Decoding the Components of Dynamics in Three-Domain Proteins

Published in final edited form as:
J Comput Chem. 2014 March 15; 35(7): 518–525. doi:10.1002/jcc.23510.

Decoding the Components of Dynamics in Three-Domain Proteins

Mateusz Maciejewski[1], **Paul N. Barlow**[2], and **Nico Tjandra**[3,*]

[1] Novartis Institutes for Biomedical Research, 250 Massachusetts Ave., Cambridge MA, 02139

[2] School of Chemistry, Joseph Black Building, West Mains Road, Edinburgh, Scotland EH9 3JJ

[3] Laboratory of Molecular Biophysics, National Heart, Lung, and Blood Institute, National Institutes of Health, 50 Center Drive, Bethesda, MD 20892

Abstract

In this study we examine the feasibility and limitations of describing the motional behavior of three-domain proteins in which the domains are linearly connected. In addition to attempting a determination of both the internal and overall re-orientational correlation times, we investigate the existence of correlations in the motions between the three domains. Since in linearly arranged three-domain proteins there are typically no experimental data that can directly report on motional correlation between the first and third domain, we address this question by dynamics simulations. Two limiting cases occur: 1) for weak repulsive potentials and 2) when strong repulsive potentials are applied between sequential domains. The motions of the first and third domains become correlated in the case of strong inter-domain repulsive potentials when these potentials do not allow the angle between the sequential domains to be smaller than about 60°. Although various modeling approaches are available, we chose to use the model-free and extended model-free formalisms of Lipari and Szabo due to their widespread application in the study of protein dynamics. We find that the motional behavior can be separated into two components; the first component represents the concerted overall motion of the three domains, and the second describes the independent component of the motion of each individual domain. We find that this division of the motional behavior of the protein is maintained only when their timescales are distinct and can be made when the angles between sequential domains remain between 60° and 160°. In this work, we identify and quantify inter-domain motional correlations.

[*]Corresponding author: Nico Tjandra. Building 50, Room 3503, NHLBI, NIH, Bethesda, MD 20892. Phone: +1 (301)-402-3029. Fax: +1 (301)-402-3405. tjandran@nhlbi.nih.gov.

Supporting Information. Tables S1A-D showing the results of model selection in the simulations in which equal potentials were applied to both interdomain hinges (cone potential results in A, cosine potentials at *i* between 1 and 3 in B-D); Table S2 showing the MF parameters from the simulations in which the MF description was preferred; Table S3A-D showing the EMF parameters that were obtained from the co-fits with the *a*-*b*+*c* vectors in the simulations in which EMF preference was observed (results for the cone potential in A, cosine potential results at *i* between 1 and 3 in B-D); Table S4 showing the EMF parameters obtained from the co-fits with the *a*-*b*+*c* vectors in the simulations in which unequal hard cone potentials were applied to the interdomain hinges, and in which the EMF preference was observed for all domains; Table S5 showing the EMF parameters obtained from the co-fit with the *a*-*b*+*c* vector in the simulation in which one of the terminal domains diffused twice slower then the remaining two domains. This material is available free of charge via the Internet.

Introduction

Many proteins rely on interdomain mobility within linear chains of three or more domains to recognize and bind to other proteins. We previously attempted to explore motions between some of the 20 domains of complement factor H that are collectively critical to its destructive engagement with its principal target, complement component C3b [1,2]. Characterizing motions in multi-domain proteins, although challenging, has the potential for more profound understanding of their functions [3]. Ever more detailed dynamic information on such proteins can be obtained in solution from a number of spectroscopic methods, including nuclear magnetic resonance spectroscopy at a range of magnetic field strengths [4], but it is not straightforward to parameterize this information. This is a growing problem given the abundance of data resulting from ongoing efforts to improve the resolution of the experimental methods and increase the size limit of the molecules that these methods can reliably characterize [4]. Several different theoretical approaches to address this problem have been developed. Examples include the slowly relaxing local structure model [5,6], which describes the dynamics of solute molecules surrounded by a coating of solvent molecules, and a multiple-state interconversion model that describes conformational exchange (such as that from varying domain orientations) between any number of discrete states [7–9]. Another notable example of analysis of interdomain dynamics is provided by coarse-grained simulations of interdomain motion, such as those carried out to analyze the Pin1 protein [10]. An alternative approach to parameterizing molecular motion that has met with significant success was proposed by Lipari and Szabo [11,12]. Their model-free (MF) formalism has been used to analyze dynamics of proteins, and in particular to extract parameters from NMR relaxation data independent of any particular model of the motion [10–14]. This latter formalism has been used to analyze the interdomain flexibility in two-domain proteins [18–21]. Here we evaluate parameters that characterize the motion of three-domain proteins.

The MF formalism is based on the assumption that one can separate the overall and internal motion as:

$$C(t) = C_0(t)\, C_I(t) \quad (1)$$

$C_0(t)$ represents the overall motion correlation function and $C_I(t)$ the internal motion correlation function. $C_I(t)$ can be rewritten as a sum of exponentials, well-approximated if only two first exponential terms are considered:

$$C_I(t) = S^2 + \left(1 - S^2\right) \cdot e^{-t/\tau_i} \quad (2)$$

Therefore, dynamics described by the MF approach are characterized by three parameters: the generalized order parameter S; the correlation time for overall motion τ_M; and an effective internal motion correlation time τ_i.

