Domain Orientation and Dynamics in Multidomain Proteins from Residual Dipolar Couplings†

Mark W. F. Fischer, Judit A. Losonczi, Jeanne Lim Weaver, and James H. Prestegard*

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

Received March 5, 1999; Revised Manuscript Received May 12, 1999

ABSTRACT: The data most commonly available for the determination of macromolecular structures in solution are NOE based distance estimates and spin−spin coupling constant based dihedral angle estimates. This information is, unfortunately, inherently short-range in nature. Thus, for many multidomain proteins, little information is available to accurately position weakly interacting domains with respect to each other. Recent studies of proteins aligned in dilute liquid crystalline solvents have shown the utility of measuring anisotropic spin interactions, such as residual dipolar couplings, to obtain unique long-range structural information. In this work, the latter approach is taken to explore the relative domain orientation in a two-domain fragment from the protein barley lectin. An approach based on singular value decomposition as opposed to simulated annealing is used to directly determine order tensors for each domain from residual $^{15}N-^{1}H$ dipolar couplings, and the limitations of the two approaches are discussed. Comparison of the order tensor principal axis frames as separately determined for each domain indicates that the two domains are not oriented as in the crystal structure of wheat germ agglutinin, a highly homologous protein (∼95% sequence identical). Furthermore, differences in the order tensor values suggest that the two domains are not statically positioned but are experiencing different reorientational dynamics and, to a large degree, may be considered to reorient independently. Data are also presented that suggest that a specific association occurs between one domain and the lipid bicelles comprising the liquid crystal solvent.

It is increasingly apparent that proteins seldom act in isolation when carrying out cellular functions: multienzyme systems cooperate in the synthesis of complex molecules, enhancers and inhibitors control the activity of enzymes, proteins assemble to make protective coats or structural entities, and proteins interact in signal-transduction systems. The protein−protein or domain−domain interactions in these systems raise structural questions about proteins to a new level. We must know not only how individual domains of proteins are internally structured, but also how they are oriented with respect to one another in functional assemblies. These assemblies can easily be fragile and transient in character and highly dependent on their immediate environment. While X-ray crystallography has made important inroads here and has provided most of information of structure at the atomic level of individual protein domains, NMR,[1] being applicable to aqueous solutions, may well have a role to play in the structural definition of protein assemblies.

Traditionally, NMR approaches to biomolecular structure have been based on distance constraints derived from NOEs (*1*). For the questions at hand, this approach is limited because of the paucity of observable constraints at protein−protein or domain−domain interfaces. This arises as a result of the short-range ($1/r^6$) interaction on which NOE-derived distance constraints are based and because of the difficulty in assigning discrete resonances for the side chains and termini most frequently involved in domain−domain contacts. Herein, we present an alternate approach based on orientational constraints derived from residual dipolar couplings that are seen in protein spectra taken in partially oriented media (*2*−*5*). We apply the approach to a question of domain−domain orientation in a two-domain fragment of the barley lectin protein.

Dipolar couplings between nuclear spins are usually averaged from maximum values in the tens of kilohertz range to zero by the isotropic tumbling of molecules in solution. If anisotropy can be reintroduced, the dipolar interaction would no longer vanish and residual dipolar couplings could be observed. When the couplings are between directly bonded nuclei, such as the ^{15}N and ^{1}H of an amide bond, the bond length can be assumed to be fixed and the primary dependence is on the average of an angular dependent function, $(3\cos^2\theta - 1)/2$, where θ is the angle between the magnetic field and the bond vector (*5*). When data on several vectors in a domain of defined structure can be collected, they can be combined to define a principal orientation frame

† This research was supported by a grant from the National Institutes of Health, GM33225.

* To whom correspondence should be addressed at the Complex Carbohydrate Research Center, University of Georgia, 220 Riverbend Road, Athens, GA 30602-4712. Telephone: (706) 542-6281. FAX: (706) 542−4412. Email: jpresteg@ccrc.uga.edu.

[1] Abbreviations: NOE, nuclear Overhouser enhancement; NMR, nuclear magnetic resonance spectroscopy; HSQC, heteronuclear single quantum correlation spectroscopy; SVD, singular value decomposition; BL, barley lection; WGA, wheat germ agglutinin; BLBC, a B and C domain fragment of BL; DMPC, dimyristoylphosphatidylcholine; DHPC, dihexanoylphosphatidylcholine; CTAB, hexadecyl(cetyl)trimethylammonium bromide; D, dipolar coupling constant; J, scalar coupling constant; PAS, principal axis system; CNS, Crystallography & NMR System; RMSD, root mean square deviation. Three-letter amino acid abbreviations used throughout.

10.1021/bi9905213 CCC: $18.00
Published on Web 06/24/1999

and degree of order for the domain. If a common interaction with the medium orients each of several domains, the resulting orientation frames and degrees of order for properly positioned rigid domains must coincide (*6*). Thus, residual dipolar couplings can provide long-range information about relative domain orientation in multidomain proteins. Carefully analyzed via an order matrix approach, these observations can also be used to investigate independent domain motion and the nature of domain-orienting medium interactions. There are now numerous examples of dipolar data collected on proteins oriented either by inherent magnetic susceptibility anisotropy (*3*, *4*, *7*) or by dissolution in liquid crystal media that include phospholipid bicelles (*8*) and bacteria phage (*9*−*11*). Herein, we use a bicelle medium to study domain orientation in a two-domain fragment of a carbohydrate binding protein, barley lectin.

Carbohydrate binding lectins, which are found in all organisms, have been shown to play key roles in such important physiological processes as fertilization, cell adhesion, and host defense (*12*, *13*). The individual domains of lectins actually have rather low affinities for the carbohydrate ligands that they recognize. But, when ligands are arrayed on cell surfaces, or are presented as part of multiantenniary conjugates on glycoproteins, they can be recognized by multidomain aggregates with high affinity and specificity (*14*). Hence, lectins usually function as multidomain aggregates and relative domain orientation is a central issue in the lectin field.

Barley lectin (BL), the subject of this study, shares approximately 95% sequence identity with another member of the highly conserved class of proteins known as the cereal lectins, namely, the well-characterized wheat germ agglutinin (WGA). WGA has been well-characterized by X-ray crystallography (*15*). Both WGA and BL consist of four homologous domains labeled A−D. Each well-defined 43 amino acid domain contains one carbohydrate binding site and is composed of two antiparallel β-strands plus a short α-helix. Each domain is stabilized by a network of four disulfide bonds. WGA and BL exist as dimers with intersubunit contacts between domains A_1-D_2, B_1-C_2, C_1-B_2, and D_1-A_2 where subscripts indicate the subunit. This dimerization results in macromolecules with as many as eight carbohydrate binding sites. However, crystal structures show only some of these sites to be occupied and also show occupancy at an interface between pairs of domains in a way that might preclude the simultaneous use of all sites (*16*). Whether this level of occupancy and mode of binding persists in solution remained in question. In an effort to explore lectin structure and carbohydrate binding in solution, a two-domain fragment of the barley lectin protein was engineered and studied by conventional NMR methods (*17*). The fragment, consisting of domains B and C, was chosen both to reduce the molecular size to one amenable to NMR study and to reduce the heterogeneity of binding sites present. This construct was named barley lectin BC or BLBC. If the construct dimerized as in the crystal structure of WGA, two identical BC interface sites would be formed. A structure for the B and C domains was determined by NMR methods (*17*).

