Journal of Biomolecular NMR, **14:** 223–230, 1999.
KLUWER/ESCOM
© 1999 *Kluwer Academic Publishers. Printed in the Netherlands.*

Recognition of protein folds via dipolar couplings

Arto Annila[a], Helena Aitio[b], Eva Thulin[c] & Torbjörn Drakenberg[a,c]
[a]*VTT Chemical Technology, P.O. Box 1401, FIN-02044 VTT, Finland;* [b]*Institute of Biotechnology, P.O. Box 56, FIN-00014 University of Helsinki, Finland;* [c]*Department of Physical Chemistry 2, University of Lund, P.O. Box 124, 221 00 Lund, Sweden*

Received 27 January 1999; Accepted 3 May 1999

Key words: alignment, bicelle, calcium-binding proteins, calerythrin, dipolar coupling, protein fold

Abstract

Alignment of proteins in dilute liquid crystalline medium gives rise to residual dipolar couplings which provide orientational information of vectors connecting the interacting nuclei. Considering that proteins are mainly composed of regular secondary structures in a finite number of different mutual orientations, main chain dipolar couplings appear sufficient to reveal structural resemblance. Similarity between dipolar couplings measured from a protein and corresponding values computed from a known structure imply homologous structures. For dissimilar structures the agreement between experimental and calculated dipolar couplings remains poor. In this way protein folds can be readily recognized prior to a comprehensive structure determination. This approach has been demonstrated by showing the similarity in fold between the hitherto unknown structure of calerythrin and sarcoplasmic calcium-binding proteins from *Nereis diversicolor* and *Branchiostoma lanceolatum* with known crystal structures.

Abbreviations: CSA, chemical shift anisotropy; D6PC, dihexanoyl phosphatidylcholine; D7PC, diheptanoyl phosphatidylcholine; DMPC, dimyristoyl phosphatidylcholine; INEPT, insensitive nuclei enhancement by polarization transfer; NOE, nuclear Overhauser enhancement; PDB, Protein Data Bank entries 1cll: calmodulin, 1rec: recoverin, 1tco: serine-threonine phosphatase B_2, 2bbm: calmodulin complexed with myosin light chain kinase, 2sas: sarcoplasmic calcium-binding protein from an amphioxus *Branchiostoma lanceolatum*, 2scp: sarcoplasmic calcium-binding protein from a sandworm *Nereis diversicolor*; TROSY, transverse relaxation optimized spectroscopy.

Introduction

The human genome project and other ventures to map genes of a single species progress so rapidly that the post-genomic era will be encountered sooner than was anticipated only a decade ago. Consequently, the emphasis will shift from DNA sequencing to protein structure determination. 'Structural genomics', i.e. the determination of all structures of all gene products from a single species, is envisioned to provide further insight into biological pathways of the whole organism. However, presently we lack powerful means to determine structures in large numbers, i.e. in hundreds of thousands, of potential interest to mankind. Therefore the focus will remain either on proteins of a particular biological, medical or biotechnical importance or on proteins with novel folds. It is estimated that there may be only on the order of one thousand protein folds. This comparatively small 'fold basis set' implies that merely a minor fraction of information embedded in a primary structure is required to govern a fold. This is also the conclusion from random mutagenesis studies. Our understanding of factors determining a fold remains impaired, and severely constricts prospects of automated homology based modeling as long as there are not representative structures for the majority of the folds (Sanchez and Sali, 1997, 1998). Based on a remote amino acid sequence homology alone it is

*To whom correspondence should be addressed. E-mail: arto.annila@vtt.fi

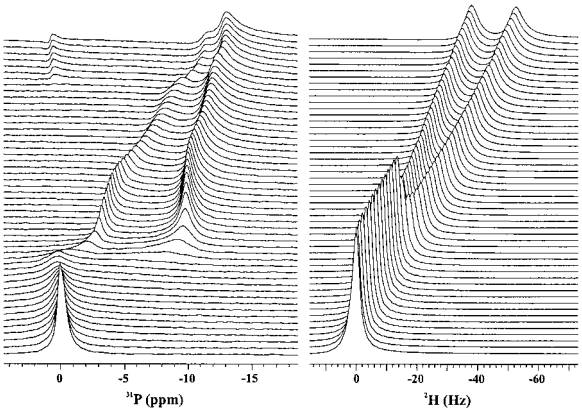

Figure 1. Series of ^{31}P (left) and ^{2}H (right) spectra from a DMPC/D7PC (3:1) mixture as a function of temperature. The spectra were acquired with 1 °C intervals starting from 10 °C up to 60 °C. The ^{31}P signal of the phosphocholine head groups was referenced to 0 ppm at 10 °C. The transition to the liquid crystalline phase at about 27 °C is identified by the appearance of shifted ^{31}P signals. The quadrupolar splitting of deuterium from partially aligned D_2O is observed in the liquid crystalline phase. The deuterium spectra were acquired unlocked as seen from the frequency drift.

