Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Magnetic Resonance 172 (2005) 85–90
JMR Journal of Magnetic Resonance
www.elsevier.com/locate/jmr

Assessment of protein alignment using ^{1}H–^{1}H residual dipolar coupling measurements

Junfeng Wang, Homayoun Valafar, and James H. Prestegard*

Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA

Received 20 October 2003; revised 10 March 2004
Available online 13 November 2004

Abstract

A quick and accurate method is described for assessing protein alignment from residual dipolar coupling (RDC) measurements. In contrast to observing D_2O resonance splitting, which reflects the orientational order of the alignment medium, the degree of alignment of a protein of interest can be estimated directly from ^{1}H–^{1}H RDCs. In this study, RDCs between aromatic protons in unlabeled Cp-rubredoxin were measured from proton homonuclear *J*-resolved experiments with high sensitivity, and the alignment was assessed without the need of extensive resonance assignment. Since labeled proteins are not needed, this method provides an efficient way for screening alignment media. In situations where the protein structure is known, as in the case of Cp-rubredoxin, a full set of order tensor parameters can be determined, allowing further studies, such as those of ligand alignment relative to a target protein.

Keywords: NMR; *J*-resolved spectra; Order tensors; Alignment media; Rubredoxin

1. Introduction

Residual dipolar couplings (RDCs) observed in high-resolution NMR spectra of partially aligned samples have become standard sources of information for probing structure and dynamics in biological macromolecules [1,2]. Their utility resides in the angular dependence of the through-space dipolar interaction between pairs of spin 1/2 nuclei. These interactions are observable as additional splittings of NMR resonances (D), or additional contributions to splittings that already exist due to scalar couplings between directly bonded nuclei ($J + D$), when a sufficient level of alignment is induced. In the latter case, molecules of interest are frequently isotopically labeled with ^{15}N or ^{13}C so that one-bond couplings between the heteronucleus and a directly attached proton is observed.

However, there are instances when introduction of isotopic labels is not convenient, and it becomes desirable to assess interactions using naturally occurring spin 1/2 nuclei, such as pairs of proximate protons. One such instance occurs when the orientation of ligands relative to a protein of known structure is of primary interest. The orientation of the protein, in addition to the ligand, must be determined in this case, but the effort required to specifically label the protein, and assign the resonances, may be unwarranted. Another instance arises when the protein is of interest, but the sacrifice of precious isotopically labeled protein in preliminary experiments to select an appropriate alignment medium may be unwarranted. There are now a dozen or more alignment media in use, and not all are compatible with a given protein [3–5]. Also, an optimum alignment level must be found that gives dipolar couplings that are neither too large nor too small.

It is tempting to assess alignment from splittings of the deuterium resonance from D_2O added to the sample [6–8]. Here a quadrupole splitting displays the same type of angular dependence as the RDCs. But, D_2O alignment really probes properties of the medium and may not relate simply to the alignment of the biomolecule of interest. We present here a procedure based on measurement of RDCs from proton pairs in the molecule of interest that may be of particular value in these instances.

* Corresponding author. Fax: 1-706-542-4412.
E-mail address: jpresteg@ccrc.uga.edu (J.H. Prestegard).

1090-7807/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jmr.2004.03.012

Protons seem ideal probes of alignment in molecules of interest. They are the nearly 100% naturally abundant isotope of hydrogen, they provide maximum NMR sensitivity, and they exhibit strong dipole–dipole interactions. However, the high proton density in most molecules presents a problem. ^{1}H–^{1}H dipolar couplings between all pairs of protons within 6–7 Å, are often observable leading to broad peaks and multiple splittings. Therefore, it is desirable to suppress most couplings while retaining a sufficient number of representative dipolar couplings to characterize the alignment. Background signals from alignment media might also be considered problematic. However, compared to resonances from the macromolecules dissolved in the alignment medium, resonances from the strongly aligned medium are split by many more strong couplings, rendering them largely unobservable. In this study, we choose to observe couplings between aromatic protons from a protein of interest. Few alignment media used for proteins have aromatic resonances minimizing background, and proteins have limited sets of well resolved aromatic resonances making retention and analysis of limited sets of couplings possible.

