Rapid Identification of Medium- to Large-Scale Interdomain Motion in Modular Proteins Using Dipolar Couplings

Demetrios T. Braddock, Mengli Cai, James L. Baber, Ying Huang, and G. Marius Clore*

Laboratories of Chemical Physics and Molecular Biology National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health Bethesda, Maryland 20892

Received May 21, 2001
Revised Manuscript Received July 26, 2001

Many proteins are modular in nature, comprising two or more independently folded domains.[1] Examples include numerous signal transduction proteins, nucleic acid binding proteins, cell surface receptors, and extracellular matrix proteins. While ^{15}N-{^{1}H} heteronuclear NOE measurements can readily pinpoint regions of high mobility,[2] providing a simple means of delineating domain boundaries, the existence of a flexible linker does not provide evidence one way or the other for the presence or absence of interdomain contacts or fixed versus variable domain orientations. In principle, analysis of heteronuclear relaxation data can provide information on the time scale and amplitude of interdomain motions.[3] However, accurate knowledge of the structures of the domains is an essential prerequisite, and in addition, the relaxation data generally have to be analyzed at multiple field strengths.[3] For example, in the case of calmodulin which consists of two modules connected by a flexible helical linker which is highly mobile in the middle, analysis of ^{15}N relaxation data at two magnetic fields was required to establish the presence of interdomain motion unambiguously on the basis of the relaxation data alone.[3] Here we present a simple method based on residual dipolar couplings (RDCs) measured in a charged liquid crystalline medium to identify medium- to large-scale interdomain motions in the absence of any detailed structural information.

RDCs yield long-range orientational information that can potentially increase the accuracy of structures determined by NMR.[4] RDCs may also provide information on internal motions in single-domain proteins[5a−c] and can be used to ascertain the presence of interdomain motions in multidomain proteins.[5d] In the case of neutral liquid crystalline media, alignment is dominated by steric effects and can be predicted on the basis of solute shape.[6] For charged liquid crystalline media electrostatic interactions also come into play.[6] In the context of a single domain or a multidomain protein with minimal interdomain motion, the average solute orientation with respect to the magnetic field will be described by a single molecular alignment tensor. Under these circumstances, internal motions will have relatively little impact on the RDCs since they scale directly with the Lipari−Szabo[7] generalized order parameter $\boldsymbol{S}$ rather than $\boldsymbol{S}^2$ as in the case of relaxation data.[4a] If, on the other hand, there are medium- to large-scale interdomain, rigid body motions, there exists the distinct possibility that the orientations of the domains will be described by different alignment tensors as a result of either different shapes or charge distributions of the domains.

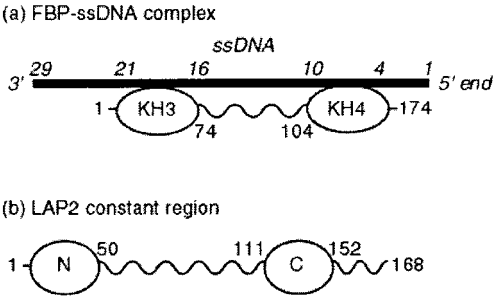

Figure 1. Modular structure of (a) the FBP−ssDNA complex and (b) the LAP2 constant region. The protein linkers, residues 75−103 in (a) and 51−110 in (b), are flexible. The ssDNA in (a) is numbered in italics.

