Simulation of fluorescence resonance energy transfer experiments: effect of the dyes on protein folding

2010 J. Phys.: Condens. Matter 22 235103

(http://iopscience.iop.org/0953-8984/22/23/235103)

J. Phys.: Condens. Matter **22** (2010) 235103 (7pp) doi:10.1088/0953-8984/22/23/235103

Simulation of fluorescence resonance energy transfer experiments: effect of the dyes on protein folding

Lucy R Allen[1] **and Emanuele Paci**[1,2,3]

[1] School of Physics and Astronomy, University of Leeds, Leeds LS2 9JT, UK
[2] Institute of Molecular and Cellular Biology, University of Leeds, Leeds LS2 9JT, UK
[3] Astbury Centre for Structural Molecular Biology, University of Leeds, Leeds LS2 9JT, UK

E-mail: e.paci@leeds.ac.uk

Received 2 March 2010, in final form 16 April 2010
Published 21 May 2010
Online at stacks.iop.org/JPhysCM/22/235103

Abstract
Fluorescence resonance energy transfer is a powerful technique which is often used to probe the properties of proteins and complex macromolecules. The technique relies on relatively large fluorescent dyes which are engineered into the molecule of interest. In the case of small proteins, these dyes may affect the stability of the protein, and modify the folding kinetics and the folding mechanisms which are being probed. Here we use atomistic simulation to investigate the effect that commonly used fluorescent dyes have on the folding of a four-helix bundle protein. We show that, depending on where the dyes are attached, their effect on the kinetic and thermodynamic properties of the protein may be significant. We find that, while the overall folding mechanism is not affected by the dyes, they can destabilize, or even stabilize, intermediate states.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Experiments based on the measurement of fluorescence have long provided crucial information on the thermodynamics and kinetics of protein folding [1–3]. Naturally occurring fluorescent residues such as tryptophan or tyrosine which are quenched in one of the folded or unfolded states are often used as probes of the folding process. Where no suitable naturally occurring residues are present, mutations which introduce fluorescence can be made. In recent years fluorescent probes have become commonly used in a different way: to probe intramolecular distances in proteins using a technique known as fluorescence resonance energy transfer (FRET). FRET is the transfer of energy between two fluorophores, a donor and an acceptor, upon excitation of the donor fluorophore. Importantly, and as shown by Förster in his original work [4], the rate of energy transfer depends strongly on the distance between the two fluorophores. Thus, if two of the residues in a protein can be made to act as fluorophores, in theory the inter-residue distance can be measured by exciting the donor fluorophore and using the number of photons emitted by both the donor and acceptor fluorophores to calculate the energy transfer efficiency. These measurements can be made at a single molecule level, for example by limiting the concentration of the protein solution so that only one molecule is present in the confocal volume at any one time [5]. Such advanced and specific measurements are not obtainable in any other way, and could provide a great deal of insight into the folding process. In fact, FRET has already provided a large amount of valuable information, particularly concerning the collapse of the denatured state on reduction of denaturant concentration [6–11]. However, the measurement and interpretation of FRET data is a complicated process, and some issues related to the process remain unresolved. One such issue is that of the orientational factor, κ^2, which has been widely debated [12–16]. Another issue, which we investigate in this paper, concerns the effect of the fluorophores themselves on the thermodynamics and kinetics of the folding process. Whilst naturally occurring fluorescent residues can often be used as probes of folding, they are rarely useful for FRET measurements. The reason for this is that resonant energy transfer only occurs between certain pairs of fluorophores.