Clore *et al.* [18] described a case that tested the limits of the two-exponential approximation of MF. In their study of staphylococcal nuclease and interleukin-1β, they found groups of

residues whose relaxation data was poorly addressed by MF. In order to improve this fit, they expanded MF by an additional exponential, effectively separating the internal motion into a fast component, characterized by correlation time τ_f, and a slow component, characterized by the correlation time τ_s, as shown in Eq. (3):

$$C_I(t) = S^2 + \left(1 - S_f^2\right) \cdot e^{-t/\tau_f} + \left(S_f^2 - S^2\right) e^{-t/\tau_s} \quad (3)$$

Residues that did not fit the MF two-exponential approximation were readily explained with the use of the "extended model-free" (EMF) formalism [18] described by Eq. (3). It should be noted that the framework of this formalism only describes rotational diffusion. Additionally the legwork of EMF is based on isotropic diffusion [11,18], but it has also been used in studies focusing on molecules that consist of highly anisotropic domains, such as Calmodulin [20,21]. In one study [22] it was shown that diffusion properties of a two-domain protein consisting of GB1 domains connected with a flexible linker is significantly altered compared to single GB1 domains, even at linker sizes as long as 24 residues. This is indicative of inter-dependence of motion of the domains, suggesting that EMF can be applicable even in the cases in which the domains are connected by long linkers.

EMF has been successfully applied as a tool to study interdomain motion in two-domain systems [20,21]. A similar framework is needed for proteins consisting of three domains and more, since the data that reflects the dynamics of such proteins can currently be analyzed only to a limited degree, especially when it comes to interdomain motion. A case in point is the study of FLNa16-21, whose structure resembles a three-bladed propeller [23]. In this study the ranges of the relative motion of the three propeller "blades" could not be obtained directly from the NMR relaxation data. This information could only have been obtained by using a proper framework to describe the dynamics.

In this study, we investigate the dynamics of three-domain proteins. First we use Brownian methods [24,25] to simulate the rotational diffusion of these domains coupled by repulsive potentials. Then, we analyze the correlation functions of the trajectories of the domains. Our motivation is three-fold: First, to study the dynamics of three-domain (or larger) systems. While such studies are rare, reports featuring multi-domain systems are beginning to appear [1,2,23,26]. Second, to understand how restrictions in the relative motion of domains, by the presence of repulsive potentials, can influence each other and are reflected in the correlation functions [19]. Third, to learn whether such coupling of domain motions could be captured within the current scope of MF and EMF, as was the case in the two-domain motion study of Chen and Tjandra [19].

We fit the time-correlation functions of the domain trajectories to various multi-exponential functions for three-domain proteins to test if coupled three-domain motions can be interpreted using the EMF formalism. It is important to assess whether the overall motion of any one of the mutually coupled domains can be characterized, and if so whether it can be described by a single decaying exponential term. Here we determine the ranges of interdomain motion in which the domains behave like individual proteins, and those in which the domains move in an interdependent fashion.

Methods

Simulation of rotational diffusion of domains

Here we represent a linear three-domain protein by unit vectors, *a*, *b*, and *c*, and we simulate their rotational diffusion using an approach based on a Brownian algorithm [24,25], in which the trajectory of each domain is affected by repulsive potentials based on the proximity of neighbor domains. In our protein, domains *a* and *c* are symmetrically connected to the opposing ends of domain *b*, and thus their rotational motion is (in the case of domain *a*):

$$a_{i+1}=a_i-[2a_i+(b_i-a_i cos\theta_{ab})\,(\mathrm{d}U/dcos\theta_{ab})]\,D^0\Delta t+[R-a_i\,(R\cdot a_i)]\,\sqrt{2D^0}\Delta t \quad (4)$$

Domain *b*, however, is influenced by domains *a* and *c* in a pairwise-additive fashion [24]:

$$b_{i+1}=b_i-\lfloor 2b_i+(a_i-b_i cos\theta_{ab})\,(\mathrm{d}U/dcos\theta_{ab})+(c_i-b_i cos\theta_{bc})\,(\mathrm{d}U/dcos\theta_{bc})\rfloor\,D^0\Delta t+[R-b_i\,(R\cdot b_i)]\,\sqrt{2D^0\Delta t} \quad (5)$$

where θ_{ab} and θ_{bc} are the angles between domains *a* and *b*, and *b* and *c*, respectively, D^0 is the diffusion constant of each of the domains, Δt is the time step, *U* is the potential between the pair of domains, and *R* is a random normally distributed number with variance of 1.

We have generated trajectories based on Eqs. (4) and (5) that consisted of $5{\cdot}10^8$ steps, with time step Δt of 1 ps, in which the pair of terminal domains (*a* and *c*) and domain *b* were initiated at opposite polar points of a spherical coordinate system. The simulations were carried out under two types of potential *U*. The first type was a hard cone potential, defined in Fig. 1. Under this potential, the domains are free to diffuse within a cone of semi-angle β with respect to an axis that is defined by the relative orientation of the domains (such that within this cone $\mathrm{d}U/\mathrm{d}\cos\theta = 0$). If a diffusive step results in the pair of domains breaching their cone, all three domains are reset to their previous positions. This potential is an extension of the hard cone potential defined earlier [19]. While Chen *et al.* [19] studied a two-domain system that was coupled under a hard-cone potential, in the linear three-domain system each of the two pairs of sequential domains interacts under such a potential (domain *a* with domain *b*, and domain *b* with domain *c*). We ran 18 independent simulations with the hard-cone potential applied symmetrically between domains *a* and *b*, and *b* and *c* (such that the semi-angles of both cones were equal), at β (= β_{ab} = β_{bc}) in the range between 5° and 90°, at 5° intervals. Another type of potential that we used was a cosine potential defined as:

$$U_i=(-1)^{i+1}k cos^i\theta \quad (6)$$

We ran 30 simulations under symmetric cosine potentials at varied scaling factors *k* = {1 to 10} and *i* = {1 to 3}. In these simulations both inter-domain angles were constrained with β = 45°.