To test this hypothesis we measured backbone RDCs for two systems dissolved in a nematic phase of negatively charged rod-shaped phage particles.[8] The first is a complex of the KH3 and KH4 domains of the FUSE binding protein FBP with a 29mer single-stranded (ss) DNA (Figure 1a).[9] The second is the constant region (residues 1−168) of the nuclear envelope protein LAP2 (Figure 1b).[10] The KH3 and KH4 domains of FBP are connected by a 29-residue flexible linker; bases 4−10 and 16−21 of the ssDNA are bound to KH4 and KH3, respectively, leaving an intervening stretch of five bases.[9] LAP2 comprises two so-called LEM domains at the N (residues 1−50)- and C (residues 111−152)-terminal ends of the molecule connected by a 60-residue flexible linker.[10] Although the protein linkers in the two cases are highly flexible, as judged from small ^{15}N{^{1}H} NOE values, it is not evident a priori that there is no interaction between these various domains. Thus, analysis of the distribution of ^{15}N T_1/T_2 values[11] for the KH3 and KH4 domains of the FBP−ssDNA complex suggests that they both tumble with an apparent rotational correlation time, $\tau_{c,app}$, of ∼11−12 ns and a diffusion anisotropy, A, of ∼1.6 at 35 °C, consistent with the ∼30 kDa molecular weight of the complex. A similar analysis for LAP2 yields values of ∼6 ns and ∼1.5 for $\tau_{c,app}$ and A, respectively, at 35 °C for both domains which again is consistent with a rigid body of molecular weight ∼18 kDa. From the perspective of an NMR structure determination, the presence of NOEs between the two domains would be indicative of minimal interdomain motions and, if sufficient in number, would permit the relative orientation of the two domains to be determined. However, for such large systems it may be difficult to unambiguously assign interdomain NOEs, at least in the initial stages of the structure determination. Further, the failure to observe interdomain NOEs cannot be regarded as absolute proof for the absence of contacts between the domains since this could be potentially attributed to conformational exchange line broadening.

* To whom correspondence should be addressed. Building 5, Room B1-30I, NIDDK, NIH, Bethesda, MD 20892-0510. E-mail: clore@speck.niddk.nih.gov.

(1) Campbell. I. D.; Downing, A. K. *Nat. Struct. Biol.* **1998**, *5*, 496−499.
(2) Tjandra, N.; Kuboniwa, H.; Ren, H.; Bax, A. *Eur. J. Biochem.* **1995**, *230*, 1014−1024.
(3) Baber, J. L.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **2001**, *123*, 3953−3959.
(4) (a) Tjandra, N.; Omichinski, J. G.; Gronenborn, A. M.; Clore, G. M.; Bax, A. *Nat. Struct. Biol.* **1997**, *4*, 732−738. (b) Clore, G. M.; Starich, M. R., Bewley, C. A.; Cai, M.; Kuszewski, J. *J. Am. Chem. Soc.* **1999**, *121*, 6513−6514. (c) Clore, G. M. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9021−9025. (d) Skrynnikov, N. R.; Goto, N. K., Yang, D. W.; Choi, W. Y.; Tolman, J. R.; Mueller, G. A.; Kay, L. E. *J. Mol. Biol.* **2000**, *295*, 1265−1273.
(5) (a) Tolman, J. R.; Flanagan, J. M.; Kennedy, M. A.; Prestegard, J. H. *Nat. Struct. Biol.* **1997**, *4*, 292−297. (b) Tolman, J. R.; Al-Hashimi, H. M.; Kay, L. E.; Prestegard, J. H. *J. Am. Chem. Soc.* **2001**, *123*, 1416−1424. (c) Meller, J.; Prompers, J. J.; Peti, W.; Griesinger, C.; Brüschweiler, R. *J. Am. Chem. Soc.* **2001**, *123*, 6098−6107. (d) Fischer, M. W. F.; Losonczi, J. A.; Weaver, J. L.; Prestegard, J. H. *Biochemistry* **1999**, *38*, 9013−9022.
(6) Zweckstetter, M.; Bax, A. *J. Am. Chem. Soc.* **2000**, *122*, 3791−3792.
(7) Lipari, G.; Szabo, A. *J. Am. Chem. Soc.* **1982**, *104*, 4546−4559.
(8) (a) Clore, G. M.; Starich, M. R.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1998**, *120*, 10571−10572. (b) Hansen, M. R.; Mueller, L.; Pardi, A. *Nat. Struct. Biol.* **1998**, *5*, 1065−1074.
(9) Braddock, D. T.; Baber, J. L.; Louis, J. M., Levens, D. A.; Clore, G. M. Manuscript submitted.
(10) Cai, M.; Huang, Y.; Ghirlando, R.; Wilson, K. L.; Craigie, R.; Clore, G. M. *EMBO J.* **2001**. In press.
(11) Clore, G. M.; Gronenborn, A. M.; Szabo, A.; Tjandra, N. *J. Am. Chem. Soc.* **1998**, *120*, 4889−4890.