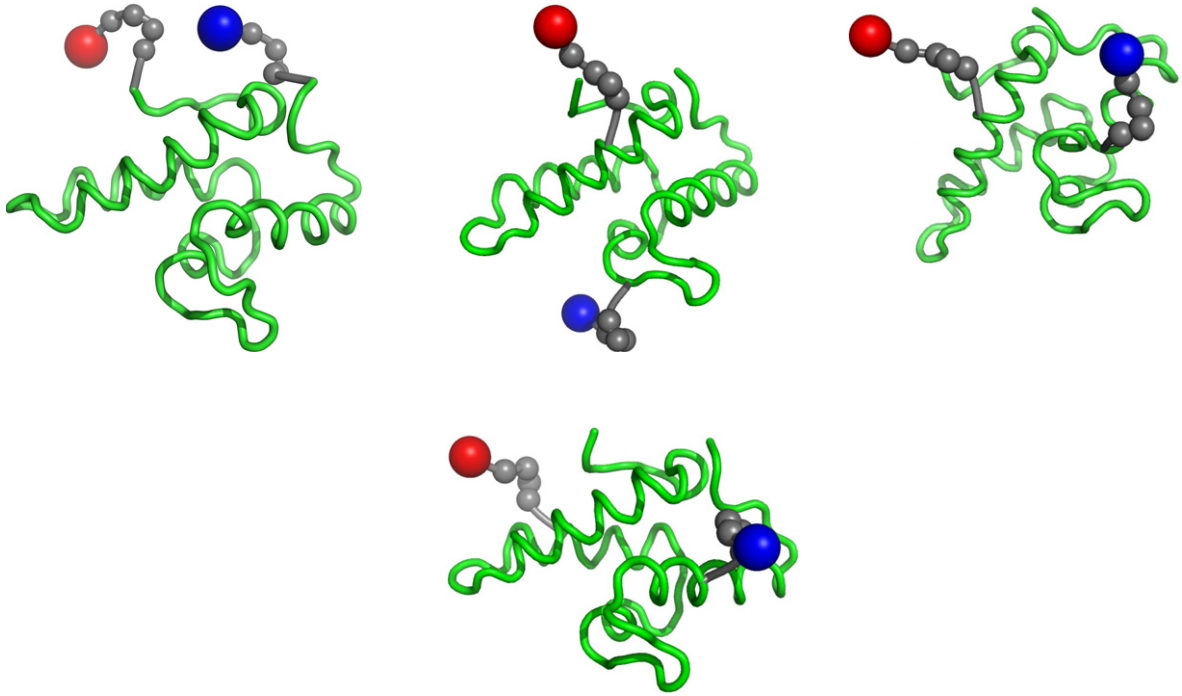

Figure 1. Im7 with protein (green), linker (grey) and dye (red/blue) beads attached. 1–86 (top left), 17–56 (top centre), 17–70 (top right) and 36–70 (bottom centre) variants shown.

For example, energy transfer will not occur unless there is some overlap between the emission spectrum of the donor fluorophore and the absorption spectrum of the acceptor fluorophore. Even if there is some overlap between the spectra, the length range over which the transfer occurs depends on a number of properties of the fluorophores and the medium between them. These requirements limit considerably the choice of natural occurring fluorophores. Usually, artificial fluorophores which have been developed specifically for the purpose of FRET measurements are attached to the protein at certain positions. The artificial fluorophores are generally attached covalently, via flexible linkers. This procedure is very convenient, as fluorophores which are suitable for the protein under investigation can be chosen, and because experiments can be carried out on variants of the same protein in which different pairs of residues are labelled, thus building up a map of inter-residue distances [17]. However, the artificial fluorophore molecules which are attached to the protein molecule are often relatively large compared to the protein under investigation. For example, the Alexa Fluor dyes which have been used to study the small Cold-Shock protein Csp*Tm* [6, 11, 17, 18], have a mass which is 20% of the mass of the wild-type protein. It is not inconceivable that attaching a mass of this magnitude could alter the folding mechanism or even the native structure of the protein. The effect of attaching the dye molecules is often assessed by simply measuring changes in stability of the native state (see SI of [17], for example); this information alone could hide changes in folding pathway or native state structure. Generally, attaching dyes has the effect of destabilizing the native state.

This issue can be, and has been to some extent, addressed by simulation: Schroder *et al* [19] found, using all-atom force fields, that attaching a dye to the loop region of a fragment of the helical protein bacteriorhodopsin had only a small effect on the native state dynamics of the protein, but did not look at the folding behaviour.

In this paper we use a structure-based, coarse-grained model of protein, linker and dye molecules to look at the issue discussed above. The specific protein studied is Im7, a small four-helix bundle, which has been recently studied experimentally [20], and which is homologous to the protein Im9, which has also been studied using FRET [21]. Most of the conclusions, however, relate to the general effects that attaching dyes to proteins may have, and do not depend on the specific properties of Im7.

A native-centric model is used for the protein, in which only native interactions are favourable and persistent non-native interactions are ruled out, i.e., the denatured state owes its stability to its high entropy alone. Such a model allows, unlike more accurate models, full equilibrium to be reached and reversible folding to be observed. The linker and dyes were included in the model by attaching a chain of hard-sphere-like beads to the protein (see section 2.1). The four variants illustrated in figure 1 are investigated: with the linker and dye beads attached to the two termini (the 1–86 variant), helices 1 and 3 (the 17–56 variant), helices 1 and 4 (the 17–70 variant) and helices 2 and 4 (the 36–70 variant).

