Cell. Mol. Life Sci. 66 (2009) 1672 – 1681
1420-682X/09/101672-10
DOI 10.1007/s00018-009-8771-9

Cellular and Molecular Life Sciences

Review

The perspectives of studying multi-domain protein folding

J. Fitter

Research Center Jülich, ISB-2: Molecular Biophysics, 52425 Jülich (Germany), Fax: +49-2461-612020,
e-mail: j.fitter@fz-juelich.de

Received 3 December 2008; accepted 23 December 2008
Online First 2 February 2009

Abstract. Most of fundamental studies on protein folding have been performed with small globular proteins consisting of a single domain. *In vitro* many of these proteins are well characterized by a reversible two-state folding scheme. However, the majority of proteins in the cell belong to the class of larger multi-domain proteins that often unfold irreversibly under *in vitro* conditions. This makes folding studies difficult or even impossible. In spite of these problems for many multi-domain proteins, folding has been investigated by classical refolding. Co-translational folding of nascent polypeptide chains when synthesized by ribosomes has also been studied. Single molecule techniques represent a promising approach for future studies on the folding of multi-domain proteins, and tremendous advances have been made in these techniques in recent years. In particular, fluorescence-based methods can contribute significantly to an understanding of the fundamental principles of multi-domain protein folding.

Keywords. Irreversible unfolding, protein aggregation, domain interaction, co-translational folding, single molecule studies, fluorescence correlation spectroscopy

Introduction

In prokaryotic and even more in eukaryotic cells the predominant fraction of the whole proteome belongs to the class of multi-domain proteins [1]. However, for methodical reasons, our existing knowledge about mechanisms and principles of protein folding results mainly from studies on smaller single-domain proteins. In particular, competing side reactions such as misfolding and aggregation of non-native states makes folding studies on larger multi-domain proteins often difficult or even impossible under *in vitro* conditions. Despite these problems, a number of multi-domain proteins have been investigated in experimental folding studies (see reviews [2–4] and references therein). Keeping in mind that domains are stable globular substructures of a polypeptide chain that generally represent autonomous folding units, it is obvious that folding of multi-domain proteins is characterized by an additional hierarchical level in the folding process that is absent in single-domain proteins. Although individual domains of multi-domain proteins seem to fold by the same basic principles as small single-domain proteins, there is at least one major feature that makes multi-domain protein folding quite different. Domain-domain interactions, which are absent in small single-domain proteins, can play a central role in protein stability, in protein folding kinetics, and in protein misfolding. In particular, the nature of domain interfaces and the length of the interdomain linker region determine whether or to what extend domain interactions have an impact on protein folding [3, 5].

From studies performed so far, there is evidence that individual domains of multi-domain proteins fold independently when the interaction between the domains is weak [3]. In other cases where the domains interact more extensively, cooperative folding proc-

esses of adjacent domains are observed [6, 7]. It is assumed that stronger domain-domain interaction during folding may result in higher rates of misfolding and aggregation. Because of this, refolding of larger multi-domain proteins *in vitro* is inefficient in many cases [8]. However, in particular by employing co-solvent-assisted folding (sugars, osmolytes), artificial chaperone systems (detergent, cyclodextrin), and *in vitro* chaperonin systems (*e.g.*, GroEL/GroES) increased folding efficiencies have been achieved with numerous multi-domain proteins [9, 10]. Even more efficient folding was observed for *de novo* synthesized multi-domain proteins where individual domains are synthesized sequentially and, apparently, also fold sequentially [8]. In contrast to single-domain proteins, for multi-domain proteins characteristic *in vivo* conditions such as the presence of molecular chaperones and co-translational folding seems to play a crucial role for an efficient folding process [11, 12]. Although valuable details about co-translational folding of multi-domain proteins have been gathered in the last few years [11, 13, 14], knowledge about *in vitro* conditions at which proper refolding of multi-domain proteins takes place is still of interest, since these conditions provide the experimental basis for most classical folding studies on multi-domain proteins.

In this review recent results and insights from folding studies on multi-domain proteins are summarized and discussed. In addition, special attention is drawn to some elaborated experimental approaches that appear promising for elucidating new and advancing knowledge about multi-domain protein folding.