Additionally, we ran six simulations in which the cone potentials of various strengths were applied between respective domains. These simulations were run using hard cone potentials

with the following angles: (β_{ab} = 90°, β_{bc} = 30°), (β_{ab} = 70°, β_{bc} = 30°), (β_{ab} = 60°, β_{bc} = 30°), (β_{ab} = 50°, β_{bc} = 30°), (β_{ab} = 40°, β_{bc} = 30°), (β_{ab} = 5°, β_{bc} = 30°).

Calculation and fitting of the correlation functions

The auto-correlation function is defined here for domain *a* as [27]:

$$C(t) = \left\langle \frac{1}{2}\left(3 \cdot (a_\tau \cdot a_{\tau+t})^2 - 1\right)\right\rangle \quad (7)$$

where the angular brackets denote the average taken over all times τ, a_τ is the orientation of domain *a* at time τ, and $a_{\tau+t}$ is its orientation at time $\tau+t$. We have calculated correlation functions of the trajectories of each of the domains in all simulations numerically, according to Eq. (7), in the τ range between 0 and $1{\cdot}10^5$ ps and the *t* intervals of 100 ps. The calculated correlation functions were fit to single-, double-, and triple-exponential curves using the constrained Levenberg-Marquardt routine available from the Grace software [28] via the models presented below.

$$C(t) = exp\left[-t/d_0\right] \quad (8a)$$

$$C(t) = p_0 \cdot exp\left[-t/d_0\right] + (1-p_0) \cdot exp\left[-t/d_1\right] \quad (8b)$$

$$C(t) = p_0 \cdot exp\left[-t/d_0\right] + (p_1-p_0) \cdot exp\left[-t/d_1\right] + (1-p_1) \cdot exp\left[-t/d_2\right] \quad (8c)$$

The model preference in the cases in which the additional exponential in the more complicated model made a significant contribution (> 5%) was established by means of an *F*-test [29] at the rejection probability threshold P_F of 0.01 (integration of rejection probability was carried out in Mathematica 6.0.1.0). Fittings and model selections were applied to the correlation functions of the trajectories of normalized vectors *a*, *b*, *c*, and *a*−*b*+*c* from each simulation.

Simultaneous fitting

In cases where a double exponential model was preferred for the correlation functions of the trajectories of domains *a*, *b*, and *c* and a single exponential model was preferred for the correlation functions of the trajectory of vector *a*−*b*+*c*, simultaneous fitting of the correlation functions of the individual domains (*a*, *b*, or *c*) and that of vector *a*−*b*+*c* were performed. In such simultaneous fittings the correlation functions of the individual domains were fit to Eq. (8b), and the correlation functions of the *a*−*b*+*c* vector were fit to Eq. (8a). The parameter d_0 was shared in these simultaneous fittings.

Results

The overall motion of the domains in a three-domain protein

The core ideas of MF formalism impose an overall reference frame [11,30]. In the case of EMF, a molecular reference frame is necessary to describe the overall motion of the protein relative to which the domains carry out their slow internal motions. The *a*−*b*+*c* vector was considered as a representation of the frame of interactions of domains in the three-domain simulations, owing to the fact that this vector, connecting the terminal tips of domains *a* and *c*, resembles the overall linear shape of the studied three-domain system. The utility of this vector was tested for the simulations run under symmetric potentials. The correlation times of the *a*−*b*+*c* vectors were expected to reflect the timescale of the overall motion, which for three linearly connected domains is expected to depend on the effective size of the protein that they constitute, or in the present vector analogy, the effective (average) length of the sum of vectors throughout the simulation [19]. This effective length of each of the terminal domains will be given by a projection of their domain vector at the average angle on the axis of linearity. In establishing such an average angle we took into account the fact that both domain *b* and each of the terminal domains can flex between angle 0 and β relative to the reference axis, which constitutes an effective flexion range between 0 and 2β:

$$cos\left[\frac{1}{2\beta}\int_0^{2\beta} arcsin\left(sin\alpha\right)\mathrm{d}\alpha\right]=cos\left[\frac{1}{2\beta}\left[\frac{\alpha^2}{2}\right]_0^{2\beta}\right]=cos\beta \quad (9)$$

If all three domains are considered in this symmetric case, the expected effective overall correlation time is given by:

$$\tau_M=\frac{1+2cos\beta}{6D^0} \quad (10)$$

For example in case of β = 10° the effective overall length is 1 + 2cos$\beta\approx$2.970, and the corresponding expected overall correlation time given by Equation (10) is $\frac{1+2cos\beta}{6\cdot D^0}\cong 30.933$ ns. The correlation times obtained from the fits of vectors *a*−*b*+*c* in the simulations symmetrically restrained with cone potentials were compared to the correlation times expected from Eq. (10). In Fig. 2 it can be observed that an *a*−*b*+*c* vector indeed provides a good representation of the overall motion of the three domains in the whole range of the cone potentials β that were used.

Model fitting in simulations run under symmetric potentials

The model preferences from the fittings to correlation functions of symmetrically restrained domain trajectories have been summarized in Table S1 in Supporting Information. These results indicate that single exponential fits were preferred (and therefore motions of the constituent domains were independent) for simulations ran with symmetric cone potentials β in the range between 65° and 90° (inclusive), and additionally at 5° (at which the domain motions were very strongly inter-dependent). Double exponential fits were preferred

(indicating that the motions of the constituent domains were inter-dependent) for simulations ran with symmetric cone potentials β in the range between 10° and 60° (inclusive). Preference for double exponential fits was also observed in all simulations run under symmetric cosine potentials, a result similar to that observed by Chen and Tjandra [19] for two-domain proteins. Triple exponential fits were not preferred in any of the simulations run with symmetric potentials, indicating that the domain interactions could be characterized using a single exponential (preference of the trajectory of domain *b* always mirrored that of the terminal domains). For all trajectories in which the double exponential model was preferred, the $a-b+c$ vectors displayed a single exponential preference.