2. Methods

2.1. Model

The model used is an extension of the structure-based protein model described in Karanicolas and Brooks [22, 23]. Each

amino acid is represented as a bead on a chain, and the interactions between the beads depend on the experimental native structure of the protein, so that the native state is the most enthalpically stable. The model includes implicitly the effects of the side chains and the solvent. Note that, as one of the 87 residues is unresolved in the PDB entry for Im7 (1AYI), the model contains only 86 protein beads. The model has been modified to include a simple representation of the linker and dye molecules, in a similar way to that described in Merchant *et al* [10]. The dye and linker molecules are based on the Alexa Fluor 488 c-5 maleimide fluorophore, and are each represented by a chain of five beads attached to a protein residue. The chain consists of four linker beads and one dye bead, each representing a group of atoms within the molecule. The masses of the beads are equal to the total mass of the group of atoms in question. An all-atom simulation of the Alexa Fluor 488 c-5 maleimide attached to a string of three amino acids was carried out with the MS modelling package and 'COMPASS' force field (Accelrys Inc., 2003). The results of this simulation were used to choose suitable values for the bond length, radii and angular and dihedral parameters for the coarse-grained linker and dye model. The dye and linker beads interact repulsively with each other and the protein beads. The forms of the potentials are the same as those in the protein model. The results shown in this report are for simulations with dyes in positions 1–86 (i.e., at the C- and N-*termini*), 17–56 (helix 1–helix 3), 17–70 (helix 1–helix 4) and 36–70 (helix 2–helix 4).

2.2. Simulations

Langevin dynamics simulations of the four models described above were carried out at temperatures ranging from 250 to 500 K. The models were implemented within the CHARMM program [24]. A friction coefficient of 1 ps^{-1}, low enough to guarantee the generation of a large sample of folding and unfolding events, but not in a ballistic, low friction regime [25], was used. The SHAKE algorithm was applied to C_α–C_α bonds, and a timestep of 10 fs was used. Coordinates were saved every 1 ps, and simulations run for 2 μs.

2.3. Analysis

The effect of attaching dyes to the model on the thermodynamics of folding was assessed using specific heat curves, and projections of the trajectory onto the two-dimensional reaction coordinate (RMSD, Q_N), where RMSD is the root mean square deviation from the experimental structure and Q_N is the fraction of native contacts. RMSD from the experimental native structure was calculated using the core residues 5–81. Contacts were considered to be present if two C_α atoms separated by more than four residues in sequence were less than 12 Å apart. The average native state structure of the wild-type (WT) protein from a short trajectory at 300 K was used to calculate the native contact map. The separation of the N, I and D states in the projections of (RMSD, Q_N) made it possible to split the trajectories into N, I and D parts according to these criteria. The accuracy of the separation into the different states was verified using the robust method [26] which uses dynamic, rather than geometric, criteria, to extract relative free energies and mutual interconversion rates from the trajectory. The specific heat capacity was calculated as a function of temperature using the weighted histogram analysis method (WHAM) [27].

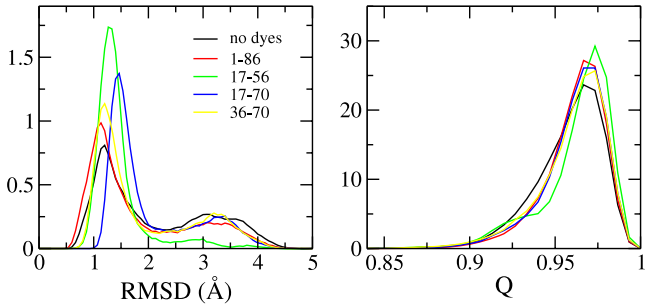

Figure 2. Histograms of RMSD from experimental native structure and fraction of native contacts formed, Q_N, for simulations of wild-type Im7 model and four variants at 250 K.

