J|A|C|S
ARTICLES

Published on Web 11/18/2008

Conformation and Dynamics of a Rhodamine Probe Attached at Two Sites on a Protein: Implications for Molecular Structure Determination *in situ*

Alfonso De Simone,*,† John E. T. Corrie,‡ Robert E. Dale,§ Malcolm Irving,§ and Franca Fraternali*,§

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, United Kingdom, MRC National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, United Kingdom, and Randall Division of Cell and Molecular Biophysics, King's College London, London SE1 1UL, United Kingdom

Received September 13, 2008; E-mail: franca.fraternali@kcl.ac.uk; ad491@cam.ac.uk

Abstract: Replica exchange molecular dynamics (REMD) calculations were used to determine the conformation and dynamics of bifunctional rhodamine probes attached to pairs of cysteines in three model systems: (a) a polyalanine helix, (b) the isolated C helix (residues 53−66) of troponin C, and (c) the C helix of the N-terminal region (residues 1−90) of troponin C (sNTnC). In each case, and for both diastereoisomers of each probe−protein complex, the hydrophobic face of the probe is close to the protein surface, and its carboxylate group is highly solvated. The visible-range fluorescence dipole of the probe is approximately parallel to the line joining the two cysteine residues, as assumed in previous *in situ* fluorescence polarization studies. The independent rotational motion of the probe with respect to the protein on the nanosecond time scale is highly restricted, in agreement with data from fluorescence polarization and NMR relaxation studies. The detailed interaction of the probe with the protein surface depends on steric factors, electrostatic and hydrophobic interactions, hydrogen bonds, and hydration effects. The interaction is markedly different between diastereoisomers, and multiple preferred conformations exist for a single diasteroisomer. These results show that the combination of the hydrophobic xanthylium moiety of bifunctional rhodamine with the carboxylate substitution in its pendant phenyl ring causes the probe to be immobilized on the protein surface, while the two-site cysteine attachment defines the orientation of its fluorescence dipole. These features allow the orientation of protein components to be accurately determined *in situ* by polarized fluorescence measurements from bifunctional rhodamine probes.

Introduction

Extrinsic probes, often small organic fluorophores covalently bound to specific sites on a protein or other biopolymer, underpin a wide range of methods in biology. Despite the importance of such probes, their use is rarely accompanied by detailed structural information on their interaction with the labeled biological macromolecule, although such information can significantly contribute to understanding the molecular basis of probe signals (see ref 1 for an example). The present work explores the contribution that molecular dynamics calculations can make in determining the orientation and dynamics of a bifunctional rhodamine (BR) probe covalently attached at two sites on isolated model α-helices and on the N-terminal lobe of troponin C from chicken skeletal muscle (sNTnC).

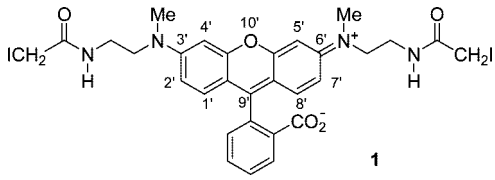
Our interest in the problem developed from studies to measure orientational changes of protein domains within functional protein complexes in their native environment. Although protein crystallography can provide detailed atomic resolution of large supramolecular complexes,[2] it cannot reveal the dynamic conformational changes associated with cellular functioning of proteins *in situ*. Single-molecule measurements can supply mechanical,[3] positional[4] and orientational[5] data in reconstituted protein assemblies (particularly for motor proteins), although not yet in their full cellular environment. In parallel with these single-molecule techniques, we have developed a complementary method based on measurements of fluorescence polarization from BR probes **1** and **2** cross-linking cysteine residues in a target domain. This method has been used to determine the orientation and movement of helices and protein domains *in situ*. Specifically, it has been applied to measurement of the orientation change of myosin domains during force generation in single muscle fibers,[6,7] and to investigate Ca^{2+}-dependent structural changes of troponin C that are involved in muscle regulation.[8,9]

† University of Cambridge.
‡ National Institute for Medical Research.
§ King's College London.

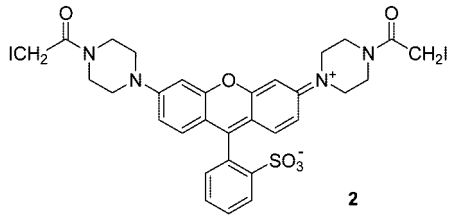
(1) Hirshberg, M.; Henrick, K.; Lloyd Haire, L.; Vasisht, N.; Brune, M.; Corrie, J. E. T.; Webb, M. R. *Biochemistry* **1998**, *37*, 10381–10385.
(2) Dutta, S.; Berman, H. M. *Structure* **2005**, *13*, 381–388.
(3) Veigel, C.; Coluccio, L. M.; Jontes, J. D.; Sparrow, J. C.; Milligan, R. A.; Molloy, J. E. *Nature* **1999**, *398*, 530–533.
(4) Yildiz, A.; Forkey, J. N.; McKinney, S. A.; Ha, T.; Goldman, Y. E.; Selvin, P. R. *Science* **2003**, *300*, 2061–2065.
(5) Forkey, J. N.; Quinlan, M. E.; Shaw, M. A.; Corrie, J. E. T.; Goldman, Y. E. *Nature* **2003**, *422*, 399–404.

10.1021/ja807264v CCC: $40.75

Chart 1. Structures of Bifunctional Rhodamine Reagents



1



2

These studies used muscle proteins (myosin regulatory light chain and troponin C) that were mutated to insert two cysteine residues, separated by 10−17 Å, at selected locations in the protein structure. These residues were cross-linked by the bis-iodoacetamidorhodamine BR-I_2 **1**. The two-site attachment was designed to constrain the orientation of the probe relative to the protein backbone. Initial assumptions that BR-labeling did not significantly perturb the structure or function of the labeled proteins were based on the absence of significant physiological perturbation.[6,8] NMR structures of N-lobe troponin C labeled either with **1** or the related but more rigid bifunctional sulforhodamine **2** (BSR-I_2) (Chart 1) confirmed that neither probe causes detectable structural perturbation of the protein.[10,11] In both cases, the NMR data showed evidence for the presence of two diastereoisomers, as expected and previously discussed.[6] However, the NMR studies did not reveal the structural relationship of either probe to the underlying protein, because NOE contacts between the probes and the protein could not be identified. Likely reasons for this are discussed below. Thus the key assumption underpinning interpretation of the fluorescence polarization data, namely that the visible-range absorption and emission dipoles of the rhodamine probes are approximately parallel to the line joining the β-carbons of the two cysteine residues to which they were attached, could not so far be tested experimentally, although the fact that BR and BSR probes report the same orientation for the C helix of troponin C in muscle fibers provided some indirect support for this assumption.[11]

Data from the experiments in muscle fibers[6–8] also showed that the protein-attached BR and BSR probes are largely immobilized on a time scale shorter than the ∼4 ns fluorescence lifetime.[12] The independent wobble of either probe with respect to the protein has an amplitude corresponding to restricted random motion in a cone with 20−30° half angle.[6,8,12] In accord with these results, recent NMR relaxation measurements in solution showed that BR bound to the C helix of N-lobe troponin C is also highly ordered with respect to the protein.[13] Overall, the data suggest that the xanthylium core of BR and BSR, the region which is responsible for their fluorescence, forms a relatively tight noncovalent interaction with the protein.