Parameters extracted from the symmetric trajectories

MF and EMF formalisms were considered for the interpretation of the simulations. In the MF and EMF formalisms, the fit exponentials constituting the correlation function are interpreted as either internal or overall motions. The fast internal motions that are characterized in the EMF correlation function expression by correlation time τ_f were not simulated in this study; due to this, Eq. (8b) is equivalent to the EMF description and Eq. (8a) corresponds to MF. MF in this study has indeed been reduced to one exponential (and one type of motion). In the cases in which the interaction between the domains is weak (simulations at symmetric cone potentials β between 90° and 65°), this exponential describes the individual independent motion of the domains. At the opposite extreme (simulation at symmetrically applied cone potentials β of 5°), it describes a situation in which the domains move together in an arrangement so rigid that their individual motions make a negligible contribution and their dynamics can be entirely described by the overall concerted motion of the three domains.

In the cases in which the double exponential model is preferred, this overall motion (reliably represented by the $a-b+c$ vector) can be factored out of the correlation function as one of the two exponentials. If such treatment is correct, the remaining exponential should correspond to the internal motion of the domain – if one transforms the entire trajectory into the frame of the overall motion and carries out a single exponential fit to the correlation function of the resulting trajectory, this single exponential will display a decay constant similar to that of the internal motion exponential mentioned above. In the simulations in which the double exponential model was preferred (those at symmetrically applied cone potentials in the range of β between 60° and 10°, as well as those run under cosine potentials) co-fits to the correlation functions of the domain trajectories were carried out with the correlation functions of the $a-b+c$ vector. Motions in this range corresponded to the EMF treatment.

The results of fittings in the range in which MF and EMF were applicable to the symmetrically restrained simulations are given in Supporting Information Tables S2 and S3, respectively.

Timescale separation of the internal and overall motion

The relationship between S^2 and the timescale separation of overall and slow internal motions was plotted in Fig. 3. It can be seen that the timescale separation diminishes with S^2, and that the fitting of the motion parameters becomes ill-defined when the ratio of the

correlation times approaches 1. A similar loss of definition is observed at S^2 values > 0.9. The results presented in Fig. 3 indicate that in highly rigid systems (characterized by high S^2 values) the overall motion of domains is much slower than their individual motions. In the high limit of S^2, the contribution of slow internal motions becomes negligibly small (and ill-defined), and overall motion becomes the vastly dominant descriptor of motion. Conversely, in the highly flexible systems (characterized by low S^2 values) motion of the domains is so highly independent that the timescale of the slow internal motion becomes very similar to that of the overall motion – the two types of motion progressively cannot be distinguished, and begin to both describe the underlying independent motions of domains, rather than provide true separation into overall and slow internal diffusion.

Motions within the overall frame

Trajectories of constituent domains from symmetrically restrained simulations for which the EMF model preference was established were transformed into their overall motion frame. In every simulation domains *a*, *b*, and *c* were transformed into the overall motion frame defined by vector $a-b+c$. Correlation functions were calculated for the trajectories transformed into their overall motion frame. It was expected that these transformed trajectories would be devoid of the overall motion component, and as such would conform to Eq. (2) with their asymptotes reproducing the S^2 of the core trajectories and the correlation times corresponding to τ_s.

Fig. 4A displays a comparison of correlation times τ_s obtained from double-exponential fits to the correlation functions of the original trajectories with those obtained from single-exponential fits (in a form similar to that in Eq. (2)) to the transformed correlation functions. The agreement between the fits from the two cases is high, and gets progressively better as the symmetric cone potentials β increase in strength (i.e. decrease in the value of the cone angle).

The asymptotes of the transformed correlation functions were compared to the S^2 of the core trajectories. This comparison is shown in Fig. 4B, in which it can be seen that (like the fits presented in Fig. 4A) the agreement between the values observed in the correlation functions of the transformed trajectories and those from the correlation functions of the core trajectories increases with the strength of the potential β.

The agreement of the τ_s and S^2 values between the transformed and core trajectories can be seen to diminish at the higher values of β, which is symptomatic of the decreased timescale separations concomitant with such increased flexibility (see Fig. 3).

Limits of motion separation

The correlation functions of domain trajectories from simulations run under cone potentials applied asymmetrically to the interdomain hinges were analyzed. In a simulation that tested the limiting case in which potentials β_{ab} and β_{bc} were 90° and 30°, respectively, domain *a* was not coupled with domain *b*, and domains *b* and *c* were coupled and adhered to the EMF model, closely matching the numerical result reported for a two-domain system [19] ($C(t) = {}^{0.62-t1.88\tau_0} + 0.38e^{-t/0.57\tau_0}$ for each of the two domains). In the simulation in which β_{ab}

was 70° and β_{bc} was 30°, the double exponential fits performed in the correlation function of domain *a* produced results in which the contribution of one of the exponentials was negligible (< 5%), suggesting that the single exponential model was sufficient to describe its motion. In the case of domains *b* and *c*, double exponential preference was observed – this suggested that in this system domain *a* moved independently, while *b* and *c* moved in an inter-dependent fashion. When the *c*-*b* vector was used to represent the frame of the inter-dependent motion of domains *c* and *b*, co-fits performed using this overall motion representation resulted in $C(t) = 0.69e^{-t/1.94\tau 0} + 0.31e^{-t/0.75\tau 0}$ for domain *b*, and $C(t) = 0.64e^{-t/1.94\tau 0} + 0.36e^{-t/0.87\tau 0}$ for domain *c*.