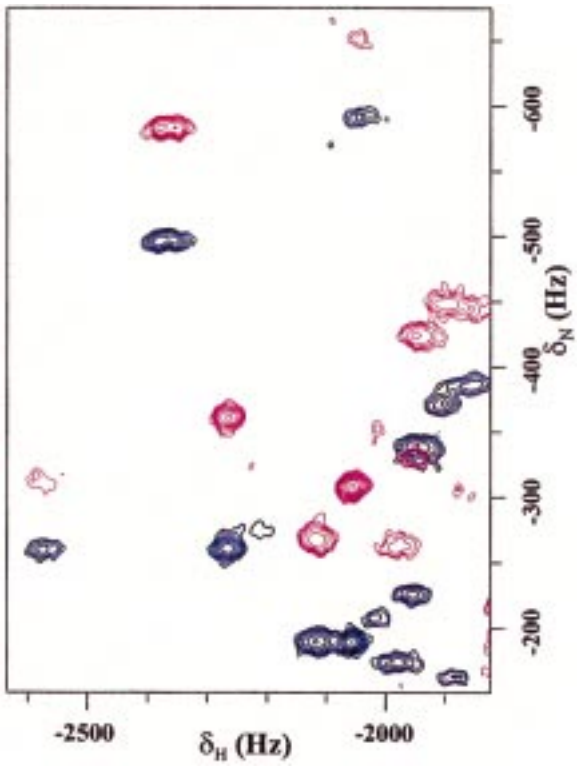

Figure 2. Two overlaid expansions of spin state selective [^{15}N-^{1}H] correlation spectra acquired at 45 °C of calerythrin dissolved in the DMPC/D7PC (3:1) liquid crystalline phase. There is a large variation in the heteronuclear splittings due to the dipolar contribution. Intensity variation among residues depends mainly on the transfer efficiency via INEPT tuned to 93 Hz. In addition, high-field multiplet components are weaker than the low-field components due to cross correlation effects between dipolar and CSA interactions.

intricate to judge a priori if a fold of a protein is new. A means for a quick characterization of a fold to evaluate the need of proper structure determination by NMR spectroscopy or X-ray crystallography would be most useful.

We have considered the use of residual dipolar couplings which arise from enhanced anisotropic tumbling of proteins in dilute liquid crystalline medium (Tjandra and Bax, 1997) for the recognition of protein folds. Taking into account the finite basis set of protein folds, main chain dipolar couplings might be sufficient to prove or exclude similarity between relative spatial organization of secondary structures. We exemplify this approach for a 20 kDa EF-hand calcium-binding protein called calerythrin from a gram-positive bacterium *Saccharopolyspora eryhtraea* (Swan et al., 1987).

Materials and methods

Calerythrin was expressed uniformly ^{13}C and ^{15}N labeled in *E. coli* K38 (pCBM1, pG-P1-2) grown in M9 minimal medium with NH_4Cl (99.1% ^{15}N) and $^{13}C_6$-d-glucose (99% ^{13}C) (Merck Bioscience) as sole carbon and nitrogen sources. The protein was purified by ion exchange and gel-filtration chromatography (a more detailed description will be given elsewhere). Resonances of HN, N, CA and CB were assigned from HNCACB (Muhandiram and Kay, 1994) and HN(CO)CACB (Yamazaki et al., 1994) spectra acquired at 45 °C from a 1 mM Ca^{2+}-loaded calerythrin sample. The liquid crystalline medium (5% w/v) was composed of dimyristoyl phosphatidylcholine (DMPC) and diheptanoyl phosphatidylcholine (D7PC) approximately in a 3:1 molar ratio. The preparation and properties of this solution for the purpose of enhancing anisotropic tumbling are quite similar to those of DMPC/D6PC (Sanders and Prestegard, 1990; Sanders and Schwonek, 1992; Bax and Tjandra, 1997; Ottiger and Bax, 1998). The anisotropic phase forms at about 27 °C as deduced from the disappearance of the ^{31}P signal referenced to 0 ppm and from the concomitant appearance of shifted signals of the phosphatidyl groups. Another phase transition takes place just above 50 °C seen as reappearance of the ^{31}P signal at 0 ppm (Figure 1). ^{15}N-labelled calerythrin was

dissolved in the DMPC/D7PC solution to 0.3 mM concentration. Heteronuclear splittings, denoted here as T_{HN}, composed of scalar couplings ($^1J_{HN} \approx -94$ Hz) and residual ^{15}N-^{1}H dipolar couplings (D_{HN}) were measured at 45 °C from multiplet components of ^{15}N-^{1}H correlations. A generalized and sensitivity enhanced TROSY sequence with ^{15}N decoupling during acquisition was employed in this study as a spin state selective α/β-filter to alleviate spectral crowding (Pervushin et al., 1997, 1998; Andersson et al., 1998; Czisch and Boelens, 1998; Weigelt, 1998; Cordier et al., 1999). Two WATERGATE (Piotto et al., 1992) elements implemented in this pulse sequence provided satisfactory water suppression.