10.1021/ja016234f
Published on Web 08/09/2001

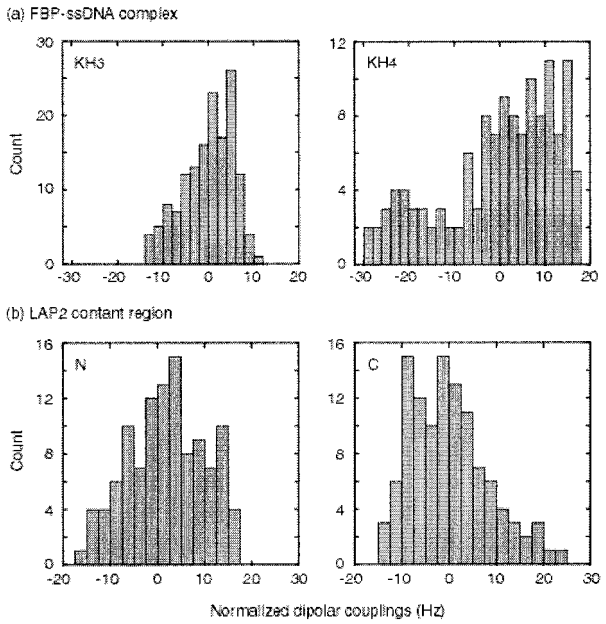

Figure 2. Histograms of the distributions of normalized backbone RDCs for (a) the KH3 and KH4 domains in the FBP−ssDNA complex and (b) the N- and C-terminal LEM domains in LAP2.[1−168] The dipolar couplings[12] have been normalized relative to the $^1D_{NH}$ couplings.[13]

The distribution of normalized backbone RDCs ($^1D_{NH}$, $^1D_{NC'}$, $^2D_{HNC'}$, and in the case of LAP2, $^1D_{C\alpha H}$ as well) for the domains is displayed in Figure 2.[12] The magnitude of the alignment tensor is readily derived from the maximum, minimum, and mode of the distribution.[13] Simple inspection of the histograms indicates that in both examples the magnitude of the alignment tensor (that is the value of the axial and rhombic components of the tensor) is very different for the two domains. Thus, in the case of the FBP−ssDNA complex (Figure 2a), the axial component (D_a^{NH}) of the tensor for the KH3 domain (−7.2 Hz) is approximately half that for the KH4 domain (−14.5 Hz), while the rhombicity η (D_r/D_a) is similar for the two domains (0.25 and 0.16, respectively). In the case of LAP2 (Figure 2b), D_a^{NH} is negative for the N-terminal domain (−7.4 Hz) but positive for the C-terminal one (11.5 Hz); simultaneously, the alignment tensor is almost fully rhombic ($\eta = 0.62$) for the N-terminal domain but nearly axially symmetric ($\eta = 0.15$) for the C-terminal one. The only explanation for these results is that in each case the two domains are oriented either independently or semi-independently by the liquid crystalline medium.

For both examples, the two domains are very similar in structure and shape. The backbone rms differences between the KH3 and KH4 domains, and between the N- and C-terminal domains of LAP2 lie in the 1.0−1.4 Å range.[9,10] Discrimination in alignment, in this instance, is therefore principally electrostatic in origin. For the FBP−ssDNA complex, the negative charge associated with KH3 is significantly greater than that associated with KH4. Although only bases 4−10 are bound to KH4 and bases 16−21 to KH3, for the purposes of the present experiments one can assume that bases 1−13 and 13−29 will impact the alignment of KH4 and KH3, respectively (Figure 1a). Thus, KH3 is effectively associated with an additional four negative charges (arising from the phosphates on the DNA backbone) relative to KH4, which is more than sufficient to account for a 2-fold decrease in the value of D_a^{NH} while leaving the rhombicity only slightly altered. In the case of the N- and C-terminal domains of LAP2, which display ∼25% sequence identity, the distribution of charged residues on the surface of the protein is completely different,[10] thereby accounting for both the difference in rhombicity and sign of D_a^{NH}.