Reversible *versus* irreversible unfolding

A prerequisite for classical (re-)folding studies and for identifying critical steps in the protein-folding pathways is the possibility of unfolding the protein of interest reversibly. As mentioned above, refolding is often hindered or slowed down by misfolding and protein aggregation. In addition, proteins undergoing irreversible unfolding in general cannot be used to determine the thermodynamic stability parameter that are typically obtained from equilibrium transitions [15, 16]. Therefore, some effort has been put in finding proper refolding conditions. For numerous multi-domain proteins that exhibit no or very poor refolding under standard *in vitro* conditions, environmental conditions or special treatments have been found that ensure proper or at least improved refolding efficiencies. Typically standard techniques such like intrinsic tryptophan fluorescence and CD spectroscopy are employed to follow unfolding and refolding transitions (see for example [17–19]). In many cases the state of aggregation was also monitored by size-exclusion chromatography, or by static or dynamic light scattering (DLS), first because this information is valuable for characterizing the unfolding process, and second because heavy protein aggregation can lead to artifacts in fluorescence and CD spectra, which have to be considered in the data interpretation [20, 21].

In the following, some established refolding approaches are briefly summarized. (i) Since the strong tendency for aggregation of non-native states is one of the major obstacles in multi-domain protein refolding, special co-solvents (*e.g.*, sugars, osmolytes) have been employed in refolding studies for proteins that notoriously unfold irreversibly. Glycerol is a very efficient co-solvent and has been shown to increase the refolding yield of unfolded multi-domain proteins [10, 22–25]. Often the type of treatment that is used to induce unfolding [elevated or low temperatures, pressure, chemical denaturants like guanidinium hydrochloride (GndHCl) or urea, alkaline or acidic pH] also plays a crucial role in whether and to what extend refolding takes place (see Fig. 1). (ii) In cells, accessory proteins such as molecular chaperones (DnaK, DnaJ, or GroEL/ES) help to achieve proper protein folding [26]. Thus, for *in vitro* experiments, purified molecular chaperones were employed to facilitate improved refolding of multi-domain proteins [27–29]. Inspired by the folding mechanism accomplished by the GroEL/ES complex, a so-called two-step artificial chaperone strategy was established using a refolding system with detergent and cyclodextrin [30–32]. (iii) In addition to the well-established approaches discussed above, in numerous cases protein-specific refolding conditions have been identified for individual multi-domain proteins. For example, phosphoglycerate kinase shows a reversible thermal unfolding at low concentrations of GndHCl, and an irreversible transition in buffers without GndHCl (see [10, 33] and Fig. 1C). Without the help of chaperones, under *in vitro* conditions the firefly luciferase only refolds at low temperatures and then at extremely slow refolding rates [34]. Under most refolding conditions, α-amylase from *Bacillus licheniformis* shows a distinct protein aggregation that most probably hinders refolding ([10, 35–37], see also Fig. 1). At a midpoint urea concentration ($c_{1/2} = 2.4$ M), reversible unfolding and refolding was observed upon the variation of calcium ions, which serve as co-factors for this three-domain protein [38].

Although many theoretical and experimental studies on the folding of multi-domain proteins have been performed, it is still not fully understood why some multi-domain proteins show reversible unfolding and others not. In case studies on tandem-linked domains,