As is often the case for multidomain proteins, however, very few interdomain NOEs were observed. Because of this scarcity of interdomain NOE contacts, it was difficult to position the two domains with respect to each other based on NOE contacts or other short-range structural reporters alone. Moreover, no multimeric species were observed from electrospray mass spectroscopy (*17*), and BLBC was found to contain two independent binding sites. This departure from dimer properties expected from the crystal structure and the absence of specific information on interdomain contacts in solution samples raised an additional question regarding whether the relative domain orientations of the B and C domains in the monomer were retained as in the WGA crystal structure or whether there was more widespread loss of crystallike properties. Answering this question would not only provide useful information about the properties of a member of the cereal lectin class of carbohydrate binding proteins but also provide a test for methodology based on orientational constraints that might be applicable to a broader range of questions involving the organization of domains in multidomain protein assemblies.

The orientational methods to be employed herein for the determination of domain orientation are, of course, capable of giving structural information about arrangements of bonds and secondary structural elements within individual domains as well. Approaches that incorporate orientational data in simulated annealing protocols for structure determination have now been applied in the refinement of several protein structures (*18*). While an NOE-based structure of the individual domains in the BLBC fragment was determined in the course of the previous work on ligand binding, these structures were not of high precision. Thus, we also explore the possibility of simultaneously determining domain orientation and refining structures of individual domains using a simulated annealing procedure. Several cautionary notes about the use of such procedures emerge.

MATERIALS AND METHODS

Bicelle Preparation. Two dilute bicelle solutions were prepared using essentially identical procedures. The method was given in detail in a previous work (*19*). To summarize briefly here, samples of dimyristoylphosphatidylcholine (DMPC Sigma, St. Louis, MO) and dihexanoylphosphatidylcholine (DHPC Avanti Polar Lipids, Birmingham, AL) were dissolved in small volumes of aqueous buffer. Vortexing and several freeze/thaw cycles were required to homogenize the DMPC dispersion, which was then heated to above the T_m of pure DMPC (23 °C). The DHPC solution was added, and the resulting mixture was vortexed and then quickly frozen in a methanol/solid CO_2 bath. The solution was then allowed to slowly come to room temperature. The final bicelle solutions contained 5% (w/v) lipid in 25 mM phosphate buffer at pH 6. A small volume of concentrated hexadecyl(cetyl)trimethylammonium bromide (CTAB, ACROS, Fair Lawn NJ), a positively charged lipid dopant chosen to discourage electrostatic association of bicelles with BLBC at pHs below its isoelectric point, was added to the prepared bicelle solutions.

The two preparations differ in two ways. First, the exact molar ratios of DMPC/DHPC are 2.9:1 and 2.8:1 in preparation 1 and preparation 2, respectively. These ratios were intended to be identical, but they differed due to difficulties in weighing DMPC and DHPC due to the hygroscopic nature of DHPC. The ratios given were determined by integration of ^{31}P spectra of the bicelle solutions. The small difference

in DMPC/DHPC ratio is not expected to have significant effects. Second, the two preparations have different ratios of uncharged to charged lipid. In preparation 1, the ratio of DMPC/CTAB was 19:1, while in preparation 2 the ratio was 34:1. As a result of the slightly different DMPC/DHPC ratios and the different DMPC/CTAB ratios, the two sample preparations undergo their ordered to disordered transitions at slightly different temperatures and experience optimal ordering at 34 and 39 °C, respectively.

Protein Preparation. The ^{15}N-labeled protein was expressed and purified as previously described (*17*) and then concentrated using a Centricon 3-SR microconcentrator (Amicon, Inc., Beverly, MA) to ~200 μL. The concentrate was subsequently diluted to ~400 μL using a 10% solution of bicelle preparation 1. The final protein concentration was estimated to be ~0.3 mM in preparation 1. BLBC was recovered from preparation 1 using a 50 kDa cutoff Centricon 50 concentrator that allowed the protein and some DHPC/CTAB to pass through, while retaining the bulk of the bicelle medium. The protein was reconcentrated and separated from any DHPC/CTAB using a 3 kDa cutoff Centricon 3-SR. The protein was then diluted with a 10% solution of bicelle preparation 2, resulting in a protein concentration of approximately 0.1 mM.

The protein gave reasonably well-resolved ^{15}N−^{1}H heteronuclear single quantum coherence (HSQC) spectra despite the low concentration and complex bicelle mixture. The spectra of BLBC dissolved in the two different bicelle preparations appear virtually identical despite differences in composition and slight differences in acquisition temperatures. Comparing the 2D {^{1}H, ^{15}N} peak positions indicates an average deviation of -0.00 ± 0.02 ppm (70% confidence interval) in the proton dimension. Similarly, the nitrogen dimension showed an average deviation of 0.0 ± 0.1 ppm. This suggests that if differences are observed in orientational properties for the two preparations, they are due to differences due to alignment as dictated by the bicelle medium and are not due to any structural changes within the domains of BLBC itself.

Measurement of Residual Dipolar Couplings. Proton−nitrogen dipolar couplings are typically extracted from ^{15}N-labeled samples in one of two ways, both of which were employed for this study. One method is based on frequency domain analysis of a modified proton−nitrogen HSQC spectrum. The proton−nitrogen dipolar (D) and scalar (J) couplings are simply not decoupled and are allowed to evolve during the ^{15}N chemical shift evolution period, resulting in a doublet in the nitrogen dimension. The experiment is performed twice: once, at a temperature below the liquid crystal transition temperature at which the bicelle/BLBC solution is not oriented, and once, at a temperature at which the bicelles and BLBC molecules are optimally aligned. The first experiment, herein run at 25 °C, results in splittings due to scalar couplings alone (J Hz), while the second, 34 or 39 °C, shows splittings due to the sum of scalar and dipolar couplings ($J + D$ Hz). Taking the difference allows the dipolar couplings to be extracted.

A second method that avoids the spectral crowding due to the doubling of the peaks in the nitrogen dimension was developed in this lab (*20*). In this method, a ^{1}H−^{15}N HSQC is modified to allow the proton−nitrogen coupling to evolve for a constant time period. The resulting signal intensity is modulated by either the sine or the cosine of the effective coupling. If the scalar couplings, J, are known or separately measured as in this case from a J-modulated HSQC of an isotropically oriented sample, the ratio of the signal intensities from the sine and cosine modulated experiments can, again, be used to extract the residual dipolar couplings. Experimental details are summarized in Table 1.

Table 1: NMR Spectral Aquisition Parameters

Gradient Selected-Coupled-HSQC[a]		
	preparation 1	preparation 2
temp:	25 °C/34 °C	25 °C/39 °C
field strength:	500 MHz	600 MHz
spectral width (^{1}H/^{15}N):	6000 Hz/2000 Hz	8000 Hz/2000 Hz
points (^{1}H/^{15}N):	2048/256	2048/160
quadrature method:	gradient selected	gradient selected
pulse lengths (^{1}H/^{15}N):	8.35 μs/31 μs	7.1 μs/42 μs
no. of transients:	32	310
total experimental time:	8 h	50 h
J-Modulated-HSQC[b]		
	preparation 1	preparation 2
temp:	25 °C/34 °C	25 °C/39 °C
field strength:	500 MHz	600 MHz
spectral width (^{1}H/^{15}N):	6000 Hz/2000 Hz	8000 Hz/2000 Hz
points (^{1}H/^{15}N):	2048/80	2048/80
quadrature method:	states-TPPI	states-TPPI
pulse lengths (^{1}H/^{15}N):	8.35 μs/30 μs	7.2 μs/42 μs
water selective G3 pulse:	2.1 ms	2.0 ms 2200 Hz bandwidth
No. of transients	168	300
total experimental time:	30 h	62 h

[a] All experiments run with water flip-back pulses, 1.5 s presat, and GARP decoupling during acquisition. [b] All experiments run with 42 ms constant time period.