3. Results

A first interesting finding is that the attachment of dyes does not greatly perturb the native state. Figure 2 shows histograms of RMSD from native structure and fraction of native contacts, Q_N for the trajectories of the wild-type protein and the four variants at 250 K (a temperature sufficiently low to ensure that the probability of unfolding is zero). Throughout the trajectories the RMSD is <5 Å and Q_N is >0.86, indicating that the variants remain in a native-like state. Interestingly, the RMSD projection reveals that for the wild-type and all the variants there are two native sub-states, the most populated (sub-state 1) having a low value of RMSD (≈1.2 Å for the wild-type) and the less populated (sub-state 2) having a slightly higher value of RMSD (≈3.2 Å for the wild-type). The position of sub-state 1 on the RMSD projection varies slightly between variants, with the 17–70 variant having the highest value (≈1.5 Å), indicating that attaching the dyes in these positions perturbs the native state the most. The relative populations of the two sub-states are approximately the same for the wild-type and all variants except 17–56, which only marginally populates sub-state 2. This can be understood by analysing the structural differences between the two states. Figure 3 shows the RMS fluctuations of the wild-type 250 K trajectory, split into the two sub-states according to their values of RMSD (RMSD < 2.25 Å or RMSD > 2.25 Å). The main difference lies in the loop region between helix 3 and helix 4, which is more flexible in sub-state 2. Residue 56 lies in this loop region: attaching the linker and dye molecule at this point has a dampening effect on the loop and destabilizes sub-state 2.

Figure 4 shows the effect of the attachment of the dyes on the heat capacity of the protein. The wild-type curve is shown in black, and the curves for the variants with dyes attached in different colours. The heat capacity curve of the wild-type has a broad peak indicating a non-cooperative native to denatured transition: this qualitatively agrees with experimental data which shows Im7 to fold via an intermediate [28–31] and

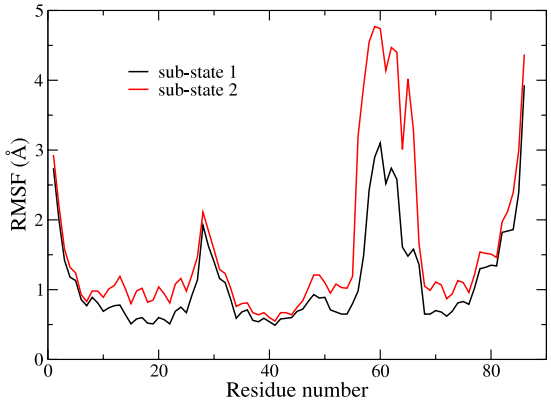

Figure 3. RMSFs of C_α atoms from simulation of wild-type Im7 model at 250 K, separated into two native sub-states according to RMSD values.

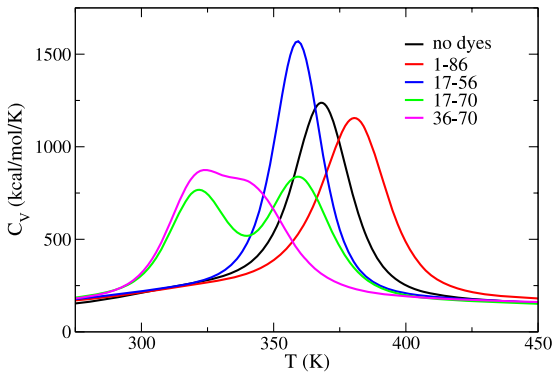

Figure 4. Heat capacities for Im7 with and without dyes.

Table 1. Enthalpy and entropy changes for the unfolding transition, estimated from the heat capacity curves.

Variant	WT	1–86	17–56	36–70
ΔH (kcal mol^{-1})	54.1	57.7	64.2	58.6
$\Delta H - \Delta H_{\mathrm{WT}}$ (kcal mol^{-1})	—	3.6	10.1	4.5
ΔS (kcal mol^{-1} K^{-1})	0.146	0.152	0.178	1.175
$\Delta S - \Delta S_{\mathrm{WT}}$ (kcal mol^{-1} K^{-1})	—	0.006	0.032	0.029

with simulations of an analogous model [32]. The effect of the dyes depends strongly on their positions, and, as the specific heat shows, can be significant. Interestingly, in the case of the 1–86 variant, the heat capacity curve peaks at a higher temperature than the wild-type curve: this indicates that when attached at the two *termini* the dyes stabilize the native state. The protein is destabilized by the dyes (i.e., the peaks of the heat capacity curves shift to lower temperature) in the three other cases (17–56, 17–70 and 36–70). The width of the heat capacity peak also varies strongly according to dye position; in some cases the peak becomes narrower than the wild-type, indicating a more cooperative transition (e.g., 17–56) and in others becoming wider, indicating a less cooperative transition (e.g., 1–86, 36–70). The bimodal shape of the 17–70 protein is particularly interesting, and indicates that the protein undergoes two distinct transitions.