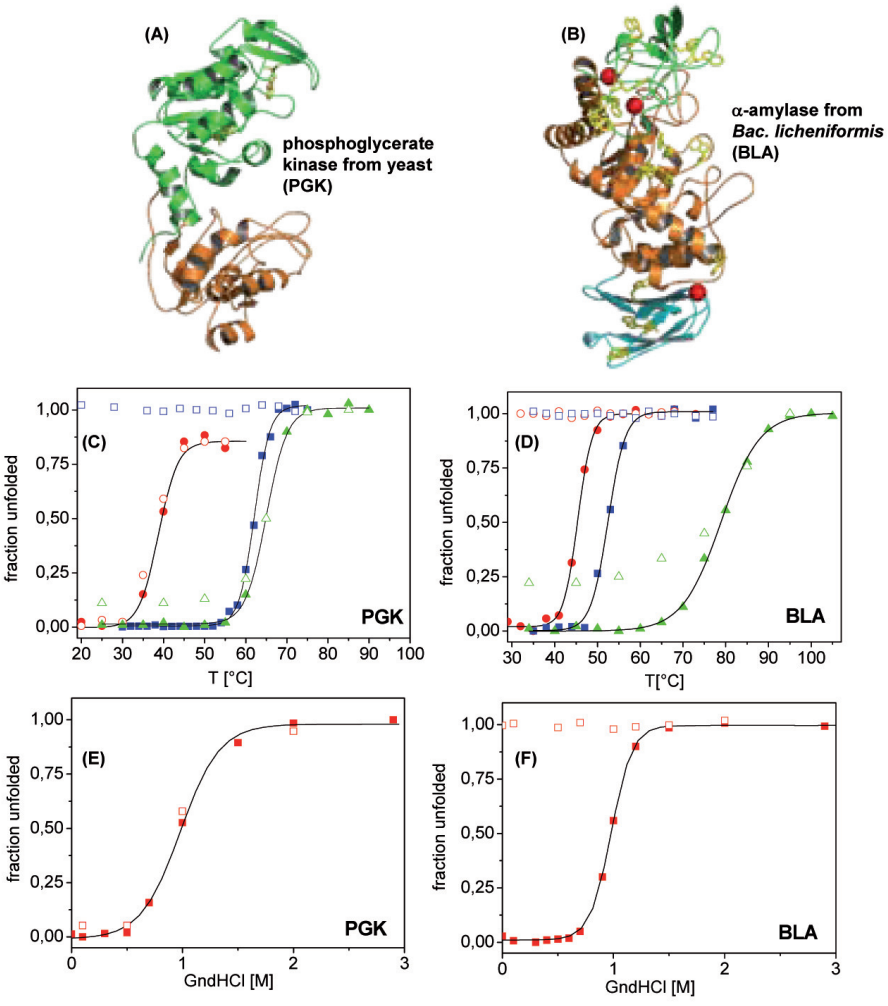

Figure 1. For reversible and irreversible unfolding, typical transitions are shown for different unfolding conditions from two multi-domain proteins. (*A*) Phosphoglycerate kinase (PGK) and (*B*) α-amylase from *B. licheniformis* (BLA). (*C, D*) Thermal unfolding (solid symbols) was measured in different buffers [red symbols, buffer with 0.7 M guanidinium hydrochloride (GndHCl); blue symbols, native buffer; green symbols, buffer with 7 M glycerol] and shows pronounced difference in the refolding (open symbols) behavior. For BLA, only in buffer with 7 M glycerol significant refolding is observed, while for PGK glycerol as a co-solvent, and a buffer with 0.7 M GndHCl ensures proper refolding. (*E, F*) GndHCl-induced unfolding (solid symbols) is reversible (refolding: open symbols) only for PGK. Most probably weaker domain interactions and smaller β-sheet contents in PGK leads to the fact that PGK is characterized by a stronger tendency for reversible unfolding as compared to BLA (see text). The presented data are from [10].

it has been demonstrated that sequence diversity of neighboring domains, properties of the domain connectivity between neighboring domains, and domain interactions in general can have a profound influence on folding routes, on folding kinetics and thereby on the reversibility of unfolding [3, 5, 39]. Another important issue is the role of β-sheet like structures as structural elements of native proteins or sheet structures formed transiently during the folding process [40]. Numerous studies have shown that β-sheet like structures lead to a higher probability of forming random inter-domain and intermolecular H-bonds. They increase the chance of conformational scrambling or of protein aggregation, and in particular the formation of amyloid-like fibrils (see [41] and references therein). Obviously, all these processes strongly determine whether unfolding transitions are reversible or irreversible.