Results and discussion

The spin state selective ^{15}N-^{1}H correlation spectrum of calerythrin in the DMPC/D7PC liquid crystalline phase displayed a large dispersion of T_{HN} from −45 to −146 Hz (Figure 2). This excessive degree of alignment was observed also as broadening or splitting of some correlations along the ^{1}H dimension due to homonuclear proton dipolar couplings. In this particular case our attempts to reduce the dipolar contribution by decreasing the concentration of bicelles resulted in loss of residual dipolar couplings. Nevertheless, coherence transfer via INEPTs tuned to 93 Hz was sufficient to detect 94% of the assigned HN signals. Values of T_{HN} and C_α secondary chemical shifts (Wishart and Sykes, 1994) as a function of the amino acid sequence display concurrent changes and reveal that calerythrin is composed of α helices at various orientations interrupted by loops and turns (Figure 3). For each helix there is a comparatively small variation in T_{HN} values because H-N bond vectors are approximately along the axis of the helix, whereas for the loops and turns there is a dispersion of T_{HN} due to the distribution of consecutive H-N bond vectors in various directions. The calcium-binding loops are further distinguished from the turns by the larger number of residues involved. Values of D_{HN} were obtained by subtracting from T_{HN} scalar couplings ($^1J_{HN}$) measured from a reference spectrum recorded in water. In spite of the unnecessary high degree of alignment the anisotropic contribution of $^1J_{HN}$ to T_{HN} was assumed to be negligible (Lounila and Jokisaari, 1982).

The described procedure illustrates that it takes only a small effort to extract backbone assignments and sequential H-N residual dipolar couplings. This piece of information, nevertheless, enables a search for similar folds among known protein structures. In the case of calerythrin we limited, for obvious reasons, the search to the members of the EF-hand family of proteins. According to the primary structure calerythrin shares 27% amino acid identity with a sarcoplasmic calcium-binding protein (SCP) from a sandworm *Nereis diversicolor* (Vijay-Kumar and Cook, 1992), 17% with calmodulin (Babu et al., 1988; Chattopadhyaya et al., 1992; Ikura et al., 1992), 13% with the B_2 domain of a serine-threonine phosphatase commonly known as calcineurin B (Griffith et al., 1995), 12% with recoverin (Flaherty et al., 1993) and 15% with another SCP from an amphioxus *Branchiostoma lanceolatum* (Cook et al., 1993). The first four of these proteins were chosen to represent different types of folds within the EF-family. The additional SCP from the amphioxus with a lower sequence identity to calerythrin was chosen to conclude if the main chain dipolar data show even minor differences between similar folds. Two calmodulin conformations were inspected, in its $4 \cdot Ca^{2+}$-loaded form (Babu et al., 1988; Chattopadhyaya et al., 1992) and in complex with a peptide from myosin kinase (Ikura et al., 1992) to see if also conformations can be discerned.

The objective was to compute from each known structure 'model dipolar data' as a function of amino acid sequence for the comparison with the data measured from calerythrin. The proper computation would require knowledge of the molecular alignment tensor *D*, which is intrinsically unknown. In order to circumvent this difficulty we assumed that for a known structure resembling calerythrin the value of the asymmetry parameter (*R*) of calerythrin, i.e. the ratio of rhombic ($D_\perp$) to axially symmetric ($D_\parallel$) component of *D*, could be used without introducing significant error. The distribution of dipolar couplings provides estimates of *R* and $D_\parallel$ (Clore et al., 1998a). In particular for an α helical protein the sampling of directions by H-N vectors is quite incomplete and only estimates of R (0.5 ± 0.1) and $D_\parallel$ (27 ± 3 Hz) could be obtained. For any known structure dissimilar to calerythrin, we reasoned that the use of the experimental *R* and $D_\parallel$ would inevitably result in an apparent discrepancy between the experimental and computed values which ascertains that the folds are indeed different. Because of the lack of the alignment tensor directions a search for the orientation of the coordinates is required. To avoid a possible trapping to local minima we simply devised a least square grid-search to produce the smallest root mean square deviation (rmsd) between