As a target for illustration we choose the Zn form of *Clostridium pasteurianum* (Cp) rubredoxin [9], a well characterized, highly soluble, small protein. Several crystal structures exist for this protein (pdb codes: 1IRN, 1IRO, 5XRN, 4XRN, 1FHH, and 1FHM) as well as close homologs from other organisms [10–12]. Its amino acid sequence includes 1 tryptophan, 2 phenylalanines, and 3 tyrosines, which comprise 1/9 of the overall sequence. The Cp-rubredoxin used in our experiments was expressed as previously documented [13]. A sample was prepared to be 1 mM rubredoxin in 50 mM potassium phosphate at pH 6.3. It was subsequently lyophilized and re-dissolved in a 7% 3:1 DMPC/DHPC (dimyristoylphosphatidylcholine/dihexanoylphosphatidylcholine) bicelle solution that was prepared in D_2O to reduce signals from amide protons.

Fig. 1 depicts the pulse sequence designed for this study. It is a simple modification of the original proton homonuclear *J*-resolved experiment [14]. Since only scalar plus dipolar coupling evolves in the t1 dimension, a narrow spectral width can be used, along with a small number of t1 points, to give adequate resolution with a modest investment in time. In the pulse sequence, a band-selective single pulsed field gradient spin echo (SPFGSE) is introduced in the middle of the t1 evolution period. The selective 180 pulse is set to affect only the proton resonances in the range from 5 to 9 ppm. Only the resonances carrying couplings between protons in the selected band, i.e., aromatic protons, now evolve with couplings and refocus at the end of each t1 increment; amide protons, when they are present, give rise primarily to a single, decoupled peak at zero frequency in the t1 dimension. Since the water resonance is outside

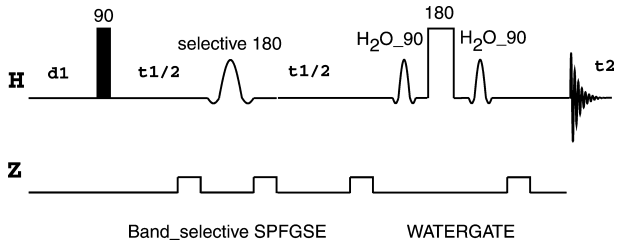

Fig. 1. Pulse scheme for obtaining couplings between aromatic protons. Narrow and wide bars correspond to 90° and 180° pulses, respectively, with phase x, unless indicated. The selection of proton–proton couplings between aromatic protons during t1 is achieved by applying a REBUP pulse to the aromatic region (5–9 ppm). The WATERGATE 3919 scheme is applied for water suppression with d3 duration of 100 μs. Phase cycling: $\phi_1 = x,-x,y,-y$; $\phi_2 = x,x,4(y),4(-x),4(-y),x,x$; $\phi_4 = x,-x,y,-y,-x,x,-y,y$; and $\phi_5 = -x,x,-y,y$. All gradients are 1 ms long with strengths $G1 = 20.4$ G/cm and $G2 = 26.6$ G/cm.

the selective envelop, the SPFGSE contributes to water suppression. However, to further suppress the water resonance, a WATERGATE segment is added after t1 evolution.

Spectra were acquired on a Varian Inova 500 MHz spectrometer. Data collection typically included 50 t1 data points and 2048 t2 data points over a period of 20 min. All data processing was done with the nmrPipe package [15]. Experiments were run at 25 °C where the bicelle medium is isotropic and at 35 °C where the medium is partially oriented. Typical splittings of the deuterium water resonance under aligned conditions were 15 Hz.

Figs. 2A and B present data on the rubredoxin sample under isotropic and oriented conditions, respectively. Sixteen well resolved vertically displaced multiplets corresponding to the 16 resonances expected for the 6 residues in CP-rubredoxin can be seen in the isotropic spectrum. The splittings in the isotropic spectrum correspond to those expected for three bond scalar couplings (7–10 Hz); the multiplets are triplets or doublets depending on whether one coupling partner exists as for tyrosine δ protons or two coupling partners exist as for phenylalanine ζ protons.