In all other asymmetric simulations, the trajectories of all domains displayed a double exponential model preference, including the simulation in which β_{ab} was 5° and β_{bc} was 30°. Due to this, the inter-dependent motion in these simulations was modeled, as in the case of symmetric potentials, by the *a*−*b*+*c* vector. No triple exponential preference was noted, furthering the applicability of the *a*−*b*+*c* overall motion representation to cases in which the applied potentials are not equal. The results of the analysis of these simulations are given in Supporting Information Table S4.

An additional “asymmetric” simulation was performed, in which the cone potentials were symmetric and equal to 60° on both interdomain hinges, but for domain *a* and *b* the $D^0_{a,b} = 16 \cdot 10^6\ s^{-1}$, while for domain *c* the $D^0_c = 8 \cdot 10^6\ s^{-1}$ (trajectory length of $5 \cdot 10^8$ ps and Δt of 1 ps). Yet again, all domains in this simulation displayed preference for the double exponential model (the obtained parameters are summarized in Supporting Information Table S5). It can be observed from this table that the domain that tumbles more slowly dominates the overall motion – its dynamics consist almost entirely of the overall motion of the protein (for domain *c* $S^2 \approxeq 1$, and $\tau_M \approxeq 1/(6\ D^0_c)$).

Discussion

This work simulated the dynamics of a three-domain system, in which the domains were connected in a beads-on-a-string fashion. Various repulsive potentials were applied between the terminal domains and the central domain, either symmetrically or asymmetrically. These simulations made it possible to distinguish the ranges of orientation in which the motion of the domains was independent versus those in which it was inter-dependent.

The simulations showed that the limiting hard cone potential at and below which all three domains are inter-dependent is 60°. If either of the terminal domains is free to explore a larger conformational space with respect to the central domain, then it becomes disconnected from the overall motion of the protein. The remaining two domains behave in an inter-dependent fashion, as long as they meet the criteria for coupled two-domain protein motion established by Chen and Tjandra [19].

The inter-dependence of the domains in the three-domain system is captured by extended model-free formalism (EMF) [18]. EMF can be used to divide the motion of the domains into overall motion, representing the concerted, inter-dependent component of the motion of the domains, and slow internal motion, representing the independent dynamics of the

domains with respect to the overall motion frame (molecular frame). In the simulations, overall and slow internal motions proved hard to separate. Nevertheless, it was possible to decompose these two types of motion by invoking a particular overall motion frame.

A normalized vector drawn between the termini of the protein (given as $a-b+c$) provided a good frame of overall motion in the considered three-domain protein. Based on this frame, the mixing of overall motion and slow internal motion was deconvoluted across most of the range of potentials in which the trajectories adhered to EMF. This mixing could not be remedied in the cases in which the motion was weakly coupled or in the cases in which the coupling was very strong. For weakly coupled motions, the correct amplitude of the overall motion could not be found, due to very low separation of the timescales of the two types of motion. For strongly coupled motions, the slow internal motion could not be separated out from the dominating overall motion.

To demonstrate the applicability of these simulations we studied interdomain motions in the three N-terminal domains of factor H (FH1-3) [1]. Here it was found that the range of motion was mostly encompassed by a cone of a 40° full-angle. This corresponds to a half-angle of 20°, which corresponds to 2β. At $\beta = 10°$ it is expected that the effective ranges of motion of all three domains will be very similar (i.e. the effective range of motion of the central domain will not be visibly lower, see Table S3A in Supporting Information). This is in excellent agreement with the equal ranges of distributions of residual dipolar couplings in the three domains of FH1-3 in this study [1].

The findings presented in this article will be applicable to the interpretation of future experimental measurements of the interdomain dynamics of multidomain proteins. Interdomain dynamics are important in many proteins, including fibrillin-1 [31], fibronectin [32], integrins [33], titin [34], myosin binding protein [35], and smooth muscle myosin [36]. In addition, many proteins in the complement system [37] (a principal component of innate immunity) contain three-domain stretches where the domains are arranged in a beads-on-a-string fashion with putative interdomain flexibility expected to be important for their activity [38].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Dr. James Ferretti and Dr. Lisa Bond for careful reading of the manuscript. This work was supported by the Intramural Research Program of the NIH, National Heart, Lung, and Blood Institute.