Data Analysis via Order Matrix Determination. An order matrix approach was used to analyze measured residual dipolar couplings in terms of an order matrix, S. The angular average that appears in the expression for the dipolar coupling between nuclei m and n,

$$D^{mn} = k^{mn}\left(\frac{3\cos^2\Theta^{mn} - 1}{2}\right) \qquad (1)$$

is over the angle θ^{mn} between the magnetic field and the dipolar interaction vector, that is, the internuclear vector connecting nuclei m and n. The constant, k^{mn}, appearing in eq 1 depends on the identities of the participating spins and the distance separating them. It is given by

$$k^{mn} = -\frac{2\mu_o\hbar}{(4\pi)^2}\gamma_m\gamma_n\frac{1}{(r_{mn}^3)} \qquad (2)$$

where μ_o is the permittivity of free space, $\hbar$ is Planck's constant/2π, and γ_m and γ_n are the magnetogyric ratios of the coupled nuclei. The term r_{mn} is assumed to be fixed at a ^{15}N−^{1}H amide bond length of 1.02 Å.

The angular average of eq 1 can be decomposed into averages over the time dependent direction cosines relating the magnetic field to a coordinate frame rigidly affixed to the domain of interest, {cos θ_x, cos θ_y, cos θ_z}, and into the time independent direction cosines {cos φ_x^{mn}, cos φ_y^{mn}, cos φ_z^{mn}} relating the mnth dipolar interaction vector to the same domain frame:

$$D^{mn} = k^{mn} \sum_{i,j=\{x,y,z\}} \left\langle \frac{3\cos\theta_i \cos\theta_j - \delta_{ij}}{2} \cos\varphi_i^{mn} \cos\varphi_j^{mn} \right\rangle \quad (3)$$

Following the formalism due to Saupe (*21*), the averages in the previous expression are identified as the elements of a symmetric, traceless tensor known as the order matrix, *S*, allowing eq 3 to be rewritten as

$$D^{mn} = k^{mn} \sum_{i,j=\{x,y,z\}} S_{ij} \cos\varphi_i^{mn} \cos\varphi_j^{mn} \quad (4)$$

Here and throughout, the convention is adopted that $|S_{zz}| > |S_{yy}| > |S_{xx}|$. It is assumed that the domain or domain core used in this analysis has a rigid, known structure, i.e., that the direction cosines of each internuclear vector in the domain frame, $\cos\varphi_i^{mn}$, are known and are time independent. This restriction need not be too severe since proteins often contain rigid, well-structured domains as in BLBC or well-defined units of secondary structure, such as the helices of myoglobin (*7*). Even units as small as peptide planes can be considered rigid entities provided sufficient data are available for each single peptide.

Given a rigid entity, the order matrix that reproduces a set of observed dipolar couplings can quickly and efficiently be determined using an analytical technique called singular value decomposition. This is done using a software program called ORDERTEN_SVD developed in this lab[2] (*6*). Since an order matrix has five independent elements, at least five dipolar couplings per rigid entity must be observed. In principle, fewer than five measurements can be used, but the resulting indeterminacy of solutions, known as the null space, increases the possible set of acceptable order matrixes. In practice, even five or more observed couplings may not be sufficient to precisely define the order matrix. For example, parallel internuclear vectors, such as the nearly parallel N−H vectors in an α-helix do not provide unique information for the determination of *S*. In short, the greater the number of observed couplings and the greater the range of coupling values observed, the more precisely ORDERTEN_SVD can determine an acceptable order matrix.

Of course, if one has large numbers of measurements, each with experimental imprecision, a single order matrix may not reproduce the observed couplings exactly. In fact, one would like to determine the set of order matrixes that reproduce the observed dipolar couplings to within their specified experimental uncertainties. ORDERTEN_SVD accomplishes this in the following way. A new set of dipolar coupling inputs is generated by randomly choosing each coupling from a distribution around its observed value. The distributions are assumed to be Gaussian with standard deviations equal to the estimated experimental uncertainties. These randomized inputs are used to generate an order matrix. This process is then repeated many times; typically >10 000 cycles are performed.

Finally, an acceptable order matrix can be diagonalized to return its three principal values, or order parameters, which reflect the degree and nature of the possibly anisotropic motion of the rigid entity. In addition, diagonalization allows one to determine the principal axis system of this motion and, from this, the most probable orientation of the rigid entity relative to the ordering force.

Structure Refinement via Simulated Annealing. An alternate method for the analysis of residual dipolar couplings is the use of dipolar couplings as restraints during a molecular dynamics/simulated annealing NMR structure refinement (*18*, *22*, *23*). Herein, we use a software package called CNS (*24*), but it is expected to share characteristics with other simulated annealing protocols. These approaches have been fully described by their authors in the literature. Briefly the simulated annealing approach introduces a residual dipolar coupling restraint energy that is minimized when the dipolar interaction vectors are oriented correctly with respect to a reference coordinate frame or, similarly, when the molecular fragment, as a unit, is correctly oriented. Equations for dipolar coupling are often expressed in a form slightly different from eq 3. In the principal axis system of the order tensor, one may write

$$D_{mn} = D_a(3\cos^2\theta^{mn} - 1) + \frac{3}{2}D_r(\sin^2\theta^{mn}\cos 2\varphi^{mn}) \quad (5)$$

where D_a and D_r are the axial and rhombic components of an alignment tensor and the angles θ^{mn} and ϕ^{mn} are the fixed polar angles of the *mn*th internuclear vector within the principal axis system. The following relationships make the connection between this formalism and the order matrix formalism:

$$k^{mn} S_{zz} = 2D_a$$

$$k^{mn} S_{yy} = -D_a\left(1 + \frac{3}{2}R\right)$$

$$k^{mn} S_{xx} = -D_a\left(1 - \frac{3}{2}R\right) \quad (6)$$

where

$$R = D_r/D_a$$

It can clearly be seen from eq 5 that if an internuclear vector were parallel to the *Z*-axis of the principal axis system, the angle θ^{mn} would be 0° and the dipolar coupling expression would reduce to

$$D^{mn}_{\max} = 2D_a \quad (7)$$

Thus, the axial component of the alignment tensor, D_a, may be estimated from the largest observed dipolar coupling. Alternatively, when $\theta^{mn} = \phi^{mn} = 90°$, a minimum value of the coupling results. In this case, eq 5 is reduced to

$$D^{mn}_{\min} = -k^{mn}\left[D_a + \frac{3}{2}D_r\right] = -k^{mn}D_a\left[1 + \frac{3}{2}R\right] \quad (8)$$

allowing for an estimation of the axial and rhombic components of the alignment tensor, D_a and D_r, respectively, given an estimate or search for the optimum rhombicity, *R* (*18*). Alternately, the axial and rhombic components may

[2] The program ORDERTEN_SVD and documentation are available for download from the Prestegard Laboratory home page at http://tesla.ccrc.uga.edu.