The change in splitting under aligned conditions is due to the addition of dipolar couplings to the isotropic scalar couplings, and the difference between splittings in isotropic and aligned spectra provides a measurement of the through-space dipolar coupling. With a signal to noise ratio of 20:1, an error of 1/3 of the line width provides a reasonable estimation of experimental error. On this basis an average error of about 1.6 Hz is predicted for the majority of the individual couplings measured. The experimental errors in the RDCs are about twice this large. The dipolar coupling can also be positive or negative depending on the orientation of inter-proton vector orientation. Under most weakly aligned conditions, however, the magnitudes of these

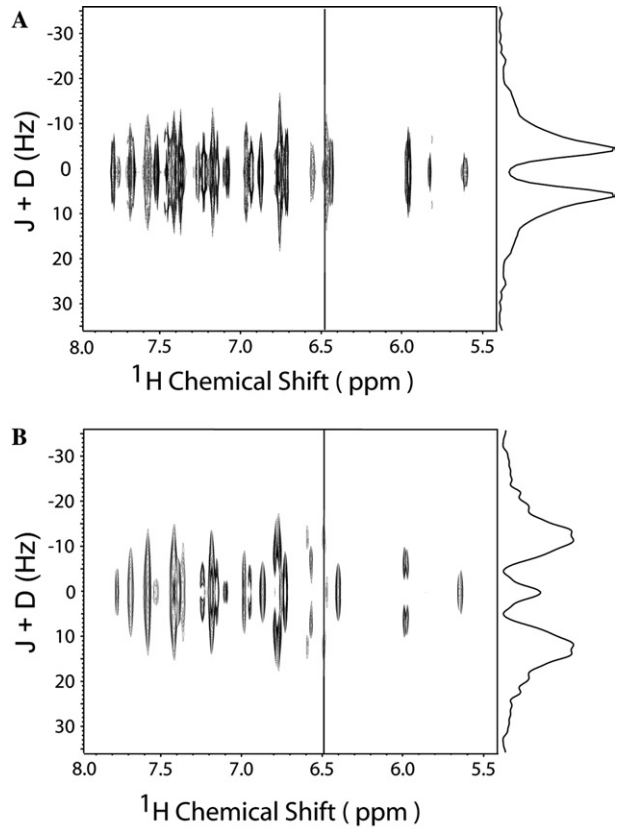

Fig. 2. Spectra of Cp-rubredoxin recorded on a Varian Inova 500 spectrometer (500 MHz proton frequency) with the pulse scheme of Fig. 1. (A) spectrum acquired at 25 °C (isotropic); (B) spectrum acquired at 35 °C (aligned). Spectral widths in the F1-(F2-) dimension = 50 (8000) Hz, number of t1 increments = 64, number of transients = 16, acquisition time (t2) = 256 ms. Data were linear predicted to 128 in the F1-dimension, and a squared cosine window function was applied in the F1- and F2-dimensions. Processing was performed using NMRPipe software package [15]. The cross-section is for H_δ of tyrosine 4 at 6.48 ppm, which is coupled to the H_ε proton of the same ring.

dipolar couplings are smaller than the corresponding J values leading to positive $J + D$ values and easily determined signs for the dipolar contribution. If the molecules align too strongly, however, a sign ambiguity in $J + D$ that results in two possible D values has to be considered. In a few cases, additional smaller splittings are observed; these most likely arise from cross-ring meta protons that have no significant J coupling. Signs of measured splittings here are truly ambiguous.

In Fig. 2 F1 slices are highlighted for Hδ of tyrosine 4, which is coupled to its neighboring Hε proton. The change in splitting from 9 Hz in the isotropic case to ±24 Hz under the aligned conditions, leads to a difference of 15 or −33 Hz that can be attributed to residual dipolar coupling. While the size of the coupling could, in principle, have lead to an observed $J + D$ that is either positive or negative, a −33 Hz dipolar coupling would have had to result from a level of alignment much higher than those typically observed under the conditions used here. Hence a value of +15 Hz is chosen. The signs of most ambiguous RDCs from Cp-rubredoxin can be resolved using similar arguments; all resulting RDCs are summarized in Table 1.