Here we used first-principle simulations to obtain an atomic-level description of the conformation and dynamics of a protein-attached BR probe. Others have used related methods to study the orientation and/or dynamics of monofunctional dyes attached to a protein or duplex DNA.[14,15] We calculated the conformation and motion of a BR probe in three systems of increasing complexity: (a) a generic helix of alanine residues that included two cysteine residues with i, $i + 7$ spacing; (b) an isolated helix of the same sequence as the C helix of sNTnC, mutated to place cysteines at positions 56 and 63 (whole sNTnC numbering); (c) the whole of sNTnC, with the same cysteine mutations as in (b). In case (c), the protein was complexed with Ca^{2+} and the switch region of troponin I (TnI; residues 115−131), which was included in order to compare with the previously determined NMR structure.[10]

In each of these systems we applied replica exchange molecular dynamics (REMD) simulations,[16] a powerful technique for sampling the conformational space of proteins and peptides.[17−21] We calculated conformational free energy surfaces that describe the preferred conformations of the probe with respect to the protein or peptide. Further MD simulations starting from each preferred conformation were performed to characterize probe dynamics for comparison with fluorescence polarization data.

Methods

Simulated Systems. The starting coordinates of the protein, the BR-labeled N-domain of troponin C from chicken skeletal muscle (residues 1−90, with the mutations E56C, E63C), complexed with Ca^{2+} and the regulatory "switch" peptide (residues 115−131) from troponin I (this complex is referred to hereafter simply as BR-sNTnC), were taken from the NMR structure of the same complex (PDB code: 1NPQ).[10] BR coordinates and charge parameters were based on those previously derived for *N*,*N*,*N*′,*N*′-tetramethylrhodamine by Density Functional Theory analysis.[22] A program developed in-house was used to generate the starting conformations of BR attached to peptide or protein, with the linker arms in an extended conformation. The three systems studied were the BR-labeled 13-mer AACAAAAAACAAA (referred to as BR-$A_{11}C_2$), the BR-labeled double-cysteine mutant C helix of TnC (TKCELDAIICEVD; referred to as BR-H_C) and the whole complex, BR-

(6) Corrie, J. E. T.; Brandmeier, B. D.; Ferguson, R. E.; Trentham, D. R.; Kendrick-Jones, J.; Hopkins, S. C.; van der Heide, U. A.; Goldman, Y. E.; Sabido-David, C.; Dale, R. E.; Criddle, S.; Irving, M. *Nature* **1999**, *400*, 425–430.
(7) Hopkins, S. C.; Sabido-David, C.; van der Heide, U. A.; Ferguson, R. E.; Brandmeier, B. D.; Dale, R. E.; Kendrick-Jones, J.; Corrie, J. E. T.; Trentham, D. R.; Irving, M.; Goldman, Y. E. *J. Mol. Biol.* **2002**, *318*, 1275–1291.
(8) Ferguson, R. E.; Sun, Y. B.; Mercier, P.; Brack, A. S.; Sykes, B. D.; Corrie, J. E. T.; Trentham, D. R.; Irving, M. *Mol. Cell* **2003**, *11*, 865–874.
(9) Sun, Y. B.; Brandmeier, B.; Irving, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 17771–17776.
(10) Mercier, P.; Ferguson, R. E.; Irving, M.; Corrie, J. E. T.; Trentham, D. R.; Sykes, B. D. *Biochemistry* **2003**, *42*, 4333–4348.
(11) Julien, O.; Sun, Y. B.; Knowles, A. C.; Brandmeier, B. D.; Dale, R. E.; Trentham, D. R.; Corrie, J. E. T.; Sykes, B. D.; Irving, M. *Biophys. J.* **2007**, *93*, 1008–1020.
(12) Dale, R. E.; Hopkins, S. C.; van der Heide, U. A.; Marszałek, T.; Irving, M.; Goldman, Y. E. *Biophys. J.* **1999**, *76*, 1606–1618.
(13) Julien, O.; Mercier, P.; Spyracopoulos, L.; Corrie, J. E. T.; Sykes, B. D. *J. Am. Chem. Soc.* **2008**, *130*, 2602–2609.
(14) Schröder, G. F.; Alexiev, U.; Grubmüller, H. *Biophys. J.* **2005**, *89*, 3757–3770.
(15) Ivanova, A.; Jezierski, G.; Vladimirov, E.; Rösch, N. *Biomacromolecules* **2007**, *8*, 3429–3438.
(16) Sugita, Y.; Okamoto, Y. *Chem. Phys. Lett.* **1999**, *314*, 141–151.
(17) Periole, X.; Mark, A. E. *J. Chem. Phys.* **2007**, *126*, 014903.
(18) Baumketner, A.; Shea, J. E. *J. Mol. Biol.* **2007**, *366*, 275–285.
(19) Cecchini, M.; Rao, F.; Seeber, M.; Caflisch, A. *J. Chem. Phys.* **2004**, *121*, 10748–10756.
(20) Gnanakaran, S.; Garcia, A. E. *Proteins* **2005**, *59*, 773–782.
(21) Sgourakis, N. G.; Yan, Y. L.; McCallum, S. A.; Wang, C. Y.; Garcia, A. E. *J. Mol. Biol.* **2007**, *368*, 1448–1457.
(22) Cavallo, L.; Moore, M. H.; Corrie, J. E. T.; Fraternali, F. *J. Phys. Chem. A* **2004**, *108*, 7744–7751.

sNTnC. The $A_{11}C_2$ and H_C helices were modeled on the backbone coordinates of BR-sNTnC. Each system was simulated twice to accommodate the two diasteroisomeric forms of BR bound to the chiral helices (see Results and Figure S1 of Supporting Information).

All systems were immersed in a box filled with extended single point charge (SPCE) waters.[23] The simulations were performed with the GROMACS[24] package, using the GROMOS96[25] force field with an integration time step of 2 fs. Nonbonded interactions were treated by the Particle Mesh Ewald method (grid spacing 1.2 Å)[26] for the electrostatic contribution, with 14 Å cutoff distance for van der Waals contributions. Neutral pH was set by selecting the appropriate protonation states of pH sensitive side-chains. BR-sNTnC was simulated by constraining the positions of protein C^α atoms (using a harmonic potential of 1000 $kJ \cdot mol^{-1} \cdot nm^{-2}$), whereas for BR-$H_C$ and BR-$A_{11}C_2$ a potential of 300 $kJ \cdot mol^{-1} \cdot nm^{-2}$ was applied to hydrogen bond donor−acceptor pairs to preserve main chain hydrogen-bond patterns.