Abbreviations

MF	model free
EMF	extended model free
NMR	nuclear magnetic resonance

References

1. Maciejewski M, Tjandra N, Barlow PN. Estimation of Interdomain Flexibility of N-Terminus of Factor H Using Residual Dipolar Couplings. Biochemistry. 2011; 50:8138–8149. doi:doi: 10.1021/bi200575b. [PubMed: 21793561]
2. Makou E, Mertens HDT, Maciejewski M, Soares DC, Matis I, et al. Solution Structure of CCP Modules 10–12 Illuminates Functional Architecture of the Complement Regulator, Factor H. Journal of Molecular Biology. 2012; 424:295–312. [PubMed: 23017427]
3. Ekman D, Björklund AK, Frey-Skött J, Elofsson A. Multi-domain proteins in the three kingdoms of life: orphan domains and other unassigned regions. Journal of Molecular Biology. 2005; 348:231–243. doi:10.1016/j.jmb.2005.02.007. [PubMed: 15808866]
4. Bascuñán J, Hahn S, Park DK, Iwasa Y. A 1.3-GHz LTS/HTS NMR Magnet–A Progress Report. IEEE transactions on applied superconductivity?: a publication of the IEEE Superconductivity Committee. 2011; 21:2092–2095. doi:10.1109/TASC.2010.2086995. [PubMed: 22081752]
5. Polimeno A, Freed JH. A many-body stochastic approach to rotational motions in liquids: a reassessment of the Hubbard-Einstein relation. Chemical Physics Letters. 1990; 174:338–346. doi:10.1016/0009-2614(90)85356-H.
6. Polimeno A, Freed JH. A many-body stochastic approach to rotational motions in liquids: complex decay times in highly viscous fluids. Chemical Physics Letters. 1990; 174:481–488. doi:10.1016/S0009-2614(90)87183-R.
7. Ryabov Y, Fushman D. Interdomain mobility in di-ubiquitin revealed by NMR. Proteins: Structure, Function, and Bioinformatics. 2006; 63:787–796. doi:10.1002/prot.20917.
8. Ryabov YE, Fushman D. A Model of Interdomain Mobility in a Multidomain Protein. Journal of the American Chemical Society. 2007; 129:3315–3327. doi:10.1021/ja067667r. [PubMed: 17319663]
9. Ryabov Y, Clore GM, Schwieters CD. Coupling between internal dynamics and rotational diffusion in the presence of exchange between discrete molecular conformations. The Journal of Chemical Physics. 2012; 136:034108. doi:10.1063/1.3675602. [PubMed: 22280745]
10. Bernadó P, Fernandes MX, Jacobs DM, Fiebig K, García de la Torre J, et al. Interpretation of NMR relaxation properties of Pin1, a two-domain protein, based on Brownian dynamic simulations. Journal of biomolecular NMR. 2004; 29:21–35. [PubMed: 15017137]
11. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 1. Theory and range of validity. Journal of the American Chemical Society. 1982; 104:4546–4559.
12. Lipari G, Szabo A. Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results. Journal of the American Chemical Society. 1982; 104:4559–4570.
13. Daragan VA, Mayo KH. Motional model analyses of protein and peptide dynamics using 13C and 15N NMR relaxation. Progress in Nuclear Magnetic Resonance Spectroscopy. 1997; 31:63–105. doi:10.1016/S0079-6565(97)00006-X.
14. Korzhnev DM, Billeter M, Arseniev AS, Orekhov VY. NMR studies of Brownian tumbling and internal motions in proteins. Progress in Nuclear Magnetic Resonance Spectroscopy. 2001; 38:197–266. doi:10.1016/S0079-6565(00)00028-5.
15. Jarymowycz VA, Stone MJ. Fast Time Scale Dynamics of Protein Backbones: NMR Relaxation Methods, Applications, and Functional Consequences. Chem Rev. 2006; 106:1624–1671. doi:10.1021/cr040421p. [PubMed: 16683748]
16. Palmer AG III. NMR Characterization of the Dynamics of Biomacromolecules. Chemical Reviews. 2004; 104:3623–3640. doi:10.1021/cr030413t. [PubMed: 15303831]
17. Palmer AG, Massi F. Characterization of the Dynamics of Biomacromolecules Using Rotating-Frame Spin Relaxation NMR Spectroscopy. Chemical Reviews. 2006; 106:1700–1719. doi:10.1021/cr0404287. [PubMed: 16683750]
18. Clore GM, Szabo A, Bax A, Kay LE, Driscoll PC, et al. Deviations from the simple two-parameter model-free approach to the interpretation of nitrogen-15 nuclear magnetic relaxation of proteins. Journal of the American Chemical Society. 1990; 112:4989–4991.