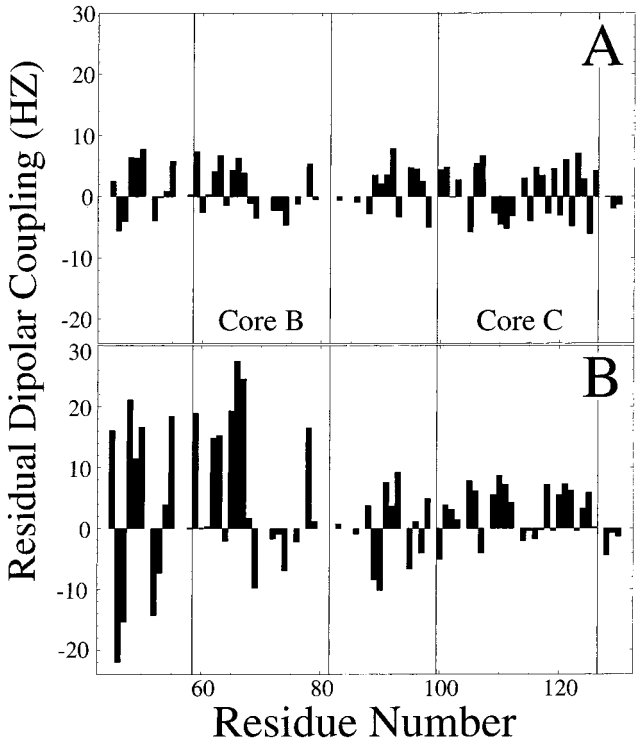

FIGURE 1: Experimentally observed dipolar couplings versus residue for BLBC in bicelle preparations 1 and 2, panels A and B, respectively. Vertical lines delineate the well-structured core regions used in this study, namely, residues 59−81 of domain B and residues 100−126 of domain C.

be deduced from the powder pattern-like distribution of observed dipolar couplings provided that is complete enough (*23*).

The estimates of D_a and D_r are fed to CNS along with a preliminary structure and any other available structural restraint information. Given a tentative orientation of the molecule with respect to a reference coordinate frame, the estimates of D_a and D_r can be used to calculate dipolar couplings and to generate a pseudoenergy

$$E_{\text{residual dipolar coupling}} = k_{\text{RDC}} \sum_{m,n} (D^{mn}_{\text{Measured}} - D^{mn}_{\text{Calculated}})^2 \quad (9)$$

where k_{RDC} is a weighting factor and d^{mn}_{measured} and $d^{mn}_{\text{calculated}}$ are the measured and calculated coupling constants, respectively. This energy is minimized during the subsequent simulated annealing/molecular dynamics calculations to determine the principal axis frame of the alignment tensor and, hopefully, to refine the molecular structure.

RESULTS

Measured Dipolar Couplings. Dipolar couplings were observed for 34 residues of domain B and 33 residues of domain C in BLBC. They are presented graphically in Figure 1 (see Losonczi et al. for experimental details (*6*)). Both positive and negative values are observed ranging from −6.0 to 7.9 Hz for preparation 1 and from −22.0 to 27.4 Hz for preparation 2. The experimental uncertainty is estimated at 0.2 Hz. Note, an observed coupling of nearly 0 Hz is a legitimate observation and, in fact, is a very powerful observation since the $(3 \cos^2 \theta - 1)/2$ angular dependence of the dipolar coupling changes rapidly as a function of θ near its zero crossing, i.e., where θ = the magic angle ≈ 54.74° (see eq 1). No dipolar couplings are reported for residues whose peaks were obscured by spectral overlap preventing precise determination of the coupling.

Order Matrix Analysis Using Simulated Annealing. The dipolar coupling data collected for this work were analyzed using both ORDERTEN_SVD and a CNS-based simulated annealing energy minimization as outlined above. The apparent advantage of the simulated annealing method is that no prior knowledge of the structure of the molecular fragment is required, and crude structures such as that which result from the NOE-based NMR data might actually be simultaneously refined. To exploit this possibility, dipolar coupling information was added in the form of orientational restraints along with all other available structural information into a CNS structure refinement protocol. These data include 279 intraresidue and 265 interresidue NOE-derived distance constraints, the latter including 84 restraints between proton pairs four or more residues apart. Upper and lower bounds for these constraints were chosen as described in our previous work (*17*). Estimates of the alignment tensor components, D_a and D_r, were made from the observed dipolar couplings: a difficult task given the relatively small number of observed couplings. Following suggestions in the literature (*18*), the weighting factor k_{RDC} (eq 9) was allowed to increase slowly over the course of the simulated annealing calculation from 0 to some maximum value. Our results were fairly insensitive to changes in the maximum value of k_{RDC} ranging from 5 to 20 kcal·Hz^{-2}. In every case, structures were obtained that satisfied all the dipolar couplings to within their experimental uncertainties.

Simulated annealing cycles requiring several hours each were successfully performed indicating that a molecular structure and/or orientation that satisfied all restraints, including the dipolar coupling restraints, could be found. However, the resulting orientations of the order tensor principal axis systems relative to the molecular frame showed little or no convergence to a single orientation (see Figure 2, panels Ax, Ay, and Az). This, we believe, was due primarily to the small number of NOE restraints available for BLBC. Only when synthetic ^{13}CO−^{13}CO, ^{13}Cα−^{13}Cα, and ^{1}HN−^{1}HN distance restraints generated from the homologous WGA X-ray structure were introduced to help immobilize the structure was the simulated annealing approach able to reorient the molecular fragment as a whole and to determine the principal axis orientation to reasonable precision. Figure 2, panels Bx, By, and Bz, shows the results for domain B of BLBC when 853 such restraints were added; 459 of the 853 restraints (153 each of ^{13}CO−^{13}CO, ^{13}Cα−^{13}Cα, and ^{1}HN−^{1}HN restraints) were from the well-defined core of domain B, residues 59−81. The synthetic distance restraints were tightly bounded at ±0.2 Å, and residues 58−60, 63−65, 99−101, 108−109, 118−120, and 122−124 were omitted since they are at or near the sites of sequence divergence between WGA and BL/BLBC. Clearly, a molecule may satisfy dipolar coupling restraints in two ways: first, by molecular reorientation; second, by subtly altering the structure near individual N−H bonds so as to satisfy the restraints without globally reorienting the domains. In the case of a molecule such as BLBC that is not restrained by many NOEs, the second option is viable. This highlights one of the potential difficulties involved in simultaneously

