of the structural model generated using the methodology described in this paper (red) compared with the crystal structure 1UBQ (blue). The N- and C-termini are indicated.

perposition to the crystal structure 1UBQ. While the final model does contain some mis-orientation of secondary structural elements (e.g. the two N-terminal β -sheets), it is still very good. Not only is the backbone fold correct, but the peptide planes and $C^\alpha-C^\beta$ bond vectors also have the correct relative orientations. This can be seen by comparing the RMSD of the final model to 1UBQ using the C^α atoms only (2.69 Å) and using the C^α and C^β atoms together (2.78 Å). Our model is accurate enough to identify the backbone fold (for purposes of structural genomics), and the fact that the sidechain $C^\alpha-C^\beta$ directions are accurately defined raises the possibility of constructing all-atom models using modern sidechain conformation prediction algorithms and molecular mechanics refinement.

We performed sidechain rotamer prediction as described in the *Theory and methods* section above using both the crystallographically determined backbone structure from 1UBQ, as well as the final model determined from the residual dipolar coupling data above. The results using the 1UBQ backbone were very good, with 94% and 87% accuracy rates for χ_1 and χ_2 angles (where ‘correct’ is defined to be within $\pm 40^\circ$ of the dihedral angle observed in the crystal structure), while the χ_3 angles were 60% correct (all RMSD and fractions correct do not include the two N-terminal and two C-terminal residues). When using the dipolar coupling modeled backbone, the error rates were higher, but still reasonably accurate: 62%, 48%, and 56% for χ_1 , χ_2 , and χ_3 , respectively. The accuracy of χ_1 was further analyzed by sidechain type (polar vs non-polar) and location (inside vs. outside, where ‘outside’ is defined as those residues whose

sidechains have more than 20% accessible surface area; Table 1). For both the X-ray and modeled backbones, the interior sidechains were predicted more accurately than the surface sidechains. With respect to amino acid type, the results suggest that the conformation of non-polar sidechains is more accurate than that of the polar ones when using the 1UBQ backbone, but this is no longer the case when the dipolar coupling modeled backbone is used. The increased accuracy of the χ_1 predictions for interior residues likely arises from steric restriction due to the backbone and sidechain packing, and similar results have been seen when modeling surface sidechain conformation in the presence and absence of crystal packing interaction (Jacobson and Friesner, manuscript submitted). In order to investigate the role of protein size on sidechain rotamer prediction accuracy, we repeated our sidechain modeling procedures using the backbone structure of the 370 residue maltodextrin binding protein (1DMB). The resulting overall accuracy rates are comparable to those for ubiquitin, and similarly show increased accuracy for interior and non-polar sidechains (Table 1). A more detailed study of the impact of protein size on sidechain packing and prediction accuracy is currently underway.

In order to evaluate the influence of the backbone accuracy on the sidechain rotamer prediction, we generated a series of perturbed backbone structures using Monte Carlo simulation at high temperature. Specifically, we used internal coordinate Monte Carlo using the OPLS/SGB energy function at 10 000 K on the 1UBQ backbone with sidechains removed. A total of 8 conformations were saved at intervals of 10 Monte Carlo steps, and the C^α RMSD from the crystal structure increased from zero to nearly 4 Å over those 8 conformations. As seen in Figure 5, the sidechain accuracy decreases rapidly as the backbone RMSD increases from 0 to 1 Å, and then remains fairly constant as the RMSD increases further. This result seems to indicate that sidechain prediction accuracy is very sensitive to small deviations from the native structure. However, it is possible that some fraction of this error is due not to distortions in the overall conformation but rather to incorrect $C^\alpha-C^\beta$ vector orientations. To determine the magnitude of this contribution, we performed a Cartesian coordinate Monte Carlo perturbation while holding C^α coordinates fixed at the positions found in the crystal structure while allowing the other backbone atoms to move (using the same energy function for 10 000 steps at 10 000 K). The accuracy of sidechain prediction in this case is

Table 1. A summary of the results for χ_1 sidechain prediction on ubiquitin and maltodextrin binding protein using native, perturbed, and dipolar coupling modeled backbone structures.