Refolding *versus* co-translational folding

Co-translational (or biosynthetic) protein folding has been studied intensively in the past, in most cases with multi-domain proteins [13, 14, 42–46]. The major difference of co-translational folding with respect to refolding of full-length polypeptide chains is the vectorial appearance of the nascent polypeptide chain, which in principle leads to a subsequent vectorial folding process. In addition, the growing nascent polypeptide chain remains bound to the ribosome during the folding process [43]. Related to these differences, for co-translational folding, a more rapid folding, different folding pathways, and higher yields of folded protein were observed. It is assumed that the differences between co-translational folding and refolding are more relevant for larger multi-domain proteins than for single-domain proteins. An important question is how does co-translational folding ensure more effective and productive folding than refolding from full-length polypeptides [43, 45]. First, during synthesis and protein folding, the elongating

nascent chain remains anchored to the ribosome and has less rotational and translational freedom. This restriction can significantly reduce intermolecular collisions with other (in particular also not folded) proteins and thereby suppress aggregation. Second, in contrast to refolding, folding of N-terminal and C-terminal regions is separated spatially and temporally during sequential synthesis (Fig. 2). Most probably this prevents non-native and unwanted interactions that often leads to off-pathway folding intermediates or aggregates. Sequential folding can be even more effective when the synthesis rate is slowed down at specific critical positions in the polypeptide sequence. For example, pauses after domains or β-sheets would allow proper co-translational structure formation, before subsequent structure elements appear that might cause aggregation during the folding process. It is assumed that codon usage is responsible for different synthesis rates at specific regions within the protein sequence. Third, several chaperones (*e.g.*, trigger factor and DnaK in bacterial systems) act already during translation on the nascent chain and prevent misfolding or aggregation. For example, the trigger factor binds to the ribosome before a chaperone-nascent chain interaction is observed. Although this process is only poorly understood, it is assumed that this type of chaperones is most efficient only in co-translational folding [8]. Finally, all these properties seem to ensure that, in co-translational folding, non-productive off-pathway intermediates and kinetic traps are avoided.

It has been demonstrated unequivocally that co-translational folding takes place for many proteins. Furthermore, it was shown that proteins can become fully folded and enzymatically active while they are bound to the ribosome through a C-terminal extension of about 30 amino acids that spans the ribosomal tunnel [13, 14, 47]. However, relative little is known about structural and dynamical details of the nascent polypeptide chain at specific steps during co-translational folding. To date, structural and dynamical information on proteins or parts of proteins that remain linked to the ribosome is mainly provided by fluorescence and NMR spectroscopy. For example, Johnson and coworkers [48] showed that for specific polypeptide sequences helix formation can take place already inside the ribosomal tunnel. More recently, structural and dynamical properties were investigated at different stages of the polypeptide elongation for a single-domain protein (apomyglobin), where release from the ribosome was inhibited [49]. In this study using dynamic fluorescence depolarization, different independent nascent chain motions were detected that reveal information on the folding status (*e.g.*, protein compactness) and on interactions with the

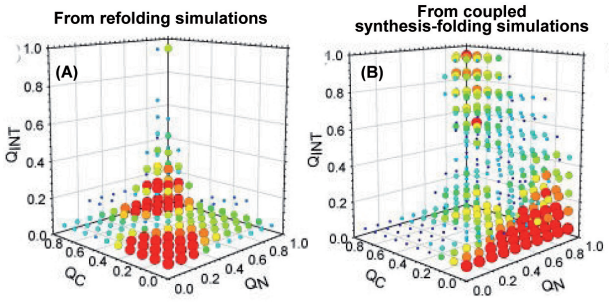

Figure 2. For the two-domain module of the Semliki forest virus protein (SFVP), (*A*) refolding of full-length proteins (149 residues) and (*B*) synthesis coupled (co-translational) folding of the protein are presented. Elcock [44] analyzed the folding of this protein by molecular simulation methods that yield folding coordinates for the N-terminal domain (Q_N), for the C-terminal domain (Q_C), and for the interface region between both domains (Q_{INT}). In both graphs it is shown how the values of Q_N, Q_C, Q_{INT} evolve during folding ("Q" is a general folding coordinate defining the fraction of native residue-residue contacts that are present for given conformations with values starting from zero for the unfolded state to a value of one for the native structure). The colored spheres represent the population points in the conformational space as determined during 100 folding trajectories (large red symbols represent highly populated states; ~100 %). For refolding of the full-length protein (*A*), native contacts for the N-terminal domain and for the C-terminal domain are formed in parallel. After the structure formation of the domains is completed, native interface contacts are formed. For synthesis-coupled folding (*B*) a different picture emerged. Here, first native contacts in the N-terminal domain are formed. After this, native contacts of the C-terminal domain and native interface contacts between both domains are formed in a concerted manner. Adapted with permission from [44].