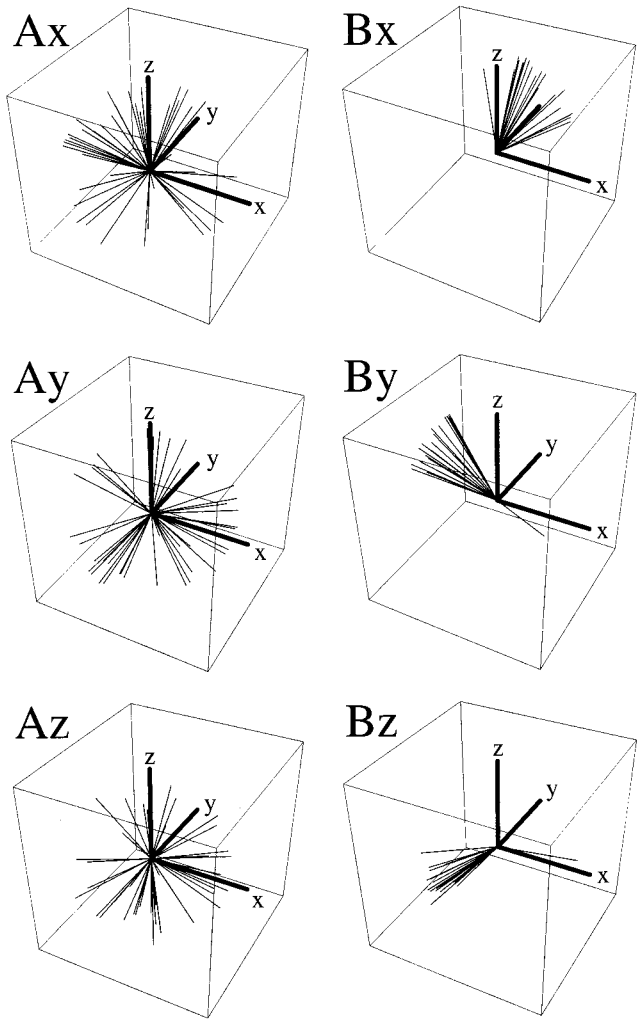

FIGURE 2: Orientations of the domain B order tensor principal axis systems in the molecular frame as determined via simulated annealing. In all panels, the possible orientations of the *x*, *y*, or *z* axes (light lines) are shown in the molecular frame (heavy lines). Panels Ax, Ay, and Az show the set of *x*-, *y*-, and *z*-axis orientations, respectively, determined from the available experimental data. No preferred alignment is readily apparent. Panels Bx, By, and Bz show the set of *x*-, *y*-, and *z*-axis orientations found after combining the experimental data with 853 synthetic distance restraints calculated from the WGA X-ray structure (300 Cα−Cα, 300 CO−CO, and 253 HN−HN restraints for domain B). 153 of each type were from the well-defined core region of domain B. Residues at or adjacent to the sites of sequence divergence between WGA and BLBC were omitted. Similar results (not shown) were obtained for domain C by augmenting the experimental data with 528 Cα-Cα, 528 CO−CO, and 496 HN−HN synthetic restraints.

determining the molecular structure and the order tensor orientation via simulated annealing. This method may become more viable if (1) structural restraints other than those from N−H dipolar couplings are sufficient to produce an extremely high resolution domain structure and (2) order parameters can be accurately estimated from the raw dipolar couplings—difficult in cases of limited dipolar coupling data. Clearly, additional data are more readily available with doubly labeled samples (^{13}C, ^{15}N). In the current study, however, we did not have the luxury of preparing these samples and chose to use data from outside sources to improve domain structure.

Order Matrix Analysis for each Domain Using ORDERTEN_SVD. Given that a domain structure must be defined to a reasonably high degree using data from outside sources, determination of the order tensor orientation and order parameters via the SVD approach presents an efficient alternative to simulated annealing. A domain, secondary structure element, or other fragment that can be considered free of large amplitude or nonuniform internal motions can serve as the required rigidly defined fragment, i.e., a rigid entity whose total motion can be described by a single order matrix. In BLBC, one might choose these entities to be the individual B and C domains in their entireties. However, it is well-known that chain termini and loops in protein structures frequently have large degrees of internal motion. We, therefore, choose more restricted fragments within each domain. Examination of the NMR structures of BLBC indicates a well-defined core region in each domain that contains a β-sheet plus a short α-helix motif. The residues in these motifs, B:His59−Cys74 and C:Leu102−Lys117, showed backbone RMSDs of 1.05 ± 0.40 and 0.44 ± 0.08 Å, respectively. These core regions are also believed to be structurally rigid because of the four disulfide bonds per domain, B:46−61, B:55−67, B:60−74, B:78−83, C:89−104, C:98−110, C:103−117, and C:121−126, two of which "cap" the well-defined core regions. This core region is also defined to higher resolution in the X-ray structure of WGA. For the β-sheet plus short α-helix core region, the backbone RMS errors between the average NMR structure of BLBC and X-ray structure of WGA are 1.47 and 1.17 Å for domains B and C, respectively. The resolution of the NMR structure, the sequence and structural homology with the high-resolution X-ray structure of WGA, and the presence of four stabilizing disulfide bonds per domain provide good justification for considering the well-defined core regions of domains B and C (B:59−81, C:100−126) to be rigid entities suitable for the SVD approach.

In BLBC, 16 of the observed dipolar couplings are from the core of domain B and 21 are from the core of domain C. Through the use of these dipolar couplings (error values of 2−3 Hz were used to allow for imperfections in the structural models) along with the domain structures and domain orientations observed in the WGA X-ray structure, the order tensors for domains B and C were determined independently using ORDERTEN_SVD. The WGA X-ray structure was taken as the structural model of the B and C domains: this model was better able to reproduce the observed dipolar couplings than was the low-resolution NMR structure. As described in a previous work that details the technical implementation of ORDERTEN_SVD (*6*), data that were found to be inconsistent with the assumed WGA structural model were removed from the calculations. It is worth noting that these residues were at or near the locations of sequence differences between WGA and BLBC. In all, 13 of 16 and 21 of 24 observed, dipolar couplings were used for domains B and C, respectively. A typical calculation of 10 000 cycles took less than 60 s on a 150 MHz Silicon Graphics Indy R4400 computer.

The orientations of the resulting order tensor principal axis systems relative to the starting structure coordinate frame are shown in Figure 3 where a mapping technique known as the equal area pseudocylindrical Sauson−Flamsteed projection was used to visualize the surface of a sphere. The top and bottom tips of the map represent +Z and −Z in the starting structure coordinate frame while +X is in the very