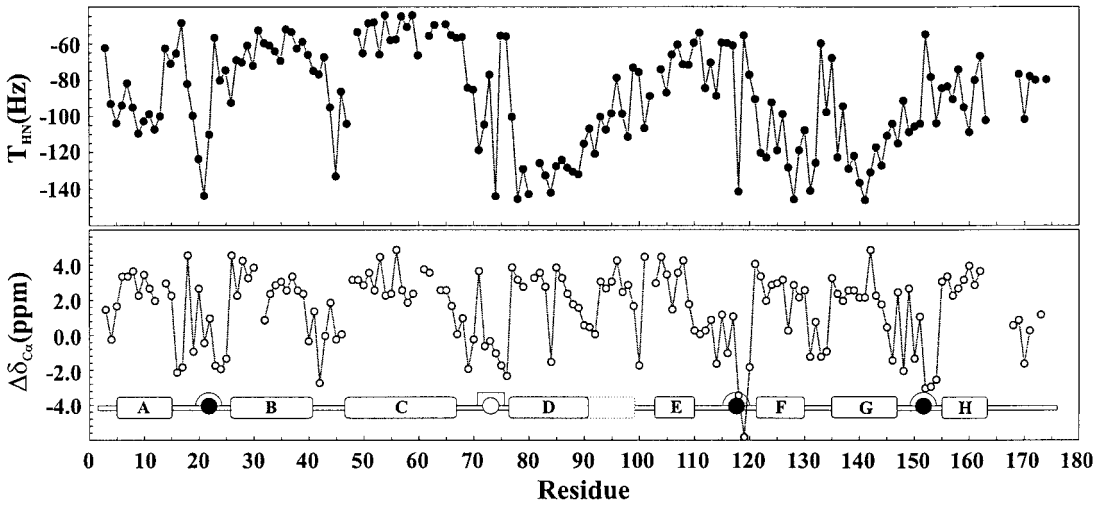

Figure 3. Heteronuclear H-N splitting (above) and C_α secondary chemical shifts (below) as a function of the amino acid sequence of calerythrin. The heteronuclear splitting is negative because of the negative gyromagnetic ratio for ^{15}N. H-N correlations for residues 1–2, 163–167 and 174–175 were not observed in the isotropic phase due to exchange phenomena. Residue 102 is a proline. The α helices, three calcium-binding sites, defunct site and turns, deduced from NOEs, are indicated schematically. Based on these data the length of the D helix is controversial.

the experimental and computed dipolar data, analogous to previously reported comparisons of measured dipolar data with corresponding values computed from the *same* known protein structure (Bax and Tjandra, 1997; Tjandra and Bax, 1997; Tjandra et al., 1997; Yang et al., 1998).

The comparison of the experimental dipolar couplings with the values computed from homologous structures is complicated because the lengths of secondary structures vary among proteins with the same fold and in particular, loops may differ considerably. An alignment of primary structures based on amino acid sequence homology qualifies for an initial guess of the equivalent segments. A more accurate alignment of secondary structures is arrived at by taking into account C_α secondary chemical shifts at hand from the backbone assignment. If no sequence homology is discernible, simple brute-force algorithms, to compare stretches of continuous amino acid sequences can be devised. Also information on the protein function, if available, could be incorporated into the initial search of 'the computational reading frame'. To recognize similarity in the case of many deletions or insertions requires more advanced and computationally intensive algorithms, analogous to those used in the sequence homology searches. On the other hand, many discontinuities in the aligned sequences themselves indicate dissimilarity. Finally, a similar structure is recognized from a low rmsd value between the measured and computed values.

Results of the computations show unambiguous similarity between the measured dipolar couplings of calerythrin and the calculated values of the two sarcoplasmic calcium-binding proteins from sandworm and amphioxus (Figure 4). The difference between the rmsd values of the two SCPs based only on D_{HN} is insignificant to argue for a distinction. We find a good match for D_{HN} of residues in the helices from B to G as well as for most residues of the intervening turns and loops. The N-terminal helix A and helix H near the C-terminus of calerythrin are in slightly different orientations compared to their counterparts in the homologous SCPs. We also notice a slight mistake in the alignment of sequences from calerythrin and SCP of the amphioxus in the loop between the B and C helices. This was corrected subsequently. The rmsd values in the vicinity of the global minimum indicate that gradients of the error function with respect to angles are small and negative. Therefore, a coarse grid-search for Euler angles with 5–15 degree intervals provides initial parameters for a convergent algorithm to accelerate the computation.

For the other structures the overall discrepancy between the measured D_{HN} of calerythrin and the computed D_{HN} is evident. The calcium-loaded conformation of calmodulin is dissimilar to calerythrin. Only the orientation of the first EF-hand comprising the A and B helices bears a resemblance. The myosin kinase bound conformation of calmodulin is also dissimilar to calerythrin. The value of rmsd is comparable to