How much interdomain motion is required for domains to be differently aligned by a liquid crystalline medium? To quantify the interdomain motion in the FBP−ssDNA complex, ^{15}N relaxation data were recorded at 600 and 750 MHz and analyzed on the basis of the refined solution structure of the complex[9] using the extended Lipari−Szabo approach,[14] as described by Baber et al.[3] This analysis yields an overall rotational correlation time and diffusion anisotropy of 21.5 ns and 1.85, respectively, for the whole complex. Interdomain motion is described by an effective correlation time τ_s and squared-order parameter S_s^2 for each domain: for KH3, τ_s and S_s^2 have values of 4.1 ns and 0.67, respectively; for KH4, the corresponding values are 3.6 ns and 0.7, respectively. The values of S_s^2 correspond to the KH3 and KH4 domains wobbling independently in cones with semi-angles of ∼30°.

We have not carried out a similar analysis for LAP2. However, thrombin cleavage of LAP2 yields two fragments corresponding to residues 1−86 and 87−168. The average molecular weight of thrombin-cleaved LAP2 derived from analytical ultracentrifugation is ∼9 kDa, indicative of the presence of two noninteracting domains.[10] The correlation coefficient for the $^1D_{NH}$ RDCs measured on intact LAP2 and thrombin-cleaved LAP2 is >0.99 for both the N- and C-terminal domains. We can therefore conclude that the N- and C-terminal domains in intact LAP2 are oriented completely independently by the liquid crystalline medium.

We have shown that the presence of medium- to large-scale interdomain motion, as exemplified by the FBP−ssDNA complex and LAP2, respectively, can be readily ascertained, independent of structure, from the distribution of backbone RDCs measured in charged liquid crystalline media where solute alignment is determined on the basis of not only shape but also charge distribution. This provides a powerful, yet simple diagnostic tool for the identification of significant interdomain motion, thereby facilitating both structure determination and a detailed analysis of relaxation data aimed at quantifying the amplitude and time scales of these interesting motions.

Acknowledgment. We thank N. Tjandra for useful discussions. This work was supported in part by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health (to G.M.C.).

JA016234F

(12) Dipolar couplings were measured at 35 °C by taking the difference in *J* couplings measured in liquid crystalline medium (25 mg/mL fd phage for the FBP−ssDNA complex and 15 mg/mL pf1 phage for LAP2$^{1-168}$) and isotropic (water) medium. The samples consisted of ∼1 mM [U-^{15}N/^{13}C] FBP−ssDNA complex in 50 mM sodium phosphate, pH 7; and ∼1 mM [U-^{15}N/^{13}C] LAP2$^{1-168}$ in 400 mM NaCl, 50 mM sodium phosphate, pH 7.2. For the FBP−ssDNA complex, 156 dipolar couplings were obtained for the KH3 domain (61 $^1D_{NH}$, 49 $^1D_{NC'}$ and 46 $^2D_{HNC'}$) and 140 for the KH4 domain (61 $^1D_{NH}$, 39 $^1D_{NC'}$ and 40 $^2D_{NC'}$). For LAP2,$^{1-168}$ 123 dipolar couplings were measured for the N-terminal domain (39 $^1D_{NH}$, 36 $^1D_{C\alpha H}$, 24 $^1D_{NC'}$ and 24 $^2D_{HNC'}$) and 102 for the C-terminal domain (34 $^1D_{NH}$, 29 $^1D_{C\alpha H}$, 19 $^1D_{NC'}$ and 20 $^2D_{HNC'}$). Note there is no differential increase in line width in the presence of phage.

(13) Clore, G. M.; Gronenborn, A. M.; Bax, A. *J. Magn. Reson.* **1998**, *133*, 216−221.

(14) Clore, G. M.; Szabo, A.; Bax, A.; Kay, L. E.; Driscoll, P. C.; Gronenborn, A. M. *J. Am. Chem. Soc.* **1990**, *112*, 4989−4990.