	Ubiquitin					Maltodextrin binding protein			
	Number of residues	Fraction correct χ_1				Number of residues	Fraction correct χ_1		
	1UBQ backbone	Dipolar coupling backbone	Perturbed 1UBQ backbone (0.5Å RMSD)	Perturbed 1UBQ backbone with fixed C^α		1DMB backbone	Perturbed 1DMB backbone (0.4Å RMSD)	Perturbed 1DMB backbone with fixed C^α	
Overall	66	94%	62%	70%	83%	297	83%	72%	76%
Inside	19	100%	74%	89%	95%	137	91%	80%	83%
Outside	47	91%	58%	62%	76%	160	76%	65%	71%
Polar	38	87%	63%	63%	84%	156	76%	67%	71%
Non-polar	28	100%	61%	79%	82%	141	91%	79%	82%

See the text for details. A χ_1 angle is considered to be correct if it is within $\pm 40^\circ$ of the rotamer conformation seen in the corresponding X-ray structure.

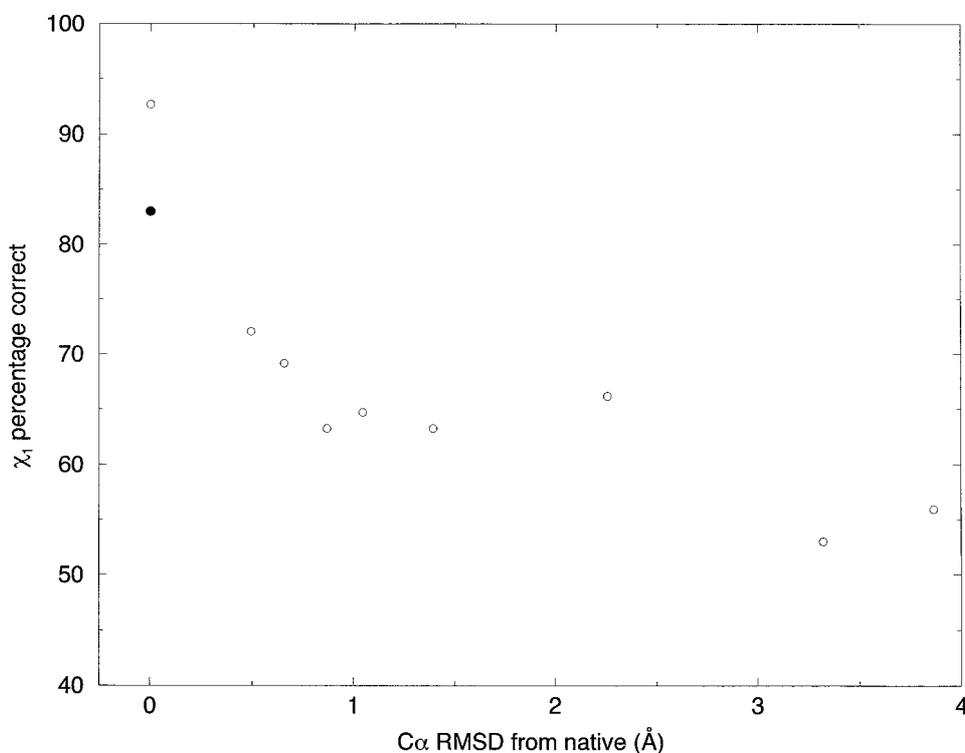


Figure 5. Results of sidechain χ_1 rotamer prediction for ubiquitin using both the native and perturbed backbone structures. The χ_1 percentage correct (where 'correct' is defined to be within $\pm 40^\circ$ of the native rotamer) is plotted against the C^α RMSD from the crystal structure. The filled square indicates the fraction correct using the crystal structure backbone. The open circles represent the perturbations of the entire backbone using high temperature Monte Carlo as described in the text. The filled circle indicates the result using a backbone structure in which all atoms except the C^α have been perturbed.

decreased to 83%, despite the fact that its C^α RMSD from the crystal structure is zero (filled circle in Figure 5). Therefore, C^α - C^β orientational errors likely

make a significant contribution to the decrease in sidechain χ_1 accuracy in going from the X-ray backbone to 1.0 Å RMSD. Since protein backbone mod-