REMD Setup. The REMD methodology simulates several copies (replicas) of the same system evolving independently at different temperatures.[16] Exchanges between neighboring replicas were attempted at time intervals, t_{swap}, of 0.5 ps on the basis of the Metropolis criterion[27] (eq 1):

$$P(1 \rightarrow 2) = \min\left(1, \exp\left[\left(\frac{1}{k_B T_1} - \frac{1}{k_B T_2}\right)(U_1 - U_2)\right]\right) \quad (1)$$

where $P(1 \rightarrow 2)$ is the exchange probability, k_B is Boltzmann's constant, and U_1 and U_2 are the potential energies of individual configurations of the replicas at the temperatures, T_1 and T_2, of the two relevant thermal baths. Since eq. 1 applies for the NVT ensemble, an additional term for the isobaric−isothermal (NPT) ensemble, as adopted here, should in principle be introduced.[28] However, this term is negligible when the volumetric fluctuations are small.[29] All runs employed 12 replicas with the following temperatures (K): 299.6, 303.0, 306.4, 309.8, 313.2, 316.8, 320.4, 324.2, 328.0, 332.0, 336.0 and 340.2 (see Supporting Information, Figure S2A). To ensure homogeneous exchange frequencies, the temperature pattern has an increment ranging from 3.4 K ($\Delta T_{1,2}$) to 4.2 K ($\Delta T_{11,12}$). This specific increasing spacing is adopted to balance the reduced exchange frequencies at low temperatures, which result from the sharpening of the potential energy distributions. For each temperature bath, the initial conformation of each system was equilibrated for 1 ns before exchanges between baths were initiated. All simulations were carried out for 10 ns per replica and the total amount of sampling performed for the set of BR-labeled systems was 0.72 μs, i.e. 120 ns per system. On account of the limited number of degrees of freedom of the linker arms between BR and the helices, these simulations are robust in describing the conformational preferences of the probe, as also indicated by the convergence of results for different systems.

All runs experienced a significant extent of random walks (Supporting Information, Figure S2C), because of extensive overlap of the potential energy distributions (not shown). This was consistent with the high exchange frequencies (~13%) (Supporting Information, Figure S2B), so each replica explored the whole temperature space by passing repeatedly throughout all thermal baths (Supporting Information, Figure S2C).

Hydration and Hydrophobicity Analysis. The hydration analysis was based on the solvent density map, the maxima of which represent hydration sites. This map represents the water population density around a molecule calculated during the course of MD sampling.[30−32] The space surrounding the protein was partitioned by considering an inner shell describing the solute hydration (extending up to 6 Å from the solute surface), and an outer shell representing the bulk solvent (extending 6−8 Å from the solute surface). The solvent density calculation is grid-based (step size 0.5 Å). To eliminate solute translation and rotation, the coordinates of each frame were transformed by superimposing the current model onto a reference one. The QUILT program[33] was used for analysis of hydrophobic clusters on the surface of the protein. This program uses a very fast algorithm to calculate hydrophobic patches, defined as clusters of neighboring apolar atoms deemed accessible on a given protein surface. Water density maps around the rhodamine were calculated using a previously described method.[30]

Coordinates for Free Energy Maps. Analysis of the REMD simulations was carried out by 2D free energy projections along two specific "reaction" coordinates that were selected to represent the position and orientation of the probe with respect to the protein. d is a measure of the separation of the probe from the surface of the C helix, calculated as the distance between the geometrical center of the xanthylium system (defined as the midpoint of the central, oxacyclic ring) and the nearest backbone atom of the helix between the labeled cysteines. θ is the orientation parameter, calculated as the angle between the line that joins atoms 3′ and 6′ of the xanthylium system (see structure **1**) and the line joining the C^β atoms of the labeled cysteines. An additional positional parameter, z, was useful to describe sliding of the probe along the helix, and was calculated by projecting the geometrical center of the xanthylium system onto the axis joining the C^β atoms of the labeled cysteines. $z = 0$ denotes the midpoint between the cysteines, and positive z denotes displacement toward Cys63.

The free-energy projections described above identified closely related sets of preferred probe conformations with respect to the helix or protein. The set of conformations with free energy within 3 kJ/mol of the local minimum is referred to as a basin. For each basin we calculated a representative structure using a clustering algorithm to analyze the rmsd matrix of 1000 conformers (out of the total of 20000 saved at 0.5 ps intervals from the 300 K bath of the REMD sampling) chosen with similar values of d and θ (typically differing by 0.5 Å and 5° respectively). These representative structures are shown in Figure 1, and their d, θ and z parameters are listed in Table 1.

Probe Dynamics. An independent MD simulation (300 K and 1 atm) of 10 ns was performed for each basin identified in the REMD samplings of the BR-sNTnC system, starting from its representative structure. The orientation parameter θ was calculated for each conformer. Orthogonal components θ_x and θ_y were calculated in a coordinate frame in which the z axis joins the C^β atoms of the labeled cysteines and the y axis is orthogonal to the plane that contains the z axis and the line joining the C^α and C^β atoms of Cys56. The variation of θ on the nanosecond time scale of the simulation describes the independent motion with respect to the protein backbone of the line joining atoms 3′ and 6′ of the xanthylium system, to which the absorption and emission dipoles

(23) Berendsen, H. J. C.; Grigera, J. R.; Straatsma, T. P. *J. Phys. Chem.* **1987**, *91*, 6269–6271.
(24) Berendsen, H. J. C.; van der Spoel, D.; van Drunen, R. *Comput. Phys. Commun.* **1995**, *91*, 43–56.
(25) Van Gunsteren, W. F.; Billeter, S.; Eising, A.; Hunenberger, P. H.; Kruger, P.; Mark, A. E.; Scott, W. R. P.; Tironi, I. G. *Biomolecular Simulations: The GROMOS96 Manual and User Guide*; Verlag der Fachvereine: Zürich, 1996.
(26) Darden, T.; Perera, L.; Li, L.; Pedersen, L. *Structure Fold. Des.* **1999**, *7*, R55–R60.
(27) Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. J. *J. Chem. Phys.* **1953**, *21*, 1087–1092.
(28) Okabe, T.; Kawata, M.; Okamoto, Y.; Mikami, M. *Chem. Phys. Lett.* **2001**, *335*, 435–439.
(29) Seibert, M. M.; Patriksson, A.; Hess, B.; van der Spoel, D. *J. Mol. Biol.* **2005**, *354*, 173–183.
(30) De Simone, A.; Dodson, G. G.; Verma, C. S.; Zagari, A.; Fraternali, F. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 7535–7540.
(31) De Simone, A.; Spadaccini, R.; Temussi, P. A.; Fraternali, F. *Biophys. J.* **2006**, *90*, 3052–3061.
(32) De Simone, A.; Dodson, G. G.; Fraternali, F.; Zagari, A. *FEBS Lett.* **2006**, *580*, 2488–2494.
(33) Lijnzaad, P.; Berendsen, H. J.; Argos, P. *Proteins* **1996**, *25*, 389–397.