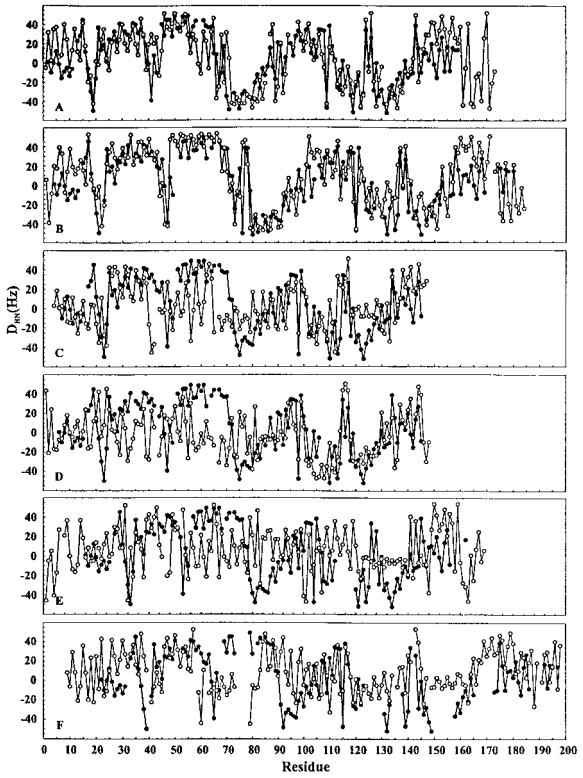

Figure 4. Comparisons of measured dipolar couplings (solid circles) of calerythrin and dipolar couplings computed (open circles) from X-ray structures of sarcoplasmic calcium-binding protein from the sandworm (rmsd 17.9 Hz) (A), SCP from the amphioxus (17.4) (B), $4 \cdot Ca^{2+}$-loaded (24.3) (C) and myosin kinase bound conformation of calmodulin (24.8) (D), B_2 domain of a serine-threonine phosphatase (27.7) (E) and recoverin (25.9) (F). The asymmetry parameter (R) and the axial component ($D_\parallel$) of the alignment tensor were estimated from the distribution of the measured dipolar couplings of calerythrin and were used in the grid-search of the orientation of coordinates which minimizes the rmsd between the measured and computed values. Only equivalent segments based on multiple sequence alignment were compared. The numbering of residues corresponds to the target sequence.

the $4 \cdot Ca^{2+}$-loaded conformation, but the agreement is in this case slightly better for the C-terminal half of calmodulin than for the N-terminus. As expected already from biological grounds, the B_2 domain of the serine-threonine phosphatase is quite unlike calerythrin. For recoverin several segments of distant amino acid sequence homology line up. The discontinuities make the comparison complex but the overall dissimilarity is unambiguous, although some of the fragments in the C-terminal half bear resemblance. In general it is important that several structures, preferably both similar and dissimilar, are submitted for the comparison in order to establish a relative scale.

For the recognition of the similar structures and for the rejection of the dissimilar structures the uncertainty in R and $D_\parallel$ was found to be inconsequential. Additional grid-searches as function of the asymmetry ($0.2 \leq R \leq 0.65$; with increment 0.05) and axially symmetric component ($20 \leq D_\parallel \leq 30$ Hz; 1 Hz) were performed. For all structures no local minima within the grid for R and $D_\parallel$ were found and Euler angles corresponding to the least square sums at each node of the grid varied typically less than 30 degrees. For the SCPs of the sandworm and the amphioxus the least square sums were within 5% whereas the corresponding values for the SCP of the sandworm were from 40 to 70% smaller than those of calmodulin, the next similar. Therefore, the SCPs were unambiguously and reliably distinguished as similar to calerythrin from the other structures by rmsd, irrespective of the uncertainty in the estimates of R and $D_\parallel$.

The distribution of the least squares in addition to the values of rmsd showed differences between the similar and dissimilar structures (Figure 5). For the similar SCP structures the computed values agree very well with the experimental data of calerythrin for the majority of residues and there are only comparatively few residues with large deviations. This can be understood as that the secondary structures closely match and contribute comparatively little to the total error, whereas the turns differ and mainly cause large errors. For the dissimilar structures the distribution of the least squares extends with amplitude to larger values. This can be interpreted as that even for the best overall agreement the mutual orientations of the secondary structures are not similar to calerythrin and errors accumulate from many residues. Therefore the distribution of the least squares can be used as an additional criterion for the recognition of similar structures.

The differences in the distribution of the least squares between the similar and dissimilar structures become more pronounced when the residues with error values larger than three standard deviations are removed from the comparisons. Then the grid-search as a function of R and $D_\parallel$ showed, for example, that the least square sum for the SCP of the amphioxus was from 5 to 20% smaller than that of the SCP from the sandworm and from 70 to 120% smaller than that of calmodulin. Consequently, the distribution can be used to refine the comparison between the structures. The conclusion is that similarity among proteins composed mainly of regular secondary structures will be easier to notice than similarity among irregular structures.