Even without specific assignments, and knowledge of the structure of the protein, we can extract useful information from the data in Table 1. Theoretically, if the internuclear vectors are distributed uniformly and isotropically in space, the axial and rhombic components of the order tensor can be extracted from the distribution of normalized residual dipolar couplings in the absence of any structural information [16]. In our case, the small set of inter-proton vectors in aromatic rings samples the space in a manner that is far from uniform. Nevertheless, it is possible to set limits on at least the axial component of the order tensor. From the data in table one, the maximum coupling of 15 Hz would suggest an order tensor with a minimum principal order parameter, S_{zz}, of -9×10^{-4} (assuming 2.5 Å distance for all neighboring aromatic protons). Knowing an order tensor is extremely valuable in choosing conditions for further structural studies on proteins. If alignment is too low, the precision of RDC measurements is degraded; if alignment is too high additional long-range couplings degrade signal quality and complicate extraction of individual RDCs. The above procedures provide ready access to information that allows optimization of alignment media with small amounts of unlabeled protein. This can be particularly important in structural genomics applications where large numbers of proteins are to be studied [17].

Much more information can be obtained with assignment of individual aromatic resonances and knowledge of the protein structure. In the case of rubredoxin, assignments of aromatic proton resonances exist for either the Cp protein on which we are working or for the closely homologous Pf protein [18]. Assigning couplings to specific sites as indicated in Table 1, and using coordinates derived from the Cp-rubredoxin crystal structure [PDB code: 1IRN (1.2 Å)] (protons were added using a tool available in the Reduce program [19]), a set of equations relating couplings and inter-proton vector geometry to order parameters can be solved. In situations where a structure is available, but assignments of resonances are not, it may still be possible to proceed. Assigning multiplets to an amino acid

Table 1
Aromatic proton RDCs (Hz) of Cp-rubredoxin extracted from *J*-resolved experiments

Y4	H_δ–H_ε	15				
Y11	H_δ–H_ε	5				
Y13	H_δ–H_ε	−2				
F30	H_δ–H_ε	−2	H_ε–H_ζ	2		
F49	H_δ–H_ε	−2	H_ε–H_ζ	1		
W37	$H_{\varepsilon 3}$–$H_{\zeta 3}$	5	$H_{\zeta 2}$–$H_{\eta 2}$	−1	$H_{\zeta 3}$–$H_{\eta 2}$	3

type (tyrosine, vs phenylalanine, vs tryptophan, vs histidine) can easily be done with additional connectivity information from a TOCSY experiment (data not shown), and within each amino acid type, assignment to a specific site can be based on consistency with RDCs [20]. For Cp-rubredoxin, there are only three tyrosines, two phenylalanines, and a single tryptophan. Permuting coupling values among the tyrosine set and the phenylalanine set presents only 12 possible sets. These can easily be checked for consistency with dipolar data and known structures.

Looking for consistency of aromatic proton RDCs with structural predictions requires certain assumptions about averaging of couplings due to ring flips about β–γ bonds. For Tyr and Phe, the existence of this type of averaging has been amply documented. For example, a recent search of the BioMagResBank suggested that the overwhelming majority of the Phe and Tyr rings (>95%) are flipping rapidly on the chemical shift time scale at ambient temperature [21]. Hence, we will assume rapid ring flipping and motional averaging for these residues. Tryptophan presents a more complex case. However, it seems likely that ring flips will occur less readily, at least for buried tryptophans, because substantial packing rearrangements would be required to accommodate an alternate tryptophan side-chain geometry. We will, therefore, assume tryptophan side-chains to be rigid. In principle, ring flips could average all H–H RDCs, complicating a structural analysis. However, when a ring flips about the C_β–C_γ bond, orientations of H_δ–H_ε vectors of Phe and Tyr do not change, allowing interpretation as in a completely rigid case. RDCs for the H_ε–H_ζ vectors of phenylalanine will be averaged by the motion. In this case, provision must be made for simulation of averaged RDCs.