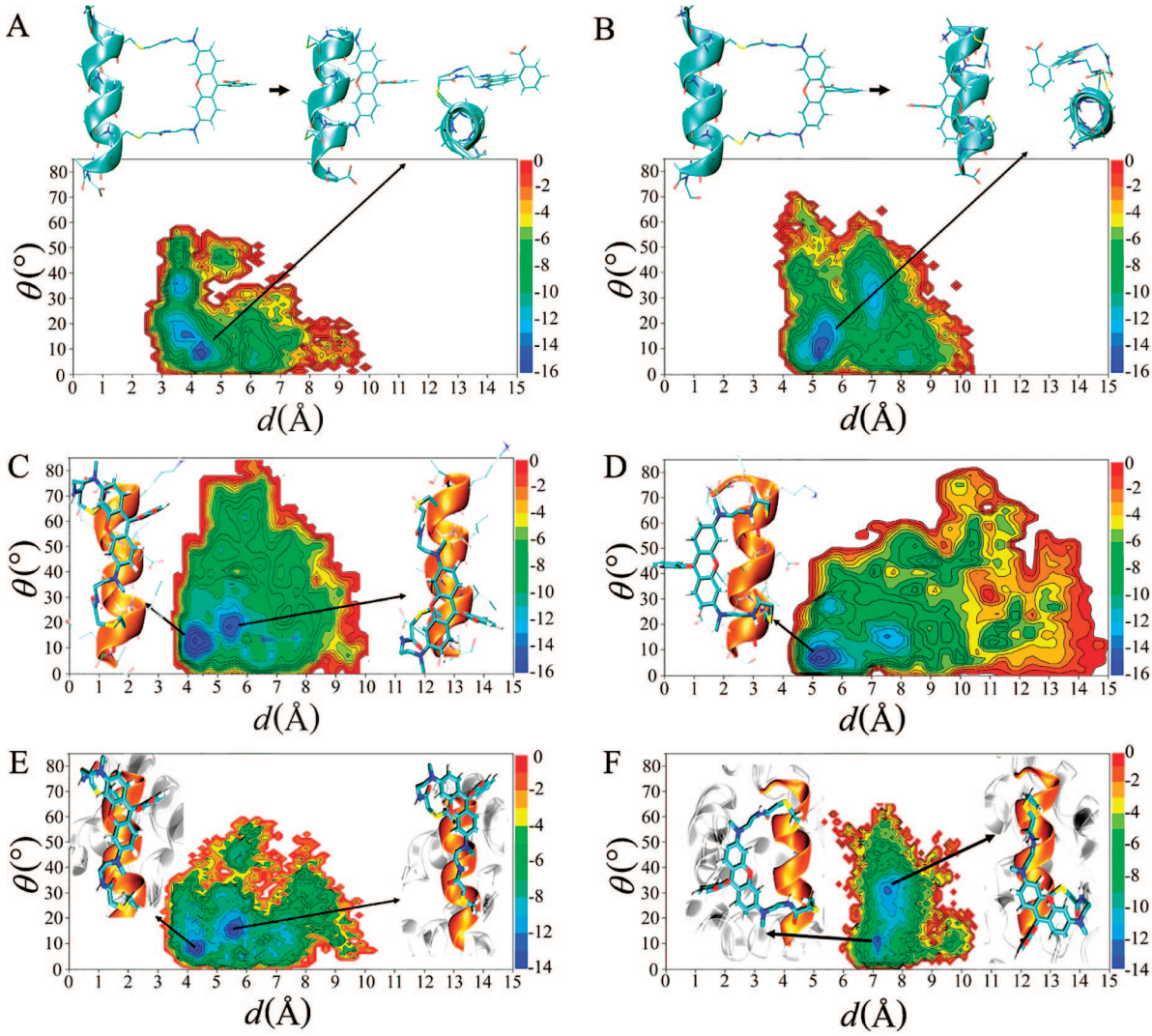


Figure 1. Free energy surfaces and structural models of the three systems studied. In all cases, the helices are oriented with their N-terminus at the top of the panel and their C-terminus at the bottom. The left- and right-hand panels represent the *syn*- and *anti*-isomers respectively. Energies are calculated in kJ/mol. (A) and (B) BR-$A_{11}C_2$ system: In this case, the structures show the starting conformation on the left-hand side for each isomer, and the folded structures to the right of the arrows are representative structures for the conformations in the basin. The end-on views show the representative structures rotated by 90° (with the C-terminus toward the front) to show the packing of the xanthylium system and the geometries of the diastereoisomers. In the direction of increasing residue number and viewed from the C-terminus, the helices have a counter-clockwise rotational sense. In (A) the vector of the bond from the pendant aryl ring to its attached carboxylate has the same sense as the helix, while in (B) the same vector is in the opposite direction to the rotational sense of the helix. (C) and (D) BR-H_C system: For the *syn*-isomer, representative structures are shown for both basins identified by the REMD simulations. (E) and (F) BR-sNTnC system: Representative structures are shown for each basin.

of the fluorophore are expected to be parallel.[34] The amplitude of this motion on the nanosecond time scale was described by the order parameter $\langle P_{2d} \rangle$ according to the formula

$$\langle P_{2d} \rangle = \frac{3}{2}\langle \cos^2 \Delta\theta \rangle - \frac{1}{2} \qquad (2)$$

where the angle brackets denote an average over the conformers in the simulation and $\Delta\theta$ denotes the angle between the line joining atoms 3′ and 6′ of the xanthylium system for an individual conformer and the corresponding line for the average of all conformers in the basin.[12]

Results

We compared three systems in order to elucidate the molecular determinants of the conformation of the BR probe with respect to the protein backbone. As defined in Methods, these were: (a) BR-$A_{11}C_2$; a polyalanine helix including two cysteines with i, $i + 7$ spacing; (b) BR-H_C; the isolated C helix of sNTnC mutated to place cysteines at positions 56 and 63 (whole TnC numbering); (c) BR-sNTnC; the whole of sNTnC with the same cysteine mutations. The comparison between (a) and (b) was expected to reveal the effects of local sequence on probe conformation, while the comparison between (b) and (c) should show effects from the whole protein structure. Each system has two diastereoisomers[6,10] which were simulated separately. The two isomers are termed *syn*- and *anti*-, with reference to the direction of the bond from the pendant aryl ring of the BR dye to its carboxylate substituent. For the *syn*-isomer (Figures 1A,C,E) this points in the same direction as the turn of the helix toward increasing residue number (toward the viewer in the end-on view at the top right of Figure 1A). For the *anti*-isomer, the ring-to-carboxylate bond points in the opposite direction to the turn of the helix (Figures 1B,D,F). The

(34) Penzkofer, A.; Wiedmann, J. *Opt. Commun.* **1980**, *35*, 81–86.

Table 1. Parameters for the Reference Structures of Each Basin

isomer	basin	*d* (Å)	*θ* (°)	*z* (Å)
		BR-$A_{11}C_2$		
syn	1	4.4	8.4	−1.3
anti	1	5.2	7.6	0.1
		BR-H_C		
syn	1	4.3	13.9	−2.6
syn	2	5.5	18.6	5.2
anti	1	5.2	6.8	0.3
		BR-sNTnC		
syn	1	4.4	10.1	−3.1
syn	2	5.5	16.2	−5.5
anti	1	7.2	10.3	−1.3
anti	2	7.5	32.7	6.0

starting conformations of the diastereoisomers in each system are illustrated in greater detail in Figure S1 (Supporting Information).

Conformation of BR on a Polyalanine Helix. The simplest system was the isolated polyalanine helix containing two cysteines cross-linked with BR (BR-$A_{11}C_2$; Figures 1A,B). The cysteines are seven residues apart, so their β-carbon atoms are on the same side of the helix and about 11 Å apart. Each simulation started with the BR linker arms in an extended conformation (top left molecular graphics in Figures 1A and 1B), and folding of the probe onto the helix surface occurred within about 1 ns.