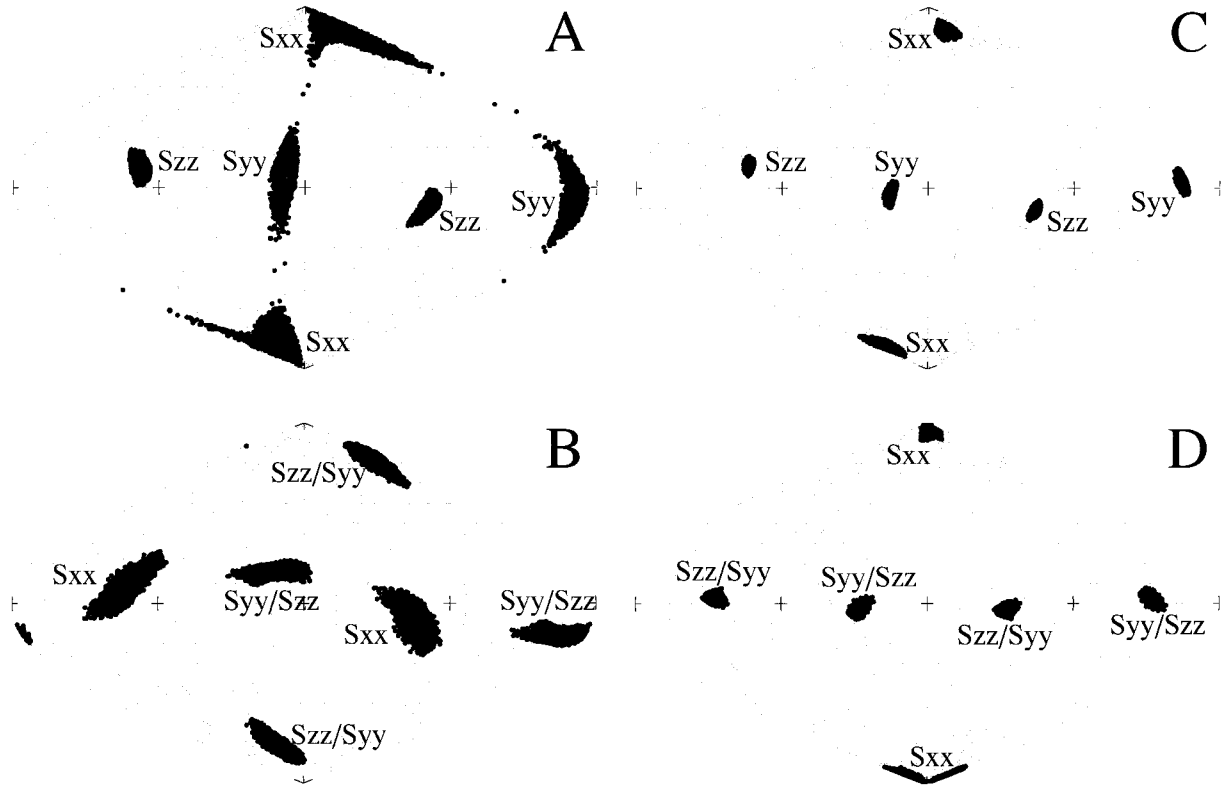

FIGURE 3: Orientations of the order tensor principal axis systems in the molecular frame as determined via SVD. Although the molecular frame is arbitrary, the same coordinate frame is used for both domains. Panels A and B show the results for domains B and C of BLBC in bicelle preparation 1. Panels C and D show the corresponding results for preparation 2. The Sauson−Flamsteed projection used here maps the surface of a unit sphere into a plane by converting latitude (ϕ) and longitude (λ) to Cartesian coordinates (x,y) via $x = \lambda \cos \phi$, $y = \phi$. The horizontal lines of latitude run from −90° to 90° in 10° increments. Vertical curved lines of longitude run from −180° to 180° in 20° increments. Each point in these plots represents the location, in the molecular frame, of the tip of the x, y, or z unit vector of the order tensor principal axis system. The labeling ambiguity in panels B and D is due to asymmetry parameters, η, near 1 as discussed in the text.

center. The positions actually correspond quite closely to those derived using the simulated annealing approach described above and shown in Figure 2, panels Bx, By, and Bz. The corresponding order parameters are shown in Figure 4. A parameter that is useful for characterizing the relationships of these parameters is the asymmetry parameter, η, defined by

$$\eta = \frac{S_{xx} - S_{yy}}{S_{zz}} \tag{10}$$

where S_{xx}, S_{yy}, and S_{zz} are the elements of the order matrix in its principal axis system. When the order parameters are ordered with $|S_{zz}| > |S_{yy}| > |S_{xx}|$, η ranges from 0 to 1 and, thus, parametrizes the asymmetry, that is, the departure from axial symmetry ($\eta = 0$). For domain B, the calculated values of η are broadly distributed about $\eta = 0.58$, indicative of a moderate departure from axial symmetry. For domain C, however, the calculated η values cluster about $\eta = 1.0$, the most nonaxially symmetric situation possible. This explains the labeling ambiguity for domain C in Figures 3 and 4. In the case of $\eta = 1.0$, the most positive and most negative order parameters are nearly equal in absolute value. As the space of possible order matrixes is mapped out, the largest absolute value order parameter is sometimes the negative parameter and sometimes the positive parameter leading to an ambiguity in assignment of S_{zz} and S_{yy}.

DISCUSSION

The derived order parameters and orientations, combined with the dipolar couplings themselves, shown in Figure 1, can be used to investigate the relative domain orientation and dynamics in BLBC. If the two domains of BLBC were rigidly oriented with respect to each other as they appear in the WGA X-ray structure, then the principal axis systems, or alignment frames, of domains B and C should be equivalent to the alignment frame of the entire molecule and, therefore, to each other. Examination of Figure 3 readily shows that this is not the case for either preparation. In fact, even if we choose an assignment for S_{yy} and S_{zz} that is most consistent between the two domains in preparation 1, the x axis must be rotated by nearly 90°. In preparation 2, the principal frame deviations are smaller but still significant.

The data shown in Figure 3 for any single preparation could, in principle, be consistent, with the possibility that the B and C domains are rigidly related to each other but with a different orientation than that observed in the WGA X-ray structure. The correct solution orientation could be recovered by simply reorienting the two domains such that the principal axis frames of the two domains coincide. The resulting structure would be, to a certain extent, putative or arbitrary, since the translational position of the two domains is undefined. However, with appropriate translations, such a structure can be made consistent with all observed short-range structural reporters, such as NOEs and J-couplings, and would be consistent with the observed residual dipolar