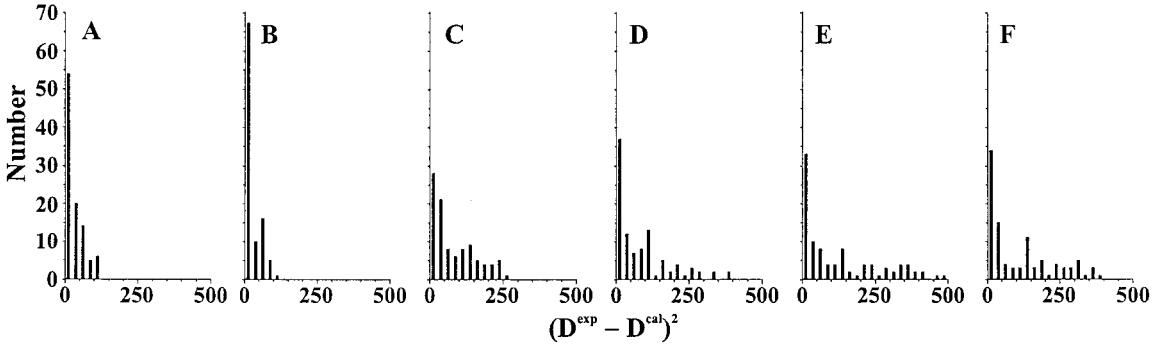

Figure 5. Distributions of the error squared between the measured dipolar coupling and corresponding value computed from the X-ray structures of sarcoplasmic calcium-binding protein from the sandworm (standard deviation 202) (A), SCP from the amphioxus (245) (B), $4 \cdot Ca^{2+}$-loaded (332) (C) and myosin kinase bound conformation of calmodulin (312) (D), B_2 domain of a serine-threonine phosphatase (370) (E) and recoverin (356) (F).

The close similarity of calerythrin to the sarcoplasmic calcium-binding protein from the amphioxus is by no means obvious from the amino acid sequence homology alone. When sequence identity falls below 30% the probability for misalignment increases rapidly. Initially calerythrin was even thought to resemble calmodulin (Swan et al., 1987). Knowledge of the fold immediately implies that calerythrin sequesters calcium rather than is involved in the calcium-mediated signaling. An approximate impression of the fold of calerythrin is provided by the ribbon presentation of the X-ray structure of SCP from the amphioxus, color coded with the values of dipolar couplings (Figure 6). The overall structural similarity between calerythrin and SCPs is independently supported by NOEs found between short β strands of adjacent EF-hands.

For homologous structures it would of course be tempting to use the differences between the measured and computed dipolar couplings to expose also small structural differences which are likely to account for the variance in biological function. For the dipolar coupling between nuclei A and B:

$$D_{AB} = D_{\parallel}[(3\cos^2\theta - 1) + {}^3/_2 R \sin^2\theta \cos 2\phi] \quad (1)$$

the error is non-linear, as explicitly seen from the partial derivatives

$$\partial D_{AB}/\partial\theta = 3D_{\parallel} \sin 2\theta[-1 + {}^3/_2 R \cos 2\phi] \quad (2)$$

$$\partial D_{AB}/\partial\phi = -3D_{\parallel} R \sin^2\theta \sin 2\phi. \quad (3)$$

For example, D_{HN} of vectors approximately collinear with the long axis of the alignment tensor ($\theta \approx 0$) are insensitive to small differences in the θ angle. This trouble of non-linear gradients can be partly alleviated provided that additional dipolar couplings between amide proton or nitrogen and adjacent carbons are available from ^{15}N and ^{13}C labeled sample. D_{HNCO} and D_{NCO} are readily obtained from a [^{15}N,^{1}H] correlation spectrum (Ottiger et al., 1998) and also D_{CACO} without explicit carbon assignments. In the comparison of the measured D_{HN} and D_{HNCO} to the values computed from an X-ray structure, peptide bond planarity is implicitly assumed. This assumption is not rigorously true and the uncertainty in the position of HN contributes to the imprecision (Ramirez and Bax, 1998). Non-redundant D_{HN} could also be extracted from a protein aligned alternatively, e.g., in another type of liquid crystalline medium (Clore et al., 1998b; Hansen et al., 1998; Losonczi and Prestegard, 1998; Wang et al., 1998) or by introducing tags to the protein to influence its alignment (Ramirez and Bax, 1998). Several non-redundant dipolar couplings per residue will improve the recognition of the fold and facilitate a coarse comparison of homologous structures. The interest in small differences is already close to the refinement of structure by dipolar couplings (Clore et al., 1998c).

Further uncertainties in the comparison of the measured and computed dipolar couplings arise from dynamics. In this study, internal motions, for example described by values of the generalized order parameter *S*, have been presumed to be uniform and included in the definition of $D_{\parallel}$. This may well account for the discrepancy between the measured and computed dipolar couplings for the residues at the N- and C-terminus of calerythrin.