The contour plots of free energy against probe orientation θ and probe-helix separation d contain conformational basins (blue regions) of low free energy. A basin was found to correspond to a closely spaced set of conformers, as detailed below. The preferred conformation of BR-$A_{11}C_2$, *syn*-isomer (Figure 1A) has a probe-helix separation parameter d close to 4 Å. The long axis of the probe is roughly parallel to the line joining the cysteines to which it is attached, shown by the orientation parameter θ of ca. 10° (see Methods for definitions of d and θ). The middle and right-hand views in Figure 1A show different projections of the representative structure of this basin.

The preferred conformation of the *anti*-isomer (Figure 1B) has d close to 5 Å and θ about 10°, as for the *syn*-isomer. Thus for both diastereoisomers the preferred conformation of the rhodamine is close to the helix surface with the xanthylium long axis roughly parallel to the helix axis. The carboxylate group points into the solvent, away from the helix. The MD analysis of hydration sites for the rhodamine moiety shows high water density near the carboxylate group, while the rest of the molecule is quite void of water (Figure S3, Supporting Information). In general the detailed structures of the two diastereoisomers are distinct, as also indicated by the different axial displacement (z) of the probe along the helix (Figures 1A,B and Table 1). These differences reflect the intrinsic asymmetry of an α-helix in which all the side chains are identical except for the labeled cysteines.

Effects of Helix Side Chains on BR Conformation. Figures 1C,D show free energy maps for BR bound to the isolated helix C of TnC (BR-H_C) that are broadly similar to those described above for the polyalanine helix. The preferred probe conformations are again close to the helix surface, with the long axis of the probe almost parallel to the helix in each case. The free energy map for the *syn*-isomer of BR-H_C (Figure 1C) is split into two closely spaced conformational basins with similar occupancy (38% and 24% of the total population for the basins with lower and higher d, respectively), corresponding to

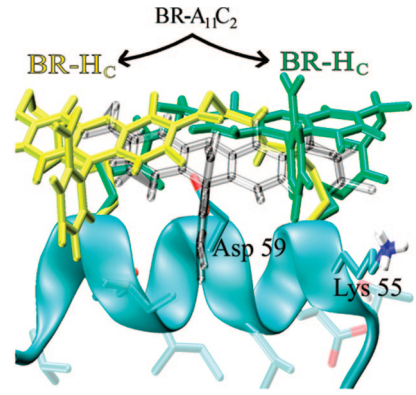


Figure 2. Local sequence effects on BR conformations. The unique basin of BR-$A_{11}C_2$ (BR displayed in white) is split into two in BR-H_C [BR in green (basin 1) and yellow (basin 2)] because of a steric clash with Asp 59. For BR in basin 1, the carboxyphenyl ring is bracketed between Lys55 and Asp59.

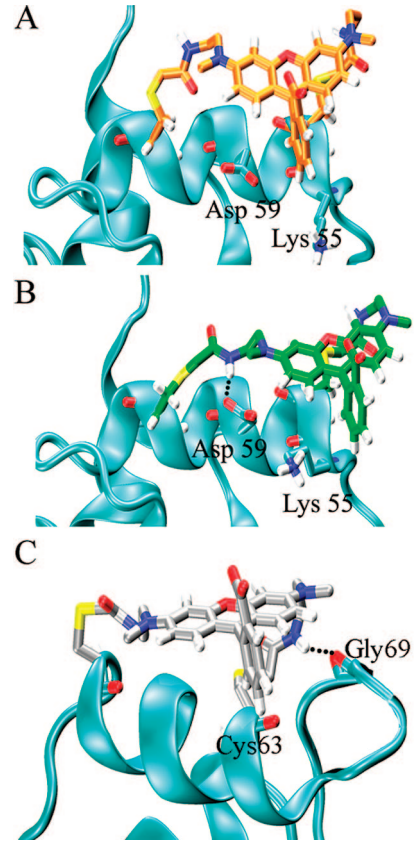


Figure 3. Close-up views of BR-sNTnC conformers. (A) *syn*-isomer, basin 1, showing bracketing of the carboxyphenyl group of BR between Lys55 and Asp59; (B) *syn*-isomer, basin 2, showing the hydrogen bond between Asp59 and the amide attached to Cys63; (C) *anti*-isomer, basin 2, showing the hydrogen bond between Gly69 and the amide attached to Cys63.

markedly different conformations, with the probe at opposite ends of the helix (Figure 1C, molecular graphic insets). The axial displacement parameter z differs by about 8 Å between the two basins (Table 1). These distinct conformations (green and yellow BR in Figure 2 for basins 1 and 2 respectively) arise from a steric clash between the Asp59 side-chain and the carboxyphenyl ring of BR. The effect is not present in *syn*-BR-$A_{11}C_2$ (white BR in Figure 2) because of the smaller size of the corresponding alanine side-chain in that structure. Lys55 also plays a role in defining z for basin 1 of *syn*-BR-H_C, in which

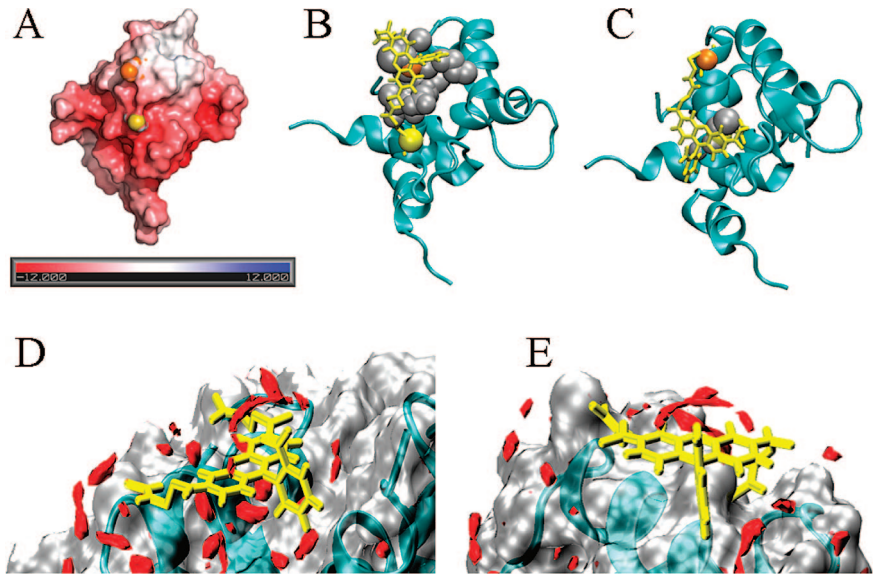


Figure 4. Hydrophobic patches and surface hydration of sNTnC. (A). Electrostatic surface representation of a representative NMR structure (PDB code 1NPQ), with orange and yellow spheres representing Cys56 and Cys63 respectively; (B) and (C): ribbon representations of the representative conformers of the *syn*-isomer, basin 2 and the *anti*-isomer, basin 2 respectively, with gray van der Waals spheres indicating hydrophobic patches determined from QUILT and BR in yellow; (D) and (E): hydration analysis of the *anti*-isomer, basins 1 and 2, respectively, with red regions indicating high water density.

the carboxyphenyl ring is effectively constrained to lie between the Lys55 and Asp59 side chains. In contrast, the *anti*-isomer of BR-H_C shows a single preferred conformation (Figure 1D), with positional and orientation parameters close to those seen for *anti*-BR-$A_{11}C_2$ (Table 1). With this diastereoisomer, the carboxyphenyl ring protrudes more toward the solvent and there is no steric clash with Asp59 or any other side chain.