The origins of the changes in native state stability upon addition of the dyes can be investigated by studying the heat capacity curves in more detail. Following suitable baseline subtractions for both the native and denatured states, the enthalpy and entropy changes for the unfolding transitions can be estimated by integrating the curves:

$$\Delta H = \int_{T_0}^{T_\mathrm{f}} \mathrm{d}T\, C_P \tag{1}$$

$$\Delta S = \int_{T_0}^{T_\mathrm{f}} \mathrm{d}\ln T\, C_P. \tag{2}$$

Note that, as our simulations do not explicitly include any solvent molecules, they are effectively carried out at infinite volume and therefore zero pressure, $C_V = C_P$. T_0 and T_f are the onset and completion temperatures of the transition. Using regions of the curve in which the protein is either completely native or completely denatured (280–300 K and 400–450 K) to calculate linear baselines, values of ΔH and ΔS were estimated for the variants with unimodal heat capacity curves (i.e., WT, 1–86, 17–56 and 36–70). The results are shown in table 1. Unsurprisingly, for all variants both values are positive i.e., the unfolding transition is enthalpically unfavourable and entropically favourable. On addition of the dyes both ΔH and ΔS increase. As the dye and linker beads make no attractive interactions, the increase in ΔH must arise from enthalpic destabilization of the denatured state relative to the native state. This relative destabilization is most likely a result of increased unfavourable steric interactions: whilst in the native state the dye/linker beads can only collide with a limited portion of the chain close to the position of attachment, in the denatured state they are free to collide with a much larger portion of the chain. The increase in ΔS on addition of the dyes indicates that the dye/linker beads have more freedom in the denatured state than in the native state. Note that the increase in ΔS is smallest for the 1–86 variant; this is because the movement of the dye/linker beads in the native state is much less restricted when they are positioned at the two *termini*. The overall stabilization or destabilization of the variants relative to the wild-type protein results from a careful balance of the enthalpic stabilization and the entropic destabilization of the native state described above. In the 1–86 variant ΔS is sufficiently small that enthalpy is the dominant factor, and the native state is stabilized. In the other variants the larger ΔSs make entropy the dominant factor, thus destabilizing the native state.

Experimentally, the attachment of dyes is generally found to destabilize native states [6, 17]; this may be partially due to the entropic stabilization of the unfolded state observed in our simulations, but it is likely that a number of interactions not included in our model are important. Alexa Fluor dyes, for example, are negatively charged and hydrophilic, whereas the dye beads in our model have only hard-sphere-like repulsive interactions. In addition, the model does not explicitly include solvent molecules, which provide a large contribution to experimentally measured enthalpy and entropy changes. Thus whilst the above results provide information about some the factors which might cause stability changes of a protein on attachment of dyes, the absolute values of enthalpy and entropy

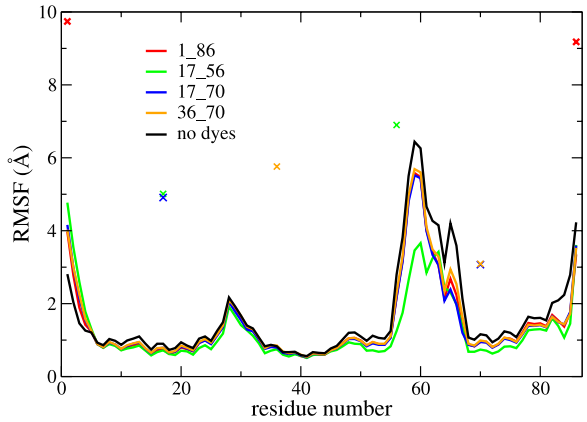

Figure 5. RMSFs of C_α atoms from simulations of Im7 and Im7 + dyes at 250 K (lines). RMSFs of dye beads are shown as crosses.