The recognition of protein folds as described above for an α helical protein should be suitable for characterizing also α/β and β structures. There should be

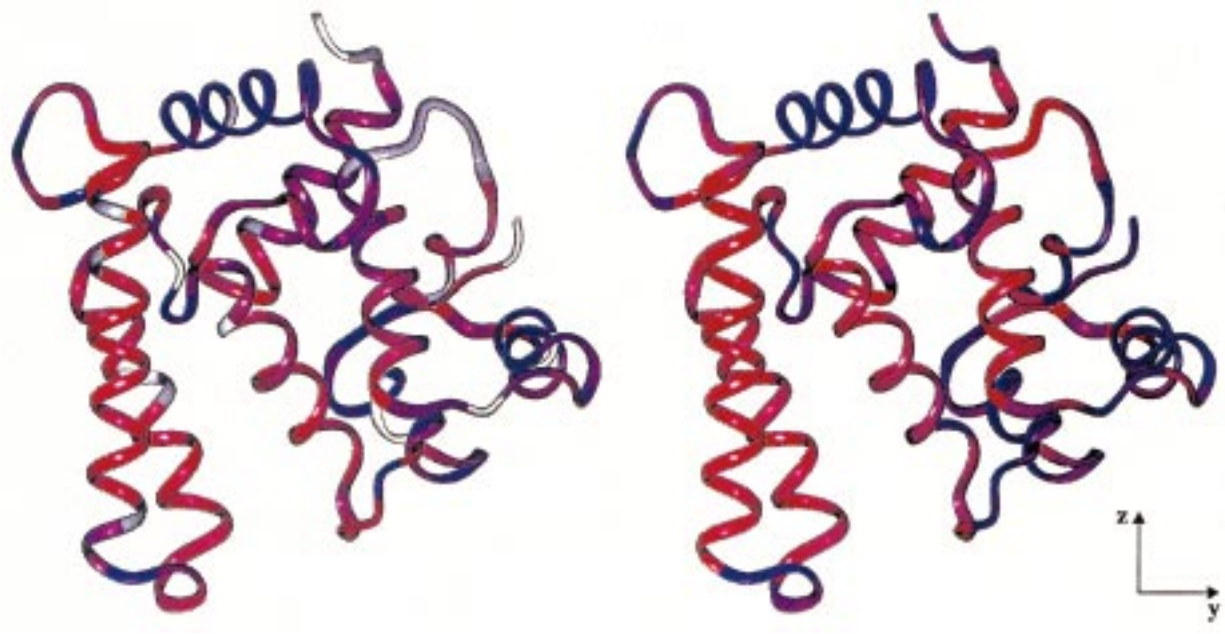

Figure 6. X-ray structure of the sarcoplasmic calcium-binding protein from amphioxus, color coded by the values of measured dipolar couplings of calerythrin (left) and by the corresponding values calculated from SCP of amphioxus (right). Blue refers to negative and red to positive dipolar coupling. The residues of SCP of amphioxus without a correspondence in calerythrin due to insertions are not colored (white). Those residues for which no data were measured are colored grey. Differences between the measured and computed values, seen as different tones of colors for the corresponding residues, arise from structural differences. In addition, differences in internal motions, which have not been considered in this study, will result in differences between the measured and computed values. Possibly some differences, in particular at loops and turns, could be attributed to misalignment of the amino acid sequences by one or two residues.

no difficulties in identifying closed α/β barrels, i.e. TIM barrels or open twisted α/β structures. However, diversity of topologies in β sheet structures might complicate the recognition of a fold. The β sheets are usually twisted and differences in the degree of twist between structures otherwise similar might obscure the recognition. To search for similar protein domains seems also feasible provided that non-redundant internuclear directions are abundant. It is to be expected that the recognition of structural similarities among small proteins or peptides, often with irregular structures, is likely to be difficult, at least if only ^{15}N-^{1}H dipolar couplings are available. Hence the method is most applicable to proteins in which residues are mainly in secondary structures.

It is of general importance that the comparison of measured dipolar couplings and values computed from homologous structures not only proves similarity but also indicates misalignments of primary structures to provide a solid foundation for the homology based modeling. This piece of information would be indispensable for the success of automated modeling of proteins because none of the algorithms presently known will recover from an incorrect alignment (Sanchez and Sali, 1997). For the recognition of folds and domains a network server linked to the protein structure and sequence data banks with programs to combine information of amino acid sequence homology, dipolar couplings, chemical shifts and three-dimensional structure seems practical.

It appears that dipolar couplings would be powerful enough to distinguish between conformations which differ from each other by mutual orientation of secondary structures. This property can be exploited in studies of complexation to show, without an extensive structure determination, which parts of the protein are subject to conformational changes. In this way valuable information, complementary to surface contact maps deduced from chemical shift changes, can be extracted. Thus a modest experimental effort to record dipolar couplings will give significant insight into biology.

Conclusions

This study demonstrates that main chain residual dipolar couplings provide means to recognize the fold of a protein without an explicit structure determination. Similarity between measured dipolar couplings and corresponding values computed from a known structure implies the same fold. Also similarities among domains can be examined. For the purpose of homology based modeling the alignment of amino acid sequences of homologous proteins can be refined with the aid of dipolar couplings. Proteins which are found to display sequential dipolar data dissimilar to computed data from any representative structure may contain a novel fold. Considering that only one thousand representative structures may constitute the fold basis set, similarity searches are feasible also in the future.

Acknowledgements

We thank Ad Bax for valuable advice. This work was supported by a grant from the Academy of Finland and by a graduate school program from the Ministry of Education of Finland.