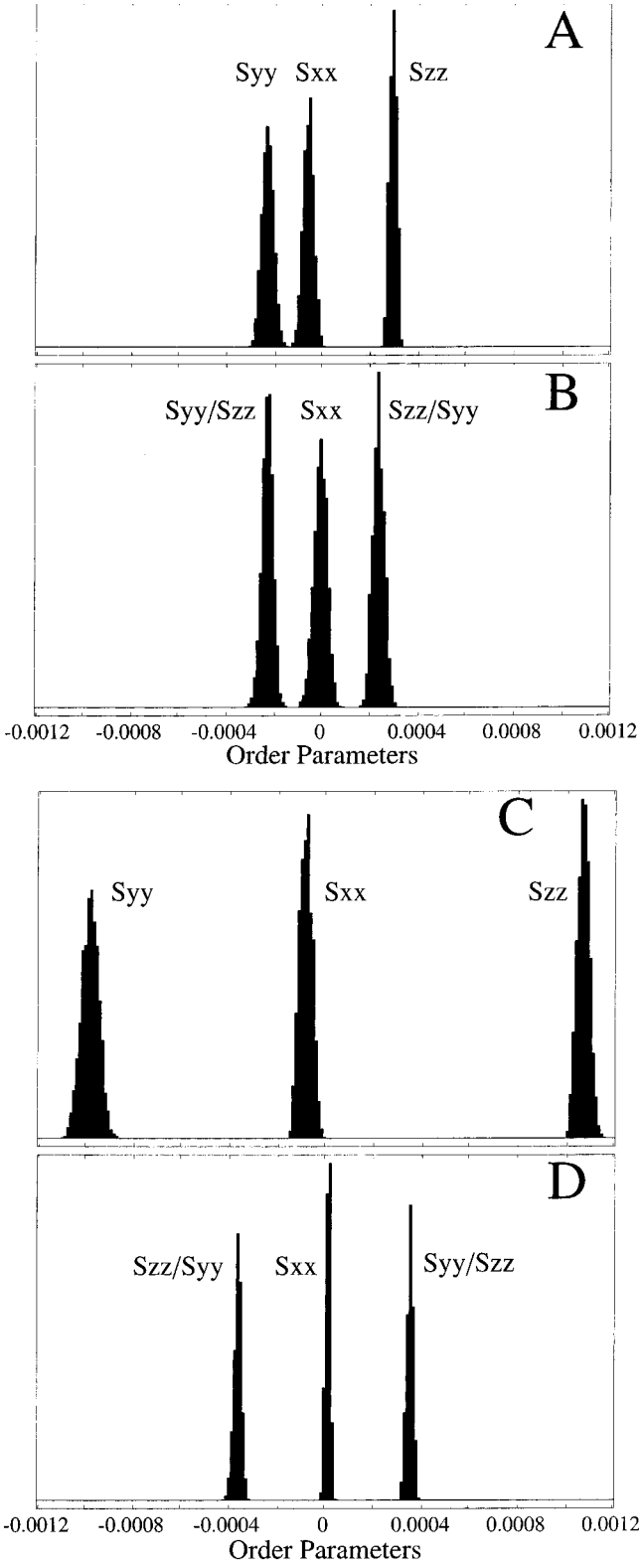

FIGURE 4: Histograms showing the calculated order parameters for domains B and C of BLBC in bicelle preparation 1, panels A and B, and preparation 2, panels C and D. All panels utilize the same horizontal scale to facilitate comparison. Herein, and throughout, order parameters are labeled according to the convention $|S_{zz}| > |S_{yy}| > |S_{xx}|$. This leads to the labeling ambiguity in panels B and D as discussed in the text and also as seen in Figure 3.

couplings. This is demonstrated for BLBC in preparation 1 in Figure 5.

This possibility is also supported by the similar ranges of coupling constants observed for domains B and C in

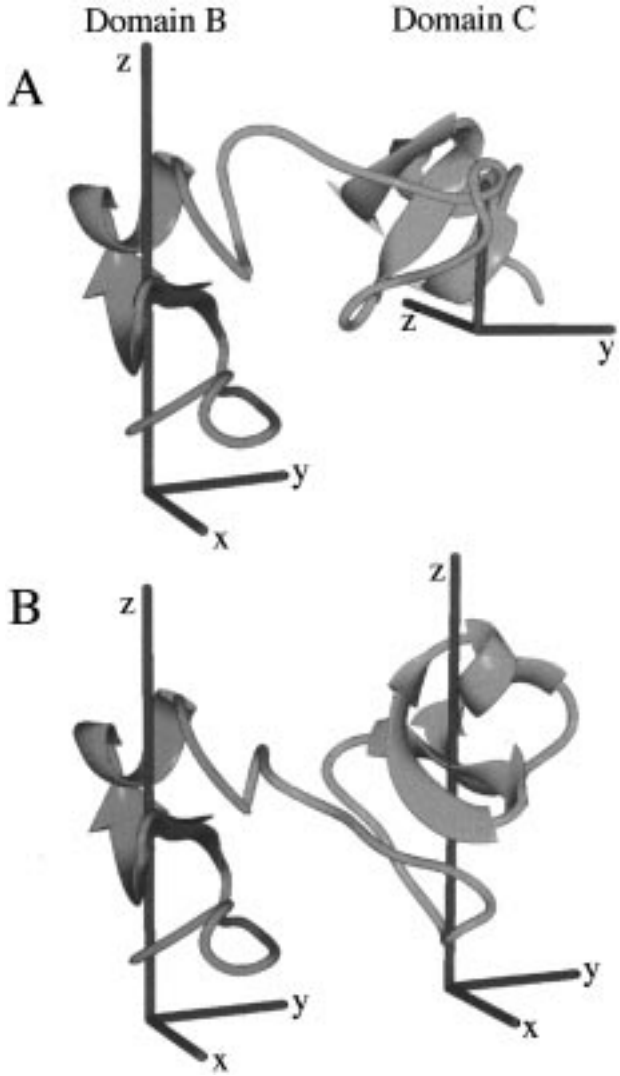

FIGURE 5: The two domains of BLBC. Superimposed on domains B and C are coordinate frames representing the order tensor principal axis systems (PAS) as determined separately for each domain via SVD (the lengths of the axes are arbitrary). Panel A shows BLBC with the domain orientation observed in the X-ray structure of the homologous protein WGA. Clearly, the order tensor principal axis frames differ approximately by a 90° rotation about the PAS *y*-axis. Panel B shows one possible structure of BLBC, obtained by reorienting domain C with respect to domain B, for which the order tensor frames do coincide.

preparation 1 (Figure 1A). If, in fact, the B and C domains were fixed with respect to each other, and one had sufficient couplings to adequately sample all orientations, one would expect to see identical ranges of observed couplings for the two domains. The similarity of ranges observed is also quantitatively reflected in the similarity of principal order parameters plotted in parts A and B of Figure 4. However, there are actually small differences in order parameters for preparation 1, and any apparent support for a simple structural explanation of differences between parts A and B and parts C and D of Figure 3 is completely undermined by the clear magnitude differences between the B-domain and C-domain dipolar couplings in preparation 2, Figure 1B. These more significant differences suggest that the two domains may be dynamically independent and that they may experience different ordering forces as manifested in the differences in the calculated order matrixes.

This dynamic conclusion is further supported by the calculated order parameters shown in Figure 4. In parts C and D of Figure 4, showing the results for preparation 2, the order parameters (S_{xx}, S_{yy}, S_{zz}) are shown to be (−0.000 09, −0.000 98, 0.001 07) for domain B versus (0.000 01, 0.000 35, −0.000 36) for domain C. This indicates that domain B experiences a significantly higher degree of alignment than does domain C. To first approximation, the order parameters are related by a scaling factor of −2.86.

It is important to note that the deviations in magnitudes of order parameters between domain B and domain C cannot be due simply to an overall increase in alignment resulting from an increase in the cooperativity and homogeneity of the bicelle medium since the dipolar couplings and order parameters obtained for domain C are of comparable magnitudes in both preparations. Instead, a mechanism whereby domain B becomes more ordered than domain C must be proposed. One such mechanism is transient binding or association of domain B with the bicelle itself.

As mentioned in the description of the bicelle sample preparations, the bicelle was doped with a positively charged lipid, CTAB, to discourage association of the bicelles and the positively charged BLBC. In fact, prior attempts to prepare BLBC samples in neutral bicelles failed due to aggregation and precipitation of a lipid−protein aggregate. The concentration of CTAB was higher (more positive charge) in preparation 1 than in preparation 2 (DMPC/CTAB = 19:1 and 34:1, respectively). Thus, BLBC would be more likely to associate with the bicelles in preparation 2 than those in preparation 1. Examination of the domain B dipolar couplings and order parameters for preparations 1 and 2 suggests that at least domain B may weakly associate with the bicelles. When the CTAB concentration was lowered in preparation 2, the association appears to become stronger, increasing the magnitudes of the observed couplings and the calculated order parameters, and slightly altering the principal axes system orientation. This can be seen most clearly in Figure 1 where the pattern of observed domain B dipolar couplings is qualitatively the same in preparations 1 and 2, differing in magnitude by a scaling factor. Similarly, the magnitudes of the calculated order parameters (Figure 4) are approximately related by a scaling factor, reflecting the increase in order upon association with the highly ordered bicelle.