Overall Effect of the Protein Fold on BR Conformation. The free energy map for the *syn*-isomer of the BR probe attached to sNTnC shows two preferred conformations in (θ, d) space (Figure 1E), representing 35% and 30% of the total number of conformers. The (θ, d) coordinates of these conformations are similar to those seen with *syn*-BR-H_C. Basin 1 (left-hand molecular graphic in Figure 1E), with $d = 4.4$ Å, also represents a similar docking of the probe onto the helix surface; z is -3.1 and -2.6 Å for BR-sNTnC and BR-H_C respectively (Table 1). In both cases the carboxyphenyl ring is bracketed by the Asp59 and Lys55 side chains (Figure 3A). Additional close-up views are shown in Figure S4A (Supporting Information).

Basin 2 of *syn*-BR-sNTnC, however, represents a new conformation, with a z value of -5.5 Å (Table 1), in marked contrast to its value ($+5.3$ Å) for basin 2 of *syn*-BR-H_C. When the rest of the protein is present, the xanthylium moiety takes up an extreme N-terminal position with respect to the C helix, with the carboxyphenyl ring on the N-terminal side of Lys55 (Figure 3B and Figure S4B, Supporting Information). This position seems to be determined by an interaction between the hydrophobic face of BR and a hydrophobic pocket on the surface of the protein. The surface of BR-sNTnC is largely hydrophilic, but has hydrophobic patches (Figure 4A). One of the largest such patches, with 260 Å^2 of solvent-accessible surface area as determined by QUILT[33] (Figure 4B), includes the carbon and sulfur atoms of residues 54 to 60. The basin 2 conformation of *syn*-BR-sNTnC seems to be primarily determined by an interaction with this accessible hydrophobic patch (Figure 4B), although it may be reinforced by a hydrogen bond between the β-carboxylate of Asp59 and the amide NH of the BR linker arm on Cys63 (Figure 3B).

The free energy map for the *anti*-isomer (Figure 1F) also shows two basins, in this case containing 26% and 19% of the total conformers. Both basins have high d values (7.2−7.5 Å), and the accessible range of d is much narrower than that for the isolated H_C system (Figure 1D, and Figures S4C,D, Supporting Information), so the two basins are separated by a difference in θ rather than in d. The basin 1 conformer, with the lower θ, has rather elongated linker arms that extend out from the protein surface, allowing a high water density on the hydrophilic surface of the protein (red contours in Figure 4D), but with no water contact with the hydrophobic face of the probe. The basin 2 conformer, in contrast, has the probe and protein surfaces closely apposed, with little intervening water density (Figure 4E). This conformer shows the largest θ (33°) of any of the systems studied here (Table 1), and the probe is displaced toward the C-terminus of helix C ($z = +6$ Å). It seems to be stabilized by a 90 Å^2 hydrophobic patch that includes the carbon and sulfur atoms of residues 63 and 64 (Figure 4C), and may be further constrained by a hydrogen bond between the backbone carbonyl of Gly69 and the amide NH of the BR linker arm on Cys63 (Figure 3C).

Although the details of the relative probe-protein conformations depend on the system and the diastereoisomer (Figure 1), all the conformations observed here have the carboxylate group of the BR facing into the solvent. An additional simulation was carried out to determine the importance of the carboxylate in determining the overall BR conformation. Starting from the representative conformer of basin 1, *syn*-BR-sNTnC, the carboxylate was manually moved to the opposite side of the pendant phenyl ring, in effect converting it to the *anti*-isomer, before running the simulation. Within a few hundred ps the xanthylium moiety moved away from the C helix surface and the BR took up the conformations described above for the *anti*-isomer, with the carboxylate pointing into the solvent (data not shown).

Probe Motion Within the Conformational Basins on the Nanosecond Time Scale. Independent molecular dynamics simulations carried out at 300 K for each of the four basins of

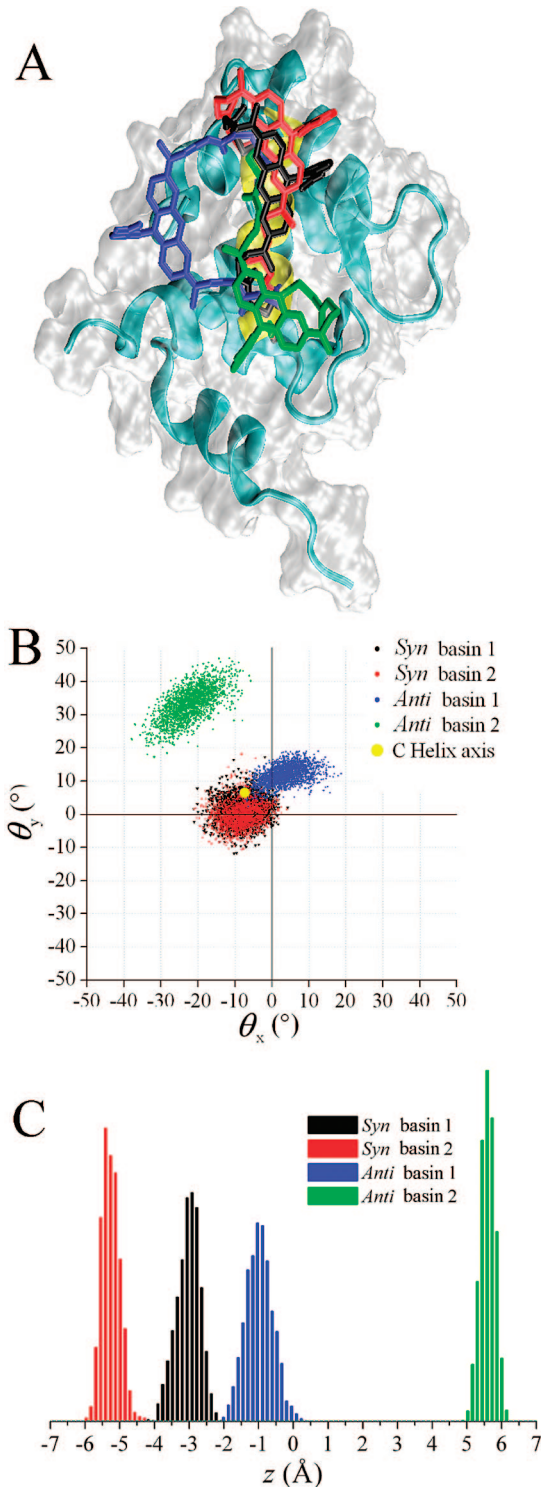


Figure 5. Four conformational basins of sNTnC. (A) Representative structures superimposed using their C^{α} positions, with BR in black, red, blue and green for *syn*-isomer basins 1 and 2 and *anti*-isomer basins 1 and 2, respectively; (B) orientational dispersion within each basin, plotted in terms of the orthogonal angles θ_x and θ_y defined in the text. Color code as in (A), with the C helix axis indicated by a yellow circle; (C) distribution of z displacements along the C helix; color code as in (A).