References

Andersson, P., Annila, A. and Otting, G. (1998) *J. Magn. Reson.*, **13**, 364–367.

Babu, Y.S., Bugg, C.E. and Cook, W.J. (1988) *J. Mol. Biol.*, **204**, 191–204.

Bax, A. and Tjandra, N. (1997) *J. Biomol. NMR*, **10**, 289–292.

Chattopadhyaya, R., Meador, W.E., Means, A.R. and Quiocho, F.A. (1992) *J. Mol. Biol.*, **228**, 1177–1192.

Clore, G.M., Gronenborn, A.M. and Bax, A. (1998a) *J. Magn. Reson.*, **133**, 216–221.

Clore, G.M., Starich, M.R. and Gronenborn, A.M. (1998b) *J. Am. Chem. Soc.*, **120**, 10571–10572.

Clore, G.M., Gronenborn, A.M. and Tjandra, N. (1998c) *J. Magn. Reson.*, **131**, 159–162.

Cook, W.J., Jeffrey, L.C., Cox, J.A. and Vijay–Kumar, S. (1993) *J. Mol. Biol.*, **229**, 461–471.

Cordier, F., Dingley, A.J. and Grzesiek S. (1999) *J. Biomol. NMR*, **13**, 175–180.

Czisch, M. and Boelens, R. (1998) *J. Magn. Reson.*, **134**, 158–160.

Flaherty, K.M., Zozulya, S., Stryer, L. and McKay, D.B. (1993) Cell, **75**, 709–716.

Griffith, J.P., Kim, J.L., Kim, E.E., Sintchak, M.D., Thomson, J.A., Fitzgibbon, M.J., Fleming, M.A., Caron, P.R., Hsiao, K. and Navia, M.A. (1995) Cell, **82**, 507–522.

Hansen, M.R., Mueller, L. and Pardi, A. (1998) *Nat. Struct. Biol.*, **12**, 1065–1074.

Ikura, M., Clore, G.M., Gronenborn, A.M., Zhu, G., Klee, C.B. and Bax, A. (1992) *Science*, **256**, 632–638.

Losonczi, J.A. and Prestegard, J.H. (1998) *J. Biomol. NMR*, **12**, 447–451.

Lounila, J. and Jokisaari, J. (1982) *Progr. NMR Spectrosc.*, **15**, 249–290.

Muhandiram, D.R. and Kay, L.E. (1994) *J. Magn. Reson.*, **B103**, 203–216.

Ottiger, M., Delaglio, F. and Bax, A. (1998) *J. Magn. Reson.*, **131**, 373–378.

Ottiger, M. and Bax, A. (1998) *J. Biomol. NMR*, **12**, 361–372.

Pervushin, K., Riek, R., Wider, G. and Wüthrich, K. (1997) *Proc. Natl. Acad. Sci. USA*, **94**, 12366–12371.

Pervushin, K., Wider, G. and Wüthrich, K. (1998) *J. Biomol. NMR*, **12**, 345–348.

Piotto, M., Saudek, V. and Sklenár, V. (1992) *J. Biomol. NMR*, **2**, 661–665.

Ramirez, B. and Bax, A. (1998) *International Conference of NMR on Biological Systems*, Tokyo.

Sanchez, R. and Sali, A. (1997) *Curr. Opin. Struct. Biol.*, **7**, 206–214.

Sanchez, R. and Sali, A. (1998) *Proc. Natl. Acad. Sci. USA*, **95**, 13597–13602.

Sanders II, C.R. and Prestegard, J.H. (1990) *Biophys. J.*, **58**, 447–460.

Sanders II, C.R. and Schwonek, J.P. (1992) *Biochemistry*, **31**, 8898–8905.

Swan, D.G., Hale, R.S., Dhillon, N. and Leadlay, P.F. (1987) *Nature*, **329**, 84–85.

Tjandra, N. and Bax, A. (1997) *Science*, **278**, 1111–1114.

Tjandra, N., Omichinski, J.G., Gronenborn, A.M., Clore, G.M. and Bax, A. (1997) *Nat. Struct. Biol.*, **4**, 732–738.

Vijay–Kumar, S. and Cook, W.J. (1992) *J. Mol. Biol.*, **224**, 413–426.

Wang, H., Eberstadt, M., Olejniczak, E.T., Meadows, R.P. and Fesik, S.W. (1998) *J. Biomol. NMR*, **12**, 443–446.

Weigelt, J. (1998) *J. Am. Chem. Soc.*, **120**, 10778–10779.

Wishart, D. and Sykes, B.D. (1994) *Methods Enzymol.*, **239**, 363–392.

Yamazaki, D., Lee, W., Arrowsmith, C.H., Muhandiram, D.R. and Kay, L.E. (1994) *J. Am. Chem. Soc.*, **16**, 11655–11666.

Yang, D., Tolman, J.R., Goto, N.K. and Kay, L.E. (1998) *J. Biomol. NMR*, **12**, 325–332.