ribosome. In another approach using NMR spectroscopy, a tandem immunoglobulin domain repeat (Ig2) was investigated [50]. With the construct used in that study, the N-terminal domain was found to be natively folded, while the C-terminal domain remained largely unfolded. In this particular case, the C-terminal domain adopted an unfolded state most probably because a terminal β-strand was missing in the construct. A more elucidating and straightforward approach would be to measure directly trajectories of the nascent chain during co-translational folding in time-resolved studies. Unfortunately, protein biosynthesis is a rather asynchronous process, which limits the sensitivity and the resolution of any kind of measurement in an ensemble of synthesizing ribosomes, either in the cell or in cell free systems. Obviously, single-molecule techniques would be the method of choice for this kind of studies, but studying co-translational folding at single-molecule level is still in its infancy [51, 52].

Although folding on the ribosome in general represents a natural method of protein folding in the cell, post-translational folding (or refolding) is a natural process as well. Protein unfolding and refolding also takes place in the cell completely decoupled from

biosynthesis. In addition, the fundamental question of the equilibration between a specific sequence of an unfolded polypeptide chain with the corresponding native three-dimensional protein structure is not fully answered yet. In any case it is valuable to study protein folding in refolding experiments with full-length polypeptide chains, even if folding routes between refolding and co-translational folding differ [53].

Working at low protein concentrations and the perspective of single-molecule studies

Essentially two major techniques have been established in the last decade to study protein folding on the single-molecule level, dynamic force spectroscopy (*e.g.*, atomic force microscopy, optical tweezers) and fluorescence-based methods [54, 55]. In the case of force spectroscopy, numerous studies on multi-domain repeat structures have been performed. In these experiments proteins are attached at both ends to surfaces and an external force is applied by increasing the distance between the two tethered ends. Unfolding and refolding is monitored by following the applied force as a function of the polypeptide chain length. For example, the unfolding of individual domains was identified in multi-domain repeat structures (see [56] and references in [54, 55]). With the advent of bright and more photostable fluorescent dyes and with enormous methodical and technical improvements, the field of high-resolution fluorescence spectroscopy and microscopy has made tremendous advances. In particular, single-molecule fluorescence studies have become a key technique, not only in protein folding, but also in molecular and structural biology [57–59].

With respect to our subject, multi-domain protein folding, this has essentially two important implications. First, extremely low protein concentrations (around nanomolar), as used for example in fluorescence correlation spectroscopy (FCS) studies, can lower or even circumvent aggregation of unfolded states. In principle, this can make refolding studies feasible, which would be hampered by aggregation at higher protein concentrations. Second, the application of single-molecule techniques has an enormous potential for studying the intrinsically heterogeneous process of protein folding, not only for multi-domain proteins [57, 58].

Since conventional techniques like intrinsic tryptophan fluorescence or CD spectroscopy give signals that are too weak at extremely low protein concentrations, highly sensitive fluorescent dyes need to be attached to the protein of interest. For protein folding studies, FCS and energy transfer techniques [Förster resonance energy transfer (FRET) and photoinduced energy transfer (PET)] have been mainly employed [57, 58, 60]. FCS is an easily applicable technique that needs only one fluorescent dye to be attached to the protein of interest. The hydrodynamic radius (R_h) of the protein can be determined from the diffusion time of the fluorescently labeled protein through the confocal volume of a tightly focused laser beam [61–65]. Typically, the native state is compact and exhibits a smaller hydrodynamic radius as compared to the more expanded unfolded state [66–68]. A similar behavior is observed for PGK, which shows an increase in hydrodynamics radius from 3.2 nm in the native state to 5.9 nm at 2 M GndHCl in the unfolded state (Fig. 3A). The radius of a refolded PGK reaches the value of the native state, which indicates a high degree of refolding. Another multi-domain protein (BLA) exhibits an increase in hydrodynamic radius upon unfolding at 4 M GndHCl (with R_h from 3.1 to 5.8 nm), but does not show a complete refolding. However, the R_h value obtained under refolding conditions (~ 3.6 nm) indicates that at least a partial refolding takes place at the extremely low protein concentration (~1 nM) employed for FCS measurements (Fig. 3B). This observation is in contrast to measurements at higher protein concentrations (~1–10 μM) at which GndHCl-induced unfolding of BLA is strictly irreversible (see Fig. 1F, and [10]). For some other α-amylases we observed much larger R_h values under refolding conditions, which is indicative for extended aggregation, even at these rather low protein concentration (Rosenkranz and Fitter, unpublished results). Recently, a similar observation was presented in a FCS study with a two-domain protein (rhodanese) under refolding conditions [69]. However, even for those proteins that aggregate in their non-native state at low concentration in solution, encapsulation techniques can help to avoid aggregation [70, 71]. If individual proteins are confined in a void or in a capsule and thereby separated from each other, aggregation in particular of unfolded proteins cannot take place. Encapsulation of single proteins into sol-gel polymer matrices, as well as into lipid or polymeric vesicles has already been performed [72–74]. In this respect it is important to minimize the interaction of the protein with the confining material. Unfortunately, for such encapsulated proteins, the state of folding or unfolding cannot by analyzed with FCS, since these proteins no longer diffuse freely in solution, which is a prerequisite for FSC. However, other single-molecule fluorescence-based approaches can be employed in this case (see below).