A program named residual dipolar coupling analysis tool (REDCAT) is available for the interpretation of RDCs in terms of molecular structure and alignment of that structure [23]. A new version, which implements simulation of motional effects has been designed and is used for the first time here. Having assigned measured RDCs to specific sites in a protein of known structure, the program can return both principal order parameters and the orientation of the order frame from the point of view of the protein. The program operates by initially solving a set of equations that relate RDCs to internal vector geometry and a set of five order parameters. The order parameters are extracted and assembled in a symmetric and traceless order matrix; diagonalization of the order matrix yields principal order parameters (S_{xx}, S_{yy}, and S_{zz}) and a transformation matrix that relates the principal order frame to the molecular frame. S_{zz} is the largest in magnitude of the principal order parameters; it defines the level of axial order. An asymmetry parameter, η is defined in terms of the other principal order parameters $((S_{yy} - S_{xx})/S_{zz}$, where $|S_{zz}| > |S_{yy}| > |S_{xx}|)$; it describes the anisotropy of order. These parameters are simply related to alignment parameters Da and R used in some analysis programs as

$$\mathrm{Da} = \frac{1}{2} D_{\max} S_{zz} = \frac{1}{2}\left(-\frac{\mu_0}{4\pi}\right)\left(\frac{\gamma_i \gamma_j h}{2\pi^2 r_{ij}^3}\right) S_{zz} \text{ and } R = \frac{2}{3}\eta,$$

where γ_i and γ_j are the gyromagnetic ratios for the spin 1/2 nuclei i and j, h is Plank's constant, and r_{ij} is the internuclear distance.

The solutions from the application of REDCAT to the RDC data on rubredoxin are shown in Fig. 3A as a Sauson–Flamsteed plot that depicts the orientation of the principal order frame as positions of x, y, and z axes on the surface of an exploded globe. Multiple solutions produced through a Monte-Carlo sampling of allowed RDC values (using error estimates of 6 Hz, which account for both experimental errors and errors in structural representation) are superimposed on the plot. The lack of a precise definition of axis directions is not sur-

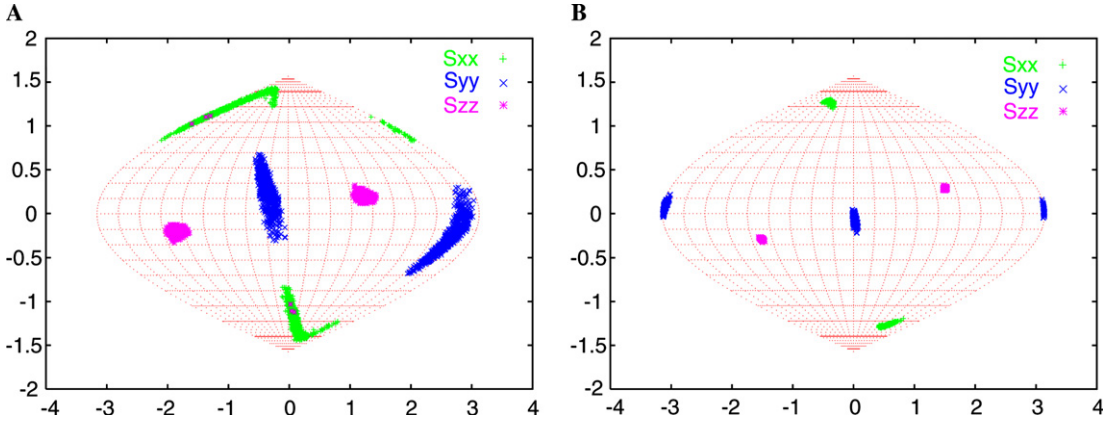

Fig. 3. Orientations of molecular alignment tensors of (A) Cp-rubredoxin by REDCAT calculation with aromatic proton RDCs and (B) Pf-rubredoxin by REDCAT calculation with NH RDCs. Orientations are visualized using Sauson–Flaumsteed projections, which maps the surface of a unit sphere into a plane by converting latitude (φ) and longitude (λ) to Cartesian coordinates (x, y) via $y = \varphi$ and $x = \lambda\cos(\varphi)$.

prising given the small number of data used in this case. The definition is, however, clearly adequate for orienting a protein domain, as would be needed in ligand binding or protein association studies.