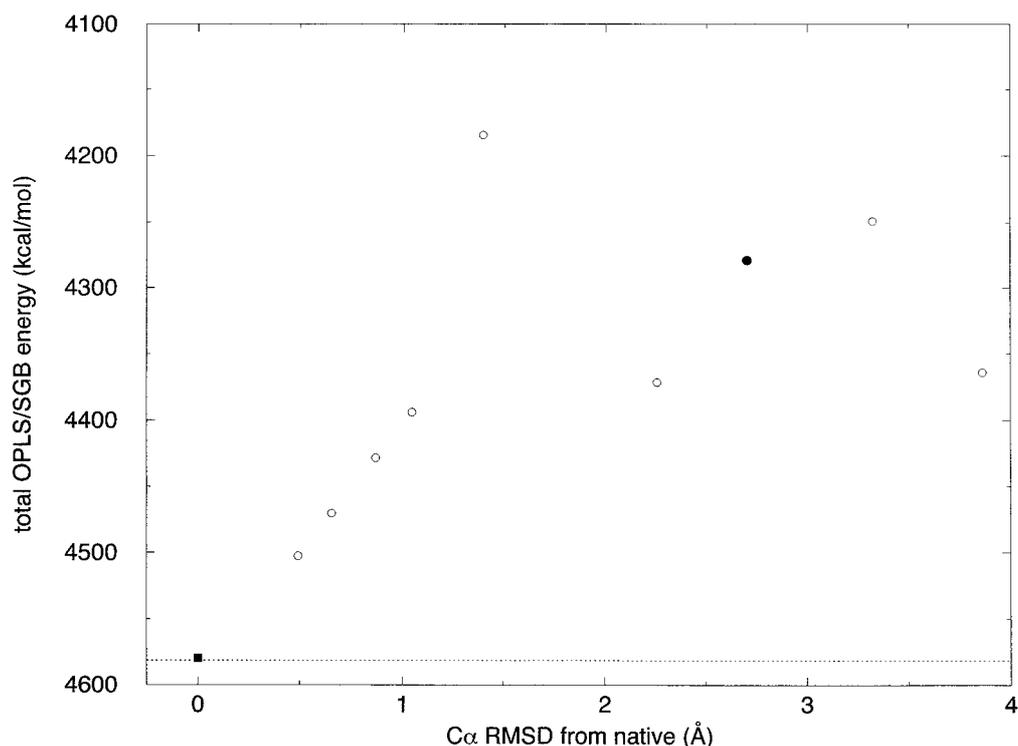


Figure 6. The total OPLS/SGB energy of the structures generated using sidechain rotamer prediction based on the native and perturbed ubiquitin backbones. The dotted line indicates the energy of the 1UBQ structure after minimization. The filled square indicates the energy of the complete structure after sidechain placement using the crystal structure backbone. The open circles represent the molecular mechanics energies of the complete structures after sidechain placement onto backbones perturbed using high temperature Monte Carlo as described in the text and Figure 5. The filled circle indicates the energy of the complete structures after sidechain placement onto backbone determined using residual dipolar couplings.

els derived from residual dipolar coupling data have generally correct $C^\alpha-C^\beta$ orientations, it is likely that further improvements in backbone model building using dipolar coupling data will lead to a greater increase in sidechain prediction accuracy than suggested by the Monte Carlo perturbation results shown in Figure 5.

Figure 6 shows the correlation between the RMSD from the crystal structure of the backbone and the total OPLS/SGB molecular mechanics energy of the all atom model generated in the high temperature Monte Carlo simulations, and also including the model based on the residual dipolar coupling data. The energy difference between the minimized X-ray structure (dotted line) and the model with sidechains predicted using the X-ray backbone is only 1.7 kcal/mol. The energy of the protein increases almost linearly for small deviations from the x-ray structure, reaching a plateau beyond 1.0 Å. This result implies that minimization of the molecular mechanics energy can result in improvements in the backbone conformation.

We have in fact observed this in sidechain modeling using the backbone structure determined using dipolar coupling data. In this case, the addition of sidechains results in steric clashes which are removed by energy minimization, causing the RMSD over core residues to improve from 2.4 to 2.2 Å. These results provide a demonstration of the synergies that are possible between protein structural database and molecular modeling methods and NMR spectroscopy, and we expect that the further development of these methods will lead to the extraction of high-resolution structural information from minimal NMR data.