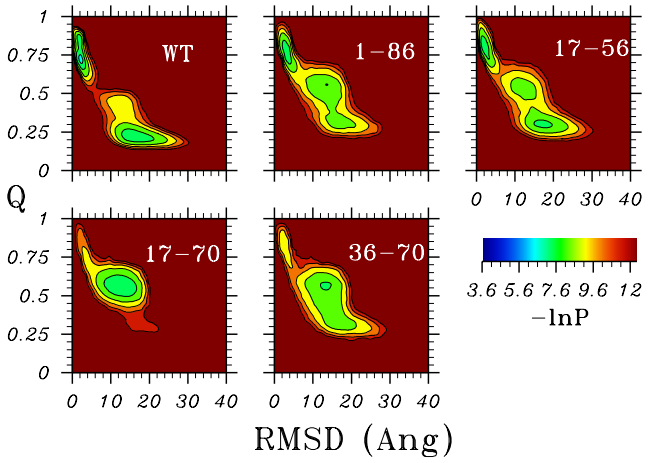

Figure 6. Projections of $-\ln p$ onto RMSD/Q_N for simulations of wild-type Im7 and Im7 with dyes in different positions at the temperature at which the heat capacity (figure 4) peaks.

do not correspond to any experimentally measurable quantities. The effects observed in our simulations are also dependant on the size of the fluorophore and the protein: we would expect that the smaller the size of the fluorophore in relation to the protein, the smaller the perturbation of the energy landscape. The structure of the protein is also important, with certain proteins having more robust energy landscapes than others.

Figure 5 compares the RMS fluctuations (RMSF) of the C_α beads of each variant at 250 K (i.e., in the native state) with the wild-type protein. The effect of the dyes is generally very small; only the 17–56 variant differs significantly from the wild-type. The difference lies in the region between helices 3 and 4: the dye in position 56 (at the end of helix 3) has the effect of dampening the fluctuations in this region. The RMSFs are higher for dye beads than for the protein beads: as they make no attractive interactions they are able to move more freely. The dye beads in the 1–86 variant fluctuate more than those in the other variants; this is in agreement with the low ΔS for this variant which arises from the increased freedom of the dyes in the native state.

Figure 6 shows $-\ln p$, where $p = p(\text{RMSD}, Q_N)$ is the probability of finding a conformation with fraction of native contacts $Q_N = Q_N$ and RMSD from the native state RMSD for simulations of the proteins at the temperature at which N and D are approximately equally populated. For all the variants, three states are visible, in the same region of RMSD/Q_N in each. This suggests that the folding mechanism does not change significantly when the dye/linkers are attached: three states are always evident and their features, as far as can be determined from the coordinates fraction of native contacts and root mean square deviation from the native state, are unchanged. Further inspection of conformations from the intermediate basins for the different variants confirms that for all variants they share the same basic structure: helices 1–3 native-like, helix 4 formed but not docked to the rest of the structure. Note that, while the basic 3-state mechanism agrees well with experimental data on Im7 [28], the intermediate state of the real protein is believed to be stabilized by non-native interactions, with helices 1, 2 and 4 formed and docked together, and helix 3 unformed [29, 33]. This difference highlights a weakness of models based solely on native interactions [34, 35].

Whilst the same three states are populated by all the variants, the occupancy of each of the states differs. The relative stability of the intermediate, and the size of the free-energy barriers clearly depend on the presence, and the positions, of the dyes. For example, in the 17–70 variant the intermediate state is strongly populated with the native and denatured basins only marginally occupied: either the intermediate state is stabilized or the free-energy barriers between I and D, and I and N, are increased by attaching the dyes in these positions. The existence of two distinct transitions (N $\rightarrow$ I and I $\rightarrow$ D) explains the bimodal heat capacity curve seen in figure 4.

4. Discussion

In recent years FRET has become a widely used tool for probing inter-residue distances in proteins, and has uncovered a great deal of valuable information, particularly about the nature of the unfolded state. FRET is a process which occurs between two fluorophores, and FRET experiments which investigate protein folding therefore require either two of the proteins' amino acids to be fluorescent, or the attachment of two fluorophores to the protein. Typically, the latter option is taken, and fluorophores such as Alexa Fluor dyes are attached to the protein via flexible linker molecules. In this paper, we have used molecular dynamics simulations to explore the effects that relatively large linkers and dyes may have on the thermodynamics and kinetics of folding of relatively small proteins.

Such a dissection is important as FRET is now a frequently used technique which has provided a large amount of detailed information, particularly concerning the denatured state of proteins. As folding and unfolding occur a number of times during a typical FRET measurement, the experiment can only be mimicked using a model for which full equilibrium can be reached during a simulation (and which folds reversibly

within the simulation time). This requirement limits the number of useable models; models which represent chemical interactions transferably based on the properties of individual amino acids only allow the observation of reversible folding for a few short peptides and preclude the exploration of the equilibrium folding of even small proteins. Here we have used a coarse-grained, structure-based model, adapted to include the linker and dye molecules as beads with only excluded volume interactions attached to various points on the protein chain.