19. Chen K, Tjandra N. Extended Model Free Approach To Analyze Correlation Functions of Multidomain Proteins in the Presence of Motional Coupling. Journal of the American Chemical Society. 2008; 130:12745–12751. doi:10.1021/ja803557t. [PubMed: 18761455]
20. Chang S-L, Tjandra N. Analysis of NMR Relaxation Data of Biomolecules with Slow Domain Motions Using Wobble-in-a-Cone Approximation. Journal of the American Chemical Society. 2001; 123:11484–11485. doi:10.1021/ja016862x. [PubMed: 11707129]
21. Chang S-L, Szabo A, Tjandra N. Temperature Dependence of Domain Motions of Calmodulin Probed by NMR Relaxation at Multiple Fields. J Am Chem Soc. 2003; 125:11379–11384. doi: 10.1021/ja034064w. [PubMed: 16220961]
22. Walsh JD, Meier K, Ishima R, Gronenborn AM. NMR Studies on Domain Diffusion and Alignment in Modular GB1 Repeats. Biophysical Journal. 2010; 99:2636–2646. [PubMed: 20959105]
23. Tossavainen H, Koskela O, Jiang P, Ylänne J, Campbell ID, et al. Model of a Six Immunoglobulin-Like Domain Fragment of Filamin A (16–21) Built Using Residual Dipolar Couplings. J Am Chem Soc. 2012; 134:6660–6672. doi:10.1021/ja2114882. [PubMed: 22452512]
24. Ermak DL, McCammon JA. Brownian dynamics with hydrodynamic interactions. Journal of Chemical Physics. 1978; 69:1352–1360.
25. Zhou H-X. Brownian dynamics study of the influences of electrostatic interaction and diffusion on protein-protein association kinetics. Biophysical Journal. 1993; 64:1711–1726. [PubMed: 8396447]
26. Schmidt CQ, Herbert AP, Mertens HD, Guariento M, Soares DC, et al. The Central Portion of Factor H (Modules 10-15) Is Compact and Contains a Structurally Deviant CCP Module. Journal of Molecular Biology. 2010; 395:105–122. [PubMed: 19835885]
27. Wallach D. Effect of Internal Rotation on Angular Correlation Functions. J Chem Phys. 1967; 47:5258. doi:10.1063/1.1701790.
28. Turner, PJ.; Grace Developement Team (n.d.). Grace Plotting Tool. version 5.1.22. Available: http://plasma-gate.weizmann.ac.il/Grace/
29. Bevington, PR.; Robinson, DK. Data reduction and error analysis for the physical sciences. McGraw-Hill; 2003. p. 344
30. Woessner DE. Nuclear Spin Relaxation in Ellipsoids Undergoing Rotational Brownian Motion. The Journal of Chemical Physics. 1962; 37:647–654.
31. Jensen SA, Robertson IB, Handford PA. Dissecting the Fibrillin Microfibril: Structural Insights into Organization and Function. Structure. 2012; 20:215–225. doi:10.1016/j.str.2011.12.008. [PubMed: 22325771]
32. Pickford AR, Campbell ID. NMR Studies of Modular Protein Structures and Their Interactions. Chemical Reviews. 2004; 104:3557–3566. doi:10.1021/cr0304018. [PubMed: 15303827]
33. Chen W, Lou J, Hsin J, Schulten K, Harvey SC, et al. Molecular Dynamics Simulations of Forced Unbending of Integrin αVβ3. PLoS Computational Biology. 2011; 7:e1001086. doi:10.1371/journal.pcbi.1001086. [PubMed: 21379327]
34. Tskhovrebova L, Walker ML, Grossmann JG, Khan GN, Baron A, et al. Shape and Flexibility in the Titin 11-Domain Super-Repeat. Journal of Molecular Biology. 2010; 397:1092–1105. doi: 10.1016/j.jmb.2010.01.073. [PubMed: 20138893]
35. Jeffries CM, Lu Y, Hynson RMG, Taylor JE, Ballesteros M, et al. Human Cardiac Myosin Binding Protein C: Structural Flexibility within an Extended Modular Architecture. Journal of Molecular Biology. 2011; 414:735–748. doi:10.1016/j.jmb.2011.10.029. [PubMed: 22041450]
36. Kast D, Espinoza-Fonseca LM, Yi C, Thomas DD. Phosphorylation-induced structural changes in smooth muscle myosin regulatory light chain. Proceedings of the National Academy of Sciences. 2010; 107:8207.
37. Soares, DC.; Barlow, PN. Structural Biology of the Complement System. CRC Press; 2005. Complement Control Protein Modules in the Regulators of Complement Activation; p. 19-62.Available: http://dx.doi.org/10.1201/9780849350368.ch2
38. Barlow PN, Stainkasserer A, Norman DG, Kieffer B, Wiles AP, et al. Solution Structure of a Pair of Complement Modules by Nuclear Magnetic Resonance. Journal of Molecular Biology. 1993; 232:268–284. [PubMed: 8331663]

Author Summary

As experimental methods improve the level of detail to which the structural and dynamic features of proteins can be elucidated, the ability to interpret these experimental studies and extract all of the information they offer often reaches its limits. These analytical limitations are common in studies of the interdomain dynamics of proteins. In particular, it is very challenging to determine the correlation of the motions of the domains in multi-domain proteins from various types of experimental data. In this study, we use Brownian dynamics simulations to study the feasibility and limitations of describing the motional behavior of three-domain proteins in which the domains are linearly connected. We find that their motion is independent when the domains interact via weak potentials, and inter-dependent in the cases when these potentials are strong. We also demonstrate that dynamics in three-domains systems can be completely described by model-free formalism, the leading framework for describing protein dynamics.

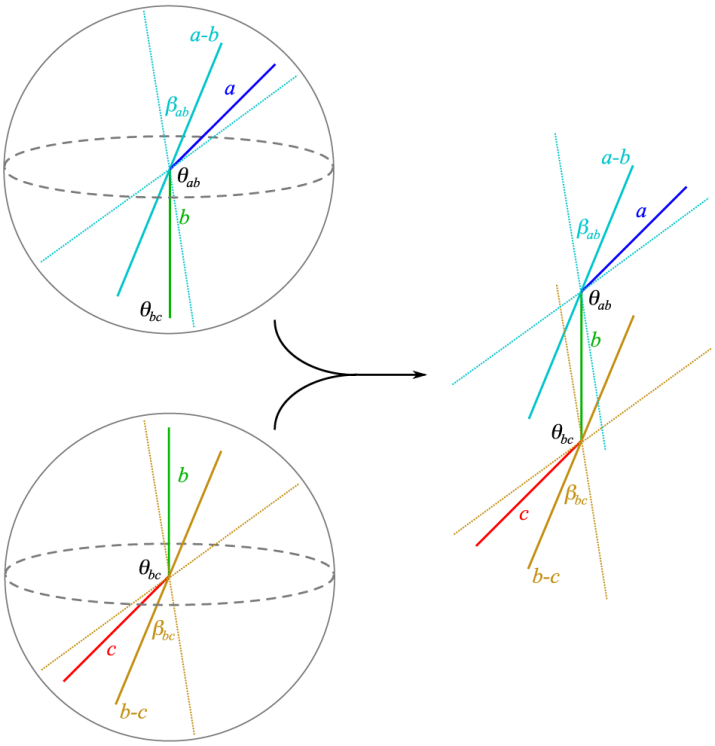

Figure 1.

Hard cone potential in a three-domain system. If at any step of the simulation the interdomain angle θ_{ab} between domains *a* and *b* falls below 180°-2β_{ab} then all three domains will be reset to their previous positions; the situation is analogous for domains *b* and *c*, their interdomain angle θ_{bc}, and the constraining cone semi-angle β_{bc}.