An assessment of the accuracy of the order direction can be made by comparing the above results to results from analysis of a set of ^{1}H–^{15}N RDCs. We do not have a complete set for Cp-rubredoxin, but we do have a set for rubredoxin from *Pyrococcus furiosus* (Pf). Pf-rubredoxin is the hyperthermophilic counterpart to Cp-rubredoxin; it shares approximately 57% primary sequence identity and 83% primary sequence homology with Cp-rubredoxin [12]. Superimposition of the two highest resolution structures, 1BRF (Pf-rubredoxin, 1.6 Å) and 1IRN (Cp-rubredoxin, 1.2 Å) shows a 0.47 Å RMS deviation in C_α atomic positions. When dissolved in a bicelle medium, where the primary interaction between protein and bicelle particles is steric, Cp- and Pf-rubredoxins are expected to have similar medium induced alignment. The order tensor for Pf-rubredoxin in a bicelle medium was, therefore, calculated based on NH RDCs collected in a bicelle solution at the same concentration used here [22]. Fig. 3B shows the alignment of the order tensor of Pf-rubredoxin from this set of H–N RDCs. An averaged error of 7 Hz was assumed. Comparing two maps, orientation distributions of the solutions determined from aromatic proton RDCs nearly overlap those from NH RDCs. Ranges of S_{zz} and η also overlap for the two determinations.

For proteins of larger size, the availability of more aromatic residues will provide more RDC restraints and the inter-proton vectors will have a better chance to sample orientation space. This will allow a better assessment of overall order and asymmetry. Where structures and resonance assignments are available, accuracy and definition of protein alignment could also improve. The current method for measurement of ^{1}H–^{1}H RDCs was tested on the somewhat larger (14 kDa) protein, lysozyme, and accurate measurement of RDCs could still be made. This means that levels of alignment can be estimated for proteins of this size and possibly larger. Making assignments needed for determination of a complete order tensor will be more problematic. However, experiments with more chemical shift resolution, including ones based on aromatic selective COSY sequences, will help. It is also possible to include RDCs from other amino acid types by selecting subsets of resonances with newly designed NMR experiments. We therefore remain very optimistic about future applications.

In this paper, we have illustrated an ability to obtain a good estimation of the alignment of a protein with a minimal investment in time (10–20 min) for a small protein at 0.5 mM concentration. Since labeled proteins are not needed, this method provides an efficient way for screening alignment media. In cases where the protein structure is known, the method can be used to deduce the direction of alignment axes. This should prove to be of particular value where ligand alignment relative to that of a protein is of primary interest.

Acknowledgments

This work was supported by grants from the National Institutes of Health, GM033225, and from the National Science Foundation, MCB0092661. We thank Dr. Catherine Bougault and Dr. Marly Edisness for assistance with sample preparation.