The first issue which has been addressed is the effect of attaching dye molecules on the thermodynamics and kinetics of folding. The Alexa Fluor 488 dye, which is often used in FRET experiments, and on which our model was based, has a molecular mass of 721, around 7% of the mass of the protein Im7. The effect of attaching two of the dyes to the protein is not only to significantly increase the mass of the protein, but also to introduce excluded volume interactions and specific dye–solvent or dye–protein interactions. In reality Alexa Fluor dyes are both negatively charged and hydrophilic while in our model they are treated simply as entropic weights with only excluded volume interactions.

The effect of attaching the dyes to the model depends strongly on their positions. Depending on the relative effects that the dyes in different positions have on the enthalpy and entropy changes associated with the folding transition, the native state is either stabilized or destabilized. In the case of the four-helix bundle Im7 protein studied here, although the size of the energy barriers is in some cases significantly increased, even causing the intermediate state to become the most stable at the melting temperature in one case, the folding pathway is unchanged by attaching the dyes. The folding of the protein Im7 and the homologous Im9, as well as large number of mutants thereof has been studied in great depth in the group of Sheena Radford in the past decade: while the population of an obligatory intermediate depends strongly on the variant studied, the folding mechanism is robust and the pathway identical [28, 30, 31]. For other proteins, the presence and position of the dyes may well affect the folding mechanisms. This is even more likely in the case of 'real' dyes as opposed to the 'ideal' dyes used in the simulation; the latter interact only via steric effects, have a non-negligible contribution to the thermodynamics of folding which depend on the position of the dyes.

Acknowledgments

This work was supported by a EPSRC studentship to LRA. We acknowledge stimulating discussions with Sheena Radford, Sara Pugh, Daniel Nettels and Ben Schuler.

References

[1] Stryer L 1968 Fluorescence spectroscopy of proteins *Science* **162** 526–9
[2] Eftink M R and Ghiron C A 1981 Fluorescence quenching studies with proteins *Anal. Biochem.* **114** 199–227
[3] Beechem J M and Brand L 1985 Time-resolved fluorescence of proteins *Annu. Rev. Biochem.* **54** 43–71
[4] Förster T 1948 Zwischenmolekulare energiewanderung und fluoreszenz *Ann. Phys., Lpz.* **2** 55–75
[5] Weiss S 1999 Fluorescence spectroscopy of single biomolecules *Science* **283** 1676–83
[6] Schuler B, Lipman E A and Eaton W A 2002 Probing the free-energy surface for protein folding with single-molecule fluorescence spectroscopy *Nature* **419** 743–7
[7] Lipman E A, Schuler B, Bakajin O and Eaton W A 2003 Single-molecule measurement of protein folding kinetics *Science* **301** 1233–5
[8] Rhoades E, Gussakovsky E and Haran G 2003 Watching proteins fold one molecule at a time *Proc. Natl Acad. Sci. USA* **100** 3197–1202
[9] Kuzmenkina E V, Heyes C D and Nienhaus G U 2005 Single molecule Förster resonance energy transfer study of protein dynamics under denaturing conditions *Proc. Natl Acad. Sci. USA* **102** 15471–6
[10] Merchant K A, Best R B, Louis J M, Gopich I V and Eaton W A 2007 Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations *Proc. Natl Acad. Sci. USA* **104** 1528–33
[11] Nettels D, Gopich I V, Hoffmann A and Schuler B 2007 Ultrafast dynamics of protein collapse from single-molecule photon statistics *Proc. Natl Acad. Sci. USA* **104** 2655–60
[12] Eisinger J and Dale R E 1974 Letter: interpretation of intramolecular energy transfer experiments *J. Mol. Biol.* **84** 643–7
[13] dos Remedios C G and Moens P D 1995 Fluorescence resonance energy transfer spectroscopy is a reliable 'ruler' for measuring structural changes in proteins. Dispelling the problem of the unknown orientation factor *J. Struct. Biol.* **115** 175–85
[14] van der Meer B W, Coker G I and Chen S Y S 1994 *Resonance Energy Transfer: Theory and Data* (Weinheim: Wiley–VCH)
[15] VanBeek D B, Zwier M C, Shorb J M and Krueger B P 2007 Fretting about FRET: correlation between kappa and R *Biophys. J.* **92** 4168–78
[16] Allen L R, Krivov S and Paci E 2009 Analysis of the free-energy surface of proteins from reversible folding simulations *PLoS Comput. Biol.* **5** e1000428
[17] Hoffmann A, Kane A, Nettels D, Hertzog D E, Baumgartel P, Lengefeld J, Reichardt G, Horsley D A, Seckler R, Bakajin O and Schuler B 2007 Mapping protein collapse with single-molecule fluorescence and kinetic synchrotron radiation circular dichroism spectroscopy *Proc. Natl Acad. Sci. USA* **104** 105–10
[18] Nettels D, Hoffmann A and Schuler B 2008 Unfolded protein and peptide dynamics investigated with single-molecule FRET and correlation spectroscopy from picoseconds to seconds *J. Phys. Chem.* B **112** 6137–46
[19] Schröder G F, Alexiev U and Grubmuller H 2005 Simulation of fluorescence anisotropy experiments: probing protein dynamics *Biophys. J.* **89** 3757–70
[20] Pugh S, Gell C, Smith D A, Radford S E and Brockwell D J 2010 Single-molecule studies of the Im7 folding landscape *J. Mol. Biol.* **398** 132–45
[21] Tezuka-Kawakami T, Gell C, Brockwell D J, Radford S E and Smith D A 2006 Urea-induced unfolding of the immunity protein Im9 monitored by spFRET *Biophys. J.* **91** L42–4
[22] Karanicolas J and Brooks C L 2002 The origins of asymmetry in the folding transition states of protein L and protein G *Prot. Sci.* **11** 2351–61
[23] Karanicolas J and Brooks C L 2003 The structural basis for biphasic kinetics in the folding of the WW domain from a formin-binding protein: lessons for protein design? *Proc. Natl Acad. Sci. USA* **100** 3954–9
[24] Brooks B R, Bruccoleri R E, Olafson B D, States D J, Swaminathan S and Karplus M 1983 CHARMM: a program