Conclusions

We have shown that a protein fragment database search approach using overlap RMSD as a filtering tool is an efficient way of generating a backbone fold from residual dipolar coupling data. Furthermore, the resulting structures are of sufficient accuracy to allow

the placement of sidechains using a rotamer prediction algorithm. There is a great deal of information contained in the residual dipolar couplings, not all of which is used by the methods described here, and the overall reliability and robustness of our procedure could be improved by incorporating this information. For example, we used the dipolar couplings to assemble the list of candidate fragments at each window position, but this information was not used at all in the fragment filtering and the construction of the initial model. Furthermore, molecular mechanics methods could be incorporated earlier into the structure determination process, as it is routinely done in NOE-based NMR structure determination. Also, the addition of chemical shift and scalar coupling data is likely to greatly increase the accuracy of the resulting models.

Acknowledgements

This research was supported by the National Institutes of Health (NRSA Fellowship GM19856-02 to MA and grant GM-30580 to RML).

References

- Moseley, H.N.B., and Montelione, G.T. (1999) *Curr. Opin. Struct. Biol.* **9**, 635–642.
- Prestegard, J.H., Tolman, J.R., Al-Hashimi, H.M., and Andrec, M. (1999) In *Structure Computation and Dynamics in Protein NMR*, (Eds., Krishna, N.R., and Berliner, L.J.) Plenum Publishers, New York. pp. 311–355.
- Tjandra, N. and Bax, A. (1997) *Science* **278**, 1111–1114.
- Hansen, M.R., Mueller, L., and Pardi, A. (1998) *Nature Struct. Biol.* **5**, 1065–1074.
- Clore, G.M., Starich, M.R., and Gronenborn, A.M. (1998) *J. Am. Chem. Soc.* **120**, 10571–10572.
- Tjandra, N. (1999) *Structure* **7**, R205–R211.
- Andrec, M., Du, P., and Levy, R.M. (2001) *J. Am. Chem. Soc.* **123**, 1222–1229.
- Annala, A., Aitio, H., Thulin, E., and Drakenberg, T. (1999) *J. Biomol. NMR* **14**, 223–230.
- Delaglio, F., Kontaxis, G., and Bax, A. (2000) *J. Am. Chem. Soc.*, **122**, 2142–2143.
- Hus, J.-C., Marion, D., and Blackledge, M. (2001) *J. Am. Chem. Soc.* **123**, 1541–1542.
- Meiler, J., Peti, W., and Griesinger, C. (2000) *J. Biomol. NMR* **17**, 283–294.
- Andrec, M., Du, P., and Levy, R.M. (2001) *J. Biomol. NMR* **21**, 335–347.
- Saupe, A. (1968) *Angew. Chem. Internat. Ed.* **7**, 97–112.
- Losonczi, J.A., Andrec, M., Fischer, M.W.F., and Prestegard, J.H. (1999) *J. Magn. Reson.* **138**, 334–342.
- Cornilescu, G., Marquardt, J.L., Ottiger, M., and Bax, A. (1998) *J. Am. Chem. Soc.* **120**, 6836–6837.
- Brenner, S.E., Chothia, C., and Hubbard, T.J.P. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 6073–6078.
- Ottiger, M., and Bax, A. (1998) *J. Am. Chem. Soc.* **120**, 12334–12341.
- Shindyalov, I.N., and Bourne, P.E. (1998) *Prot. Eng.* **11**, 739–747.
- Kabsch, W. (1978) *Acta Cryst. A* **34**, 827–828.
- Moltke, S., and Grzesiek, S. (1999) *J. Biomol. NMR* **15**, 77–82.
- Jorgensen, W.L., Maxwell, D.S., and Tirado-Rives, J. (1996) *J. Am. Chem. Soc.* **118**, 11225–11236.
- Kaminski, G.A., Friesner, R.A., Tirado-Rives, J., and Jorgensen, W.L. (2001) *J. Phys. Chem. B* **105**, 6474–6487.
- Ghosh, A., Rapp, C.S., and Friesner, R.A. (1998) *J. Phys. Chem. B* **102**, 10983–10990.
- Xiang, Z., and Honig, B. (2001) *J. Mol. Biol.* **311**, 421–430.
- Bruccoleri, R.E., and Karplus, M. (1987) *Biopolymers* **26**, 137–168.
- Xie, D., and Schlick, T. (1999) *SIAM J. Optim.* **10**, 132–154.
- Vijay-Kumar, S., Bugg, C.E., and Cook, W.J. (1987) *J. Mol. Biol.* **194**, 531–544.