References

[1] J.H. Prestegard, New techniques in structural NMR—anisotropic interactions, Nat. Struct. Biol. 5 (1998) 517–522.
[2] A. Bax, Weak alignment offers new NMR opportunities to study protein structure and dynamics, Protein Sci. 12 (1) (2003) 1–16.
[3] J. Tolman, A novel approach to the retrieval of structural and dynamic information from residual dipolar couplings using several oriented media in biomolecular NMR spectroscopy, J. Am. Chem. Soc. 124 (40) (2002) 12020–12030.
[4] W. Peti, J. Meiler, R. Bruschweiler, C. Griesinger, Model-free analysis of protein backbone motion from residual dipolar couplings, J. Am. Chem. Soc. 124 (20) (2002) 5822–5833.
[5] E.R. Zartler, F.E. Jenney, M. Terrell, M.K. Eidsness, M.W.W. Adams, J.H. Prestegard, Structural basis for thermostability in aporubredoxins from *Pyrococcus furiosus* and *Clostridium pasteurianum*, Biochemistry 40 (24) (2001) 7279–7290.
[6] D.M. Chen, F.Y. Fujiwara, L.W. Reeves, Studies of behavior in in magnetic-fields of some lyomesophase systems with respect to electrolyte addictions, Can. J. Chem. 55 (12) (1977) 2096–2403.
[7] H.H. Mantsch, H. Saito, I.C.P. Smith, Deuterium magnetic-resonance: application in chemistry, physics and biology, Prog. Nucl. Magn. Reson. Spectrosc. 11 (1977) 211–271.
[8] B.J. Forrest, L.W. Reeve, New lyotropic liquid-crystals composed of finite nonspherical micelles, Chem. Rev. 81 (1) (1981) 1–14.
[9] Z. Dauter, K.S. Wilson, L.C. Sieker, J.M. Moulis, J. Meyer, Zinc- and iron-rubredoxins from *Clostridium pasteurianum* at atomic resolution: a high-precision model of a ZnS4 coordination unit in a protein, Proc. Natl. Acad. Sci. USA 93 (17) (1996) 8836–8840.
[10] Z. Dauter, L.C. Sieker, K.S. Wilson, Refinement of rubredoxin from *Desulfovibrio-vulgaris* at 1.0-Å with and without restraints, Acta Crystallogr. B 48 (1992) 42–59.
[11] M. Frey, L. Sieker, F. Payan, R. Haser, M. Bruschi, G. Pepe, J. Legall, Rubredoxin from *Desulfovibrio-gigas*—a molecular-model of the oxidized form at 1.4 Å resolution, J. Mol. Biol. 197 (3) (1987) 525–541.
[12] K.D. Watenpaugh, L.C. Sieker, L.H. Jensen, Crystallographic refinement of rubredoxin at 1.2 a resolution, J. Mol. Biol. 138 (3) (1980) 615–633.
[13] M.K. Eidsness, K.A. Richie, A.E. Burden, D.M. Kurtz, R.A. Scott, Dissecting contributions to the thermostability of *Pyrococcus furiosus* rubredoxin: beta-sheet chimeras, Biochemistry 36 (1997) 10406–10413.
[14] W.P. Aue, J. Karhan, R.R. Ernst, Homonuclear broad band decoupling in two-dimensional J-resolved NMR spectroscopy, J. Chem. Phys. 64 (10) (1976) 4226–4227.
[15] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, Nmrpipe—a multidimensional spectral processing system based on unix pipes, J. Biomol. NMR 6 (3) (1995) 277–293.

[16] G.M. Clore, A.M. Gronenborn, A. Bax, 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. 133 (1) (1998) 216–221.
[17] M.W.W. Adams, H.A. Dailey, L.J. Delucas, M. Luo, J.H. Prestegard, J.P. Rose, B.C. Wang, The southeast collaboratory for structural genomics: a high-throughput gene to structure factory, Accounts Chem. Res. 36 (3) (2003) 191–198.
[18] C.M. Bougault, M.K. Eidsness, J.H. Prestegard, Hydrogen bonds in rubredoxins from mesophilic and hyperthermophilic organisms, Biochemistry 42 (15) (2003) 4357–4372.
[19] J.M. Word, S.C. Lovell, J.S. Richardson, D.C. Richardson, Asparagine and glutamine: using hydrogen atom contacts in the choice of side-chain amide orientation, J. Mol. Biol. 285 (4) (1999) 1735–1747.
[20] M. Zweckstetter, Determination of molecular alignment tensors without backbone resonance assignment: aid to rapid analysis of protein–protein interactions, J. Biomol. NMR 27 (1) (2003) 41–56.
[21] J.J. Skalicky, J.L. Mills, S. Sharma, T. Szyperski, Aromatic ring-flipping in supercooled water: implications for NMR-based structural biology of proteins, J. Am. Chem. Soc. 123 (3) (2001) 388–397.
[22] F. Tian, H. Valafar, J.H. Prestegard, A dipolar coupling based strategy for simultaneous resonance assignment and structure determination of protein backbones, J. Am. Chem. Soc. 123 (47) (2001) 11791–11796.
[23] H. Valafar, J.H. Prestegard, REDCAT: a residual dipolar coupling analysis tool, J. Magn. Res. 167 (2004) 228–241.