for macromolecular energy, minimization and dynamics calculations *J. Comput. Chem.* **4** 187–217
[25] Rhee Y M and Pande V S 2008 Solvent viscosity dependence of the protein folding dynamics *J. Phys. Chem.* B **112** 6221–7
[26] Krivov S V, Chekmarev S F and Karplus M 2002 Potential energy surfaces and conformational transitions in biomolecules: a successive confinement approach applied to a solvated tetrapeptide *Phys. Rev. Lett.* **88** 038101
[27] Kumar S, Bouzida D, Swendsen R H, Kollman P A and Rosenberg J M 1992 The weighted histogram analysis method for free-energy calculations on biomolecules. 1. The method *J. Comput. Chem.* **13** 1011–21
[28] Ferguson N, Capaldi A P, James R, Kleanthous C and Radford S E 1999 Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9 *J. Mol. Biol.* **286** 1597–608
[29] Capaldi A P, Kleanthous C and Radford S E 2002 Im7 folding mechanism: misfolding on a path to the native state *Nat. Struct. Biol.* **9** 209–16
[30] Friel C T, Capaldi A P and Radford S E 2003 Structural analysis of the rate-limiting transition states in the folding of Im7 and Im9: similarities and differences in the folding of homologous proteins *J. Mol. Biol.* **326** 293–305
[31] Paci E, Friel C T, Lindorff-Larsen K, Radford S E, Karplus M and Vendruscolo M 2004 Comparison of the transition states ensembles for folding of Im7 and Im9 determined using all-atom molecular dynamics simulations with ϕ value restraints *Proteins* **54** 513–25
[32] Sutto L, Latzer J, Hegler J A, Ferreiro D U and Wolynes P G 2007 Consequences of localized frustration for the folding mechanism of the IM7 protein *Proc. Natl Acad. Sci. USA* **104** 19825–30
[33] Gsponer J, Hopearuoho H, Whittaker S B-M, Spence G R, Moore G R, Paci E, Radford S E and Vendruscolo M 2006 Determination of an ensemble of structures representing the intermediate state of the bacterial immunity protein Im7 *Proc. Natl Acad. Sci. USA* **103** 99–104
[34] Paci E, Vendruscolo M and Karplus M 2002 Native and non-native interactions along protein folding and unfolding pathways *Proteins* **47** 379–92
[35] Paci E, Vendruscolo M and Karplus M 2002 Validity of Gō models: comparison with a solvent-shielded empirical energy decomposition *Biophys. J.* **83** 3032–8