The observed dipolar couplings and calculated order parameters also increase for domain C. This is to be expected since domains B and C are tethered and restriction of B should lead to some restriction of C. However, it appears that domain C may not associate so strongly with the well-ordered bicelle since the magnitude change is much smaller (∼30−40% increase in order). Unlike for domain B, the pattern of domain C couplings (Figure 1) is changed completely, indicating a change in the nature of the reorientational dynamics. The dramatic change could result if in preparation 1 each domain was largely ordered by independent interaction with the medium, while in preparation 2, domain B became so well-ordered that the tethering began to dominate overall order of domain C.

A domain specific association, such as that proposed, is possible despite the high level of structural homology between domains B and C. It should be noted that the amino acid sequences of domains B and C are not identical, allowing for differences in protein−bicelle interactions. Furthermore, the two domains are asymmetric in that they are linked head-to-tail: domain B is attached via its C-terminus to domain C whereas domain C is linked via its N-terminus. If, for example, the site of protein−lipid interaction were at the N terminus, domain B would be able to associate whereas domain C might not.

Could such a specific association of loosely tethered domains explain more quantitatively the observed order parameters? As alluded to earlier, the calculated order parameters for domains B and C in preparation 2 (−0.00009, −0.00098, 0.00107) and (0.00001, 0.00035, −0.00036), respectively, are nearly related by a simple scaling factor and exhibit similar order tensor orientations (Figure 3) except that labels for the *zz* and *yy* parameters are interchanged. In general, such an interchange cannot be disregarded. However, in cases such as this, with η near 1, it is not significant. It is merely a manifestation of the relabeling of axes that occurs at $\eta = 1$. It appears, then, that the relative motion of domains B and C may be such that the order parameters of domain C are simply scaled down relative to those of domain B. One possible motion that can produce this effect is wobbling within a cone. In a previous work from this lab (*7*), it was shown that such a motion produces a uniform scaling of all order parameters by an amount

$$S_{\text{cone}} = \frac{1}{2}\cos\psi_{\text{cone}}\,(1 + \cos\psi_{\text{cone}}) \qquad (11)$$

where ψ_{cone} is the half angle of the bounding cone. Thus, to obtain the relative scaling of order parameters observed for domains C and B by cone motion alone, domain C must undergo motion within a cone of half angle ∼60° relative to domain B. Given the length of the unstructured linker region connecting domains B and C up to 18 residues, simple modeling indicates that this degree of relative motion is plausible.

One concern that remains is the high asymmetry in the determined order tensor (i.e., η near 1). One might, at first, expect that if domain B were associating with the bicelle surface, its order matrix would be axially symmetric as is the bicelle itself. This would indeed be the case if domain B were bound only to the faces of the bicelle disks with a single binding geometry. However, calculations indicate that if one BLBC molecule binds through a specific site to the edge of the bicelle disk for every two molecules bound through the same site to the face, the resulting order tensor would be characterized by $\eta = 1$. Other ratios of binding to the face versus edge can produce any value of η from 0 to 1. Similarly, if BLBC were specifically, but loosely, associated and spent part of the time ordered by direct connection to the bicelle and part of the time ordered by other forces in a very different direction, such as steric and electrostatic alignment, the resultant order matrix could assume asymmetry parameters other than $\eta = 0$. Clearly, while the exact nature of the proposed association of domain B with the bicelle cannot be known, the nonaxially symmetric nature of the domain B order tensor need not rule out the possibility of such an association.

CONCLUSIONS

Thus, we conclude that anisotropic spin interactions, such as dipolar couplings, are a rich source of information,

particularly for determining long-range order in those cases where short-range structural reporters, such as NOE and torsion angle information do not suffice. In addition, analysis of dipolar couplings through the order matrix approach, as implemented, for example, in the program ORDERTEN_SVD, reveals dynamic as well as structural information. Application of these methods to the two domain protein BLBC clearly demonstrates the utility of this approach.

While it was expected that BLBC might retain the interdomain orientation observed in the WGA crystal structure, observation of residual dipolar couplings in BLBC aligned in two different dilute bicelle solutions shows that it definitely does not. Moreover, the studies provide information on both the domain orientation and the relative domain motion. The two domains of BLBC are clearly not rigidly oriented with respect to each other and even the average orientations differ from those seen in the WGA X-ray structure. Comparison of the data from the two different bicelle preparations also indicates that a change in the bicelle conditions can result in large changes in the alignment forces. In particular, the selective increase in ordering of domain B relative to domain C seen in preparation 2 allows one to infer interaction of BLBC with the bicelle itself.

REFERENCES

1. Wagner, G. (1997) *Nat. Struct. Biol. 4*, 841−844.
2. Bastian, E. W., Maclean, C., Zilj, P. C. M. V., and Bothner-By, A. A. (1987) *Annu. Rep. NMR Spectrosc. 19*, 35−77.
3. Tolman, J. R., Flanagan, J. M., Kenedy, M. A., and Prestegard, J. H. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9297−9283.
4. Tjandra, N., Grzesiek, S., and Bax, A. (1996) *J. Am. Chem. Soc. 118*, 6264−6272.
5. Prestegard, J. H. (1998) *Nat. Struct. Biol. NMR Suppl.*, 517−522.
6. Losonczi, J. A., Andrec, M., Fischer, M. W. F., and Prestegard, J. H. (1999) *J. Magn. Reson.*, in Press.
7. Tolman, J. R., Flanagan, J. M., Kennedy, M. A., and Prestegard, J. H. (1997) *Nat. Struct. Biol. 4*, 292−296.
8. Tjandra, N., and Bax, A. (1997) *Science 278*, 1111−1114.
9. Clore, G. M., Starich, M. R., and Gronenborn, A. M. (1998) *J. Am. Chem. Soc. 120*, 10571−10572.
10. Hansen, M. R., Rance, M., and Pardi, A. (1998) *J. Am. Chem. Soc. 120*, 11210−11211.
11. Hansen, M. R., Mueller, L., and Pardi, A. (1998) *Nature Struct. Biol. 5*, 1065−1074.
12. Gabius, H.-J. (1997) *Eur. J. Biochem. 243*, 543−576.
13. Lis, H., and Sharon, N. (1998) *Chem. Rev. 98*, 637−674.
14. Mammon, M., Choi, S. K., and Whitesides, G. M. (1998) *Angew. Chem., Int. Ed. Engl. 37*, 2755−2794.
15. Wright, C. S. (1990) *J. Mol. Biol. 215*, 635−651.
16. Wright, C. S. (1977) *J. Mol. Biol. 111*, 439−457.
17. Weaver, J. L., and Prestegard, J. H. (1998) *Biochemistry 37*, 116−128.
18. Clore, G. M., Gronenborn, A. M., and Tjandra, N. (1998) *J. Magn. Reson. 131*, 159−162.
19. Losonczi, J. A., and Prestegard, J. H. (1998) *J. Biomol. NMR 12*, 447−451.
20. Tolman, J. R., and Prestegard, J. H. (1996) *J. Magn Res., Ser. B 112*, 245−252.
21. Saupe, A. (1968) *Angew. Chem., Int. Ed. Engl. 7*, 97.
22. Tjandra, N., Omichinski, J. G., Gronenborn, A. M., Clore, G. M., and Bax, A. (1997) *Nat. Struct. Biol. 4*, 732−738.
23. Clore, G. M., Gronenborn, A. M., and Bax, A. (1998) *J. Magn. Reson. 133*, 216−221.
24. Brünger, A. T., Adams, P. D., Clore, G. M., Delano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1997−1998) *CNS−Crystallography & NMR System*, Yale University: New Haven CT.

BI9905213