BR-sNTnC, starting from the four representative conformations (Figure 5A), showed in each case that the probe remained within the basin throughout the 10 ns simulation. The 3D angle between the probe's long axis (which is parallel to its visible-range absorption and emission dipoles[34]) and the line between the C^{β} protein attachment points that define the z axis, was resolved into orthogonal components θ_x and θ_y (Figure 5B), where the y axis is perpendicular to the plane that contains the z axis and the C^{α}-C^{β} vector of Cys56. When $\theta_x = 0$, $\theta_y = 0$ the probe dipole is parallel to the line between the two cysteines, as assumed in published studies in which the BR probes were used to estimate the orientation of helices or protein domains in muscle fibers.[6–9] The present results show that this is a reasonably good approximation, within ca. 10°, for the two basins of *syn*-BR-sNTnC (red, black) and basin 1 of *anti*-BR-sNTnC (blue), but less good for basin 2 of *anti*-BR-sNTnC (green), which is centered near $\theta_x = -20°$, $\theta_y = -30°$. The z distributions (Figure 5C) are narrower for the basins in which the BR is displaced from the midpoint of the helix (red and green). This is consistent with the constraints associated with hydrogen bonds to the Cys63-attached BR amide group, as discussed above.

The orientation distributions of the four basins (Figure 5B) were used to calculate the dispersion parameter $\langle P_{2d} \rangle$ that represents independent motion of the probe with respect to the protein on time scales short compared to the 4 ns fluorescence lifetime. $\langle P_{2d} \rangle$ values were 0.97, 0.99, 0.97 and 0.95 for basins 1 and 2 of *syn*-BR-sNTnC and basins 1 and 2 of *anti*-BR-sNTnC, respectively. $\langle P_{2d} \rangle$ values close to unity correspond to a highly restricted angular range of independent probe motion.

Implications of Probe Orientations for in situ Structural Studies. The BR probe studied here was designed to measure intracellular orientations of protein domains from fluorescence polarization data, under the assumption that the BR dipole axis is parallel to the line joining the β-carbons of the two cysteines to which the probe is attached. The results described above allow this assumption to be removed in the case of BR attached to the C helix of sNTnC, and the polarized fluorescence data to be reinterpreted using the BR orientations determined by REMD.

We consider the orientation of the N-lobe of TnC in actively contracting skeletal muscle cells[8] as an example of the analysis. These cells have effective cylindrical symmetry, and the orientation of the N-lobe of TnC can be described in terms of two orientation parameters: β, the angle between its D helix and the muscle fiber or actin filament axis, and γ, describing rotation around the D helix axis (Figure 6C). $\gamma = 0$ denotes coplanarity of the A and D helix axes with the fiber axis. Data from more than one pair of BR attachment sites are required to derive the (β,γ) orientation distribution, and this was achieved by attaching BR to the N, A and C helices in separate experiments.[8] Assuming that the BR dipole is parallel to each pair of attachment sites, it is then possible to calculate the distribution of β and γ in the muscle cell, shown as a contour plot in Figure 6A, where hotter colors denote a higher probability of a particular orientation. This (β,γ) distribution is the broadest distribution, calculated by maximizing informational entropy,[35] that is consistent with the fluorescence polarization data from the three probes, and can be considered as a low resolution view of the real distribution. The peak of the distribution is close to $\beta = 65°$; $\gamma = 5°$; these values are slightly different from those reported previously ($\beta = 82°$; $\gamma = 10°$)[8] because of a revised definition of the maximum entropy distribution in which the previous $\sin\beta$ weighting term is omitted.[11]

This analysis was repeated using the representative orientations of the BR probe on the C helix calculated in the

(35) van der Heide, U. A.; Hopkins, S. C.; Goldman, Y. E. *Biophys. J.* **2000**, *78*, 2138–2150.

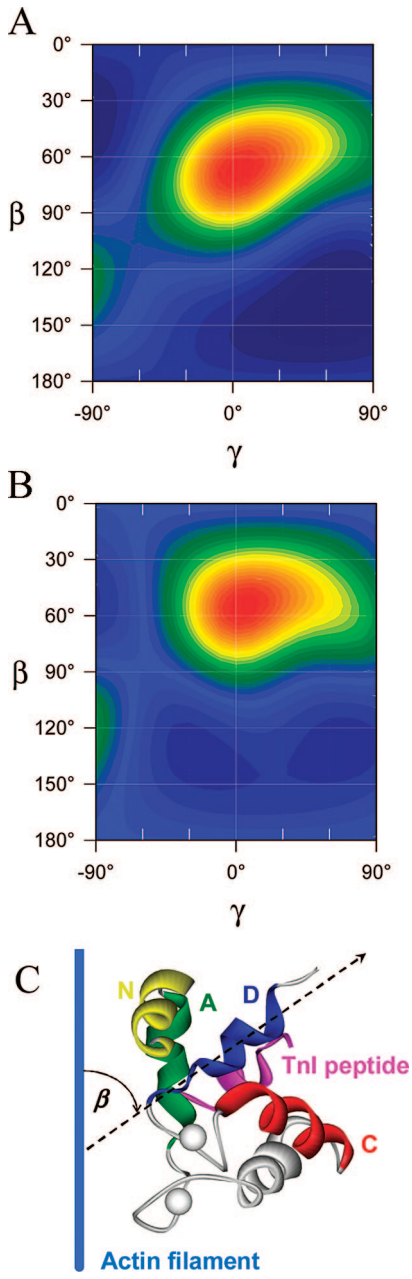


Figure 6. Orientation of sNTnC in an actively contracting skeletal muscle fiber, determined from polarized fluorescence data. (A) and (B) Contour plots of the maximum entropy distributions of angles β and γ as defined in the text. Plot A was calculated under the assumption that the BR dipole on helix C is parallel to the line joining the C^{β} residues to which it is attached; plot B was calculated using the REMD calculations of its orientation (see text for details); (C) graphical representation of the peak orientation in (B).

present study, retaining the previous assumption that the fluorescence dipoles of the N- and A-helix probes are parallel to the lines joining the labeled cysteines. To simulate the conditions of the muscle fiber experiments, we used a 1:1 mixture of *syn*- and *anti*-isomers, with relative occupancies of basins 1 and 2 as determined above by REMD. The resulting (β,γ) distribution (Figure 6B) peaks at $\beta = 55°$; $\gamma = 5°$. This conformation is shown graphically in Figure 6C, and is quite similar to that calculated by assuming that the long axis of the probe is parallel to the cysteine residues to which it is attached. The peaks of the two distributions have the same γ and differ in β by about 10°.

Discussion

The orientations of protein domains in their native environment have been estimated previously by measurements of the fluorescence polarization of BR probes **1** and **2** cross-linking cysteine residues in a target domain.[6–11] This approach has depended on a previously untested assumption about the orientation of the BR probe dipole relative to the protein to which it is attached. Moreover, quantitative interpretation of the fluorescence polarization data requires additional assumptions about the extent of the relative mobility of the probe with respect to the protein backbone on the nanosecond time scale of the fluorescence lifetime. The present results provide the first *ab initio* estimates of both the conformation and the rapid mobility of a BR probe attached to a protein, using the example of the C helix of troponin C.