Another advantage of FCS (but also for other fluorescence-based single-molecule techniques) is that not only highly dilute proteins can be studied,

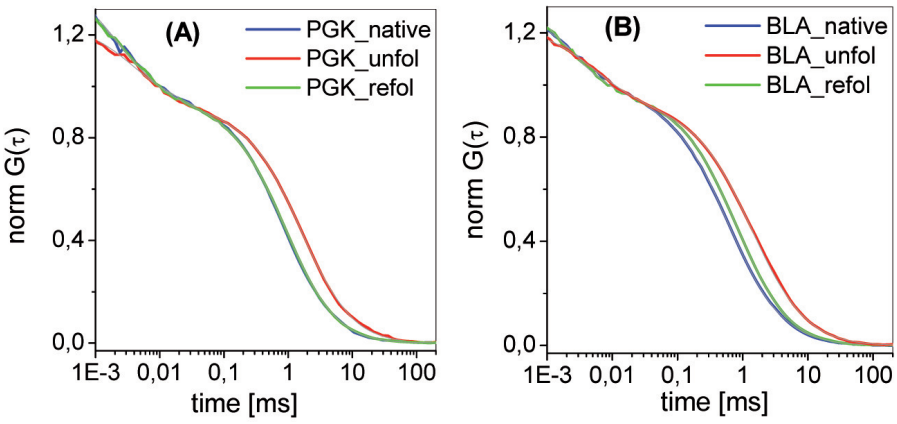

Figure 3. Normalized autocorrelation curves as obtained from FCS measurements using a MicroTime 200 from PicoQuant, Berlin (Rosenkranz and Fitter, unpublished results). Both proteins PGK (*A*) and BLA (*B*) were labeled with Alexa647 and were measured at protein concentrations in the nanomolar range for three different folding states [native state, blue lines; unfolded state (2 M GndHCl for PGK and 4 M GndHCl for BLA), red lines; refolded state (native buffer conditions), green lines]. For both proteins, the hydrodynamic radius increases upon unfolding as seen by curves shifted to longer times. In the case of PGK, the refolded state reaches a hydrodynamic radius that was observed already for the native state, which indicates a complete refolding. BLA exhibits only a partial refolding since the curve for the refolded state is not fully back shifted.

but also proteins at extremely high concentrations or in crowding media [75–77]. Only the fraction of labeled proteins must be at low concentration to provide a sufficiently strong signal fluctuation caused by translational protein diffusion. Here the impact of molecular crowding on protein stability and on protein folding would be an interesting task to analyze [78].