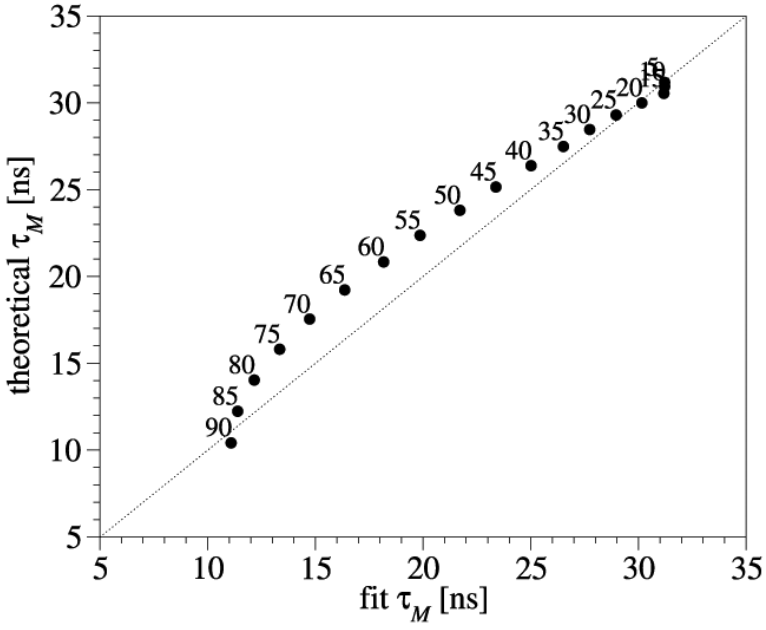

Figure 2.

Correlation times of the overall motion representation. Values from vectors $a-b+c$ from the simulations restrained with symmetrically applied cone potentials are shown. The numbers above the points in the plot show the magnitude of the applied symmetric potential β corresponding to each given point. The three overlapped points near the high end of the plot correspond to β values of 15°, 10°, and 5°.

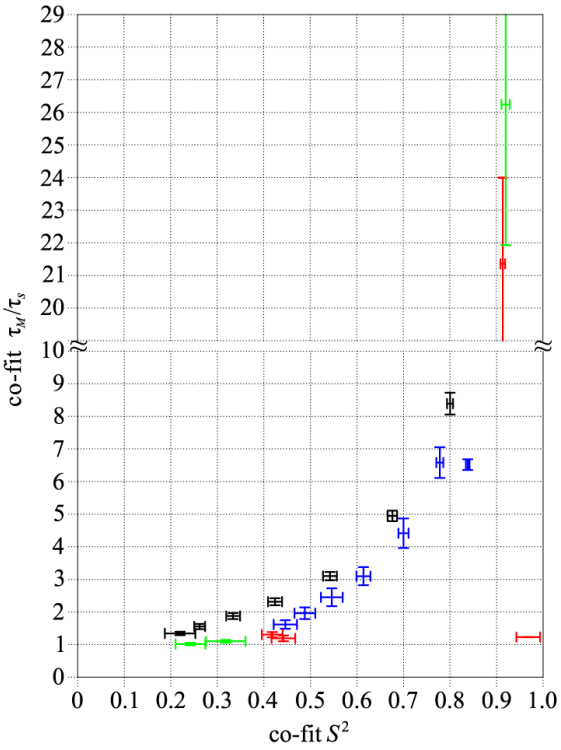

Figure 3.

Values from trajectories of terminal domains are presented in black (those in the ill-defined range in green); values from the central domains in blue (those in the ill-defined range in red).

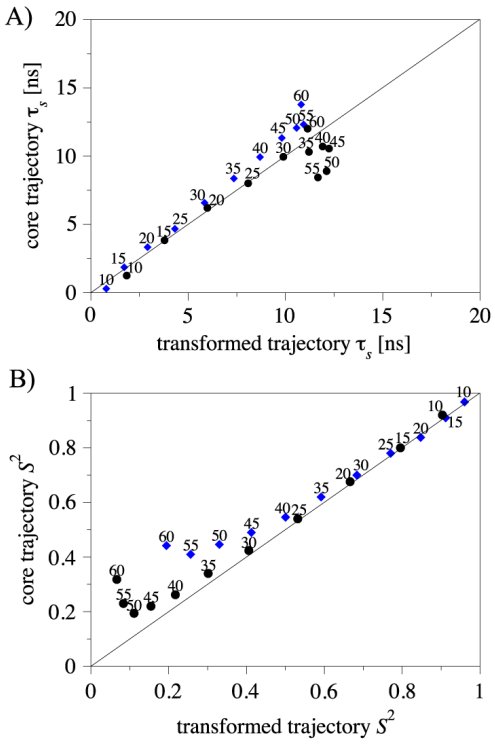

Figure 4.

A) Correlation time in trajectories transformed into the overall motion reference frame vs τ_s in core trajectories. The black circles correspond to fits to correlation functions of trajectories of domain *a* (equivalently *c*), and blue diamonds are from fits to correlation functions of trajectories of domain *b*. The numbers above each point reflect semi-angle β in the cone potential that was symmetrically applied to both interdomain hinges during the corresponding simulation. B) S^2 in trajectories transformed into the overall motion reference frame vs. S^2 in core trajectories. The black circles correspond to fits to correlation functions of trajectories of domain *a* (equivalently *c*), and blue diamonds correspond to fits to correlation functions of trajectories of domain *b*. The numbers above each point reflect semi-angle β in the cone potential applied to both interdomain hinges during the corresponding simulation.