The amplitude of probe motion on a time scale faster than the ca. 4 ns fluorescence lifetime is conveniently quantified by the order parameter $\langle P_{2d} \rangle$.[12] Fluorescence polarization measurements of the BR probe on the C helix of TnC in an actively contracting muscle fiber gave $\langle P_{2d} \rangle = 0.91 \pm 0.03$.[8] In a separate NMR study, the amplitude of a 0.8 ns component of the motion of the N-CH_3 axis of the BR probe attached to the same site on the N-terminal lobe of TnC in solution (sNTnC·$2Ca^{2+}$·$TnI_{115-131}$·BR_{56-63}) determined by ^{13}C relaxation measurements was characterized by an order parameter with a similar value, 0.89 (note that the result given in that work was 0.79 for the squared order parameter, S^2; its square root corresponds to the 0.89 value quoted here).[13] In the present work with the same probe-protein conjugate, $\langle P_{2d} \rangle$ for the long axis of the xanthylium system, which corresponds to the orientation of the fluorescence dipole, was in the range 0.95−0.99 for motion on the subnanosecond time scale within the different conformational basins identified by the molecular dynamics analysis. The $\langle P_{2d} \rangle$ values estimated by these three approaches are not directly comparable. For example, the comparison between the optical and MD values might be affected by depolarization in the experimental optics, or effective nonparallel orientation of the probe absorption and emission dipoles. However all three approaches show that the independent nanosecond motion of the BR probes is highly restricted, and that residual probe motion is sufficiently small to allow reliable correction of the fluorescence data for depolarization produced by probe motion between absorption and emission of a photon.[12] In consequence, fluorescence polarization data from these bifunctional probes give useful information about the orientation and motion of the local region of the protein where they are attached.

Immobilization of the bifunctional probes is not a direct consequence of their two-site attachment to the protein; the length and flexibility of the linkers between the xanthylium and the protein attachment sites are not sufficient to explain the observed immobilization. Moreover, monofunctional iodoacetamidotetramethylrhodamine probes attached via a single cysteine to myosin regulatory light chains have similarly restricted motion to that of bifunctional rhodamine probes on the same region of this protein.[6,36] In both cases immobilization is likely to be mediated by an interaction with the protein surface, and the molecular dynamics analysis presented here provides detailed

(36) Hopkins, S. C.; Sabido-David, C.; Corrie, J. E. T.; Irving, M.; Goldman, Y. E. *Biophys. J.* **1998**, *74*, 3093–3110.

information about these interactions and how they constrain the orientation of a bifunctional probe with respect to its protein attachment sites.

Although many aspects of the probe/protein/solvent interaction play a role in this constraint, there is a clear hierarchy of factors. The dominant effect is associated with projection of the carboxylate group of BR into the solvent as the hydrophobic face of the xanthylium moiety closely approaches the protein surface. This effect is preserved throughout the systems of increasing complexity that are reported here. The proximity of the probe to the protein surface then brings into play specific probe-protein interactions, including steric factors, electrostatic and hydrophobic interactions, torsion angles for the labeled cysteines and within the linker arms, hydrogen bonds and hydration effects, which determine the detailed probe conformation. The steric factors include both local effects of the side chains on the labeled helix (Figure 2) and interactions with other parts of the protein surface (Figure 3), and these can be taken into account in designing probe attachments to proteins of known structure. Hydrophobic patches on the surface of the protein are also important in stabilizing probe conformations on the protein surface (Figure 4), and there is good agreement between the MD-based conformations and predictions of hydrophobic regions based on protein structure. Although hydrogen bonds seem not to play a dominant role in the overall folding, they may constrain the final preferred conformation in some cases.

The existence of two diastereosiomers in BR-sNTnC, each of which has two conformational minima with limited conformational mobility around the minima, gives insight into the lack of evidence for specific probe-protein contacts in NMR studies of the mixture of diastereoisomers.[13] Neither NOEs involving ^{13}C-labeled methyl groups of BR, nor protection of backbone amides by the probe against hydrogen/deuterium exchange were observed in those experiments. Since the footprint of the probe is widely distributed over the protein surface (Figure 5A), its effects on any particular atom of the protein ensemble population will be limited and thus the ability to detect such effects will be severely constrained.

In each of the four conformations of BR-sNTnC, the long axis of the xanthylium system, corresponding to the fluorescence dipole, is roughly parallel to the line joining the two cysteine residues to which the probe is attached, the largest deviation being for basin 2 of the *anti*-isomer (Table 1 and Figure 5A). At this level of approximation the results thus confirm the basic assumption used for the interpretation of BR probe data in terms of helix orientation.[6,8,9] A quantitative analysis of the effect of the probe conformations deduced from molecular dynamics on the calculated conformation of the orientation of the N-terminal lobe of TnC in a muscle fiber showed that application of this assumption led to an error of about 10°. Since the orientational information obtained by polarized fluorescence is limited to determination of terms in $\cos^2 \theta$ and $\cos^4 \theta$,[12] and therefore of relatively low angular resolution, the 10° error will not in general have serious consequences for the interpretation of such data. In practice, the data are usually combined from a set of probes attached to different sites on a protein domain,[6,8] further reducing the error in domain orientation.

Although the molecular dynamics analysis reported here refers to one specific protein, troponin C, the comparison of the results for the whole domain with those for isolated helices with different sequences, and for the two diasteroisomers in each system, shows that the conclusions described above are likely have general application to other proteins. In extrapolating the present results it will be important to consider potential steric and hydrophobic interactions in the conformational space available to the probe. The physicochemical properties of rhodamine, in particular the combination of the hydrophobic xanthylium moiety and the carboxylate substitution on the pendant aryl ring, appear to be generally advantageous in defining the orientation of the fluorescence dipole with respect to the protein. The presence of probe-protein diastereoisomers which are not readily separable complicates the analysis, but may ameliorate the effect of specific hydrophobic patches on the protein surface which are unlikely to be available to both isomers.

Conclusion

The results described here, while influenced in detail by specific features of the particular protein environment, are likely to have wider generality for understanding the interactions of rhodamine probes with proteins to which they are attached. In the area of our immediate interest, the results provide important support to the use of fluorescence polarization measurements with bifunctional rhodamine probes to determine *in situ* orientations of protein domains, but also provide insights on the origins of the restricted motion of monofunctional rhodamines attached to proteins. The developing understanding of probe-protein interactions from this and related work is likely further to inform and reinforce the widespread use of fluorescence probes in biophysics.

Acknowledgment. A.D. acknowledges a grant from the European Molecular Biology Organization, Fellowship No. 584-2007. F.F., M.I. and J.C. acknowledge the U.K. Medical Research Council for funding.

Supporting Information Available: Detailed views of the starting conformations of each of the simulated systems (Figure S1); REMD statistics (Figure S2); Hydration density profile of isolated rhodamine (Figure S3); Close-up views of BR conformations in the main basins of BR-sNTnC samplings. (Figure S4). This material is available free of charge via the Internet at http://pubs.acs.org.

JA807264V