Using single-molecule fluorescence techniques, the greatest amount of information on unfolded and native states of proteins has been obtained from FRET measurements. FRET between two fluorophores, the donor and the acceptor, occurs when both are at a distance of a few nanometer. This energy transfer is due to non-radiative Coulombic dipole-dipole interactions and has proven to be a powerful spectroscopic technique for measuring distances in the range 2–8 nm [79, 80]. Due to the fact that two fluorescent dyes are attached at defined positions in the protein structure, larger structural changes that occur upon protein folding/unfolding are monitored by measuring the energy transfer efficiency. Such FRET measurements yield not only structural information about the compactness of individual proteins at defined equilibrium conditions (*e.g.*, specific concentrations of chemical denaturant). They also provide valuable information on different subpopulations, for example the folded or unfolded state, which are present under the given conditions [57, 58, 81, 82]. Until now, studies have been performed mainly with small single-domain proteins that behave as two-state folders. Only very recently have single-molecule FRET studies been performed with larger two-domain proteins [83, 84]. In typical single molecule studies using confocal microscopy, a tightly focused laser beam in combination with confocal detection defines a small volume out of which bursts of photons are collected from individual traversing proteins. In this set-up, the observation time for a single protein is limited by the translational diffusion time of the molecule through the detection volume. However, a possibility of extended observation times (>milliseconds) would allow rare or slow dynamic events as well as repetitive processes with one and the same molecule to be monitored. This is of particular interest for larger multi-domain proteins, which often exhibit more complex and thereby slower folding or unfolding transitions [85]. To achieve longer observation times, proteins need to be immobilized. One requirement for almost all single molecules studies with immobilized proteins is to reduce, as much as possible, interactions of the protein with the surfaces of the enclosing cavity used for immobilization. In protein folding, perturbing interactions of proteins with surfaces may significantly alter the polypeptide structure and dynamics and can give rise to artifacts in the obtained results [58]. Several studies have demonstrated successful confinement of proteins inside agarose gel matrices [71, 86] or within surface-tethered vesicles [74, 87, 88] without altering their structure and dynamics. Even anchoring proteins directly to polymer-coated surface with a specific single-point attachment has given satisfactory results [89].

For individual proteins encapsulated in surface-tethered lipid vesicles kept at thermodynamic midtransition points (*i.e.*, $GndHCl_{1/2}$), multiple successive unfolding/refolding transitions were observed with FRET [87, 90]. For a small protein (cold shock protein

CspTm, single-domain protein, 67 residues) the transitions were too fast to be resolved, even with the shortest integration time (~100 μs). Therefore, transitions appear in FRET trajectories as single steps for complete folding or unfolding processes. In contrast, a larger protein (adenylated kinase, two-domain protein, 214 residues) exhibits transitions on different time scales. In particular, slower transitions (>1 s) show small and partial steps during individual unfolding and refolding transitions (Fig. 4). This behavior seems to represent a higher complexity of the folding/unfolding process and is indicative for populated intermediate states and for a higher roughness of the funnel-shaped folding energy landscape [87]. Obviously, single-molecule FRET studies with surface-immobilized proteins offer a large potential, in particular for continuing studies on other, even larger multi-domain proteins. Typically, the folding pathways of these proteins are populated by intermediate states and transitions are slow enough that corresponding FRET trajectories can disclose information on folding and unfolding events. In particular, the following subjects are worth more detailed study: (i) In the past detailed single-molecule FRET studies have been performed only with two-state folders. It would be interesting to investigate proteins that exhibit observable intermediate states and to answer the question whether the intermediate states are productive on-pathway conformations or off-pathway traps. In principle, FRET trajectories from single-molecule measurements can give an unambiguous answer to this question, provided the intermediate state is spectroscopically distinguishable from the native and the unfolded state and all investigated states live long enough to be characterized by a sufficient number of emitted photons [58]. However, even for faster interconverting species or for more complex reaction schemes suitable approaches have been developed to obtain valuable information on dynamical processes from FRET data [91–94]. (ii) Another important advance would be to follow folding and unfolding transitions of individual domains as part of a multi-domain protein. Such a goal can be achieved in successive measurements with alternating positions of donor and/or acceptor labeling in the protein structure. In this case individual samples with a specific pair of label positions can yield information on folding transitions of individual domains and of the whole protein, respectively. A more elegant way would be the use of a three-color FRET (or higher order FRET) scheme, which in principle can give information on the folding of two individual domains simultaneously in one measurement (see for example [95]). Using multi-color FRET allows not only one distance but two or even more distances within the protein structure to be measured at the same time [96, 97]. With this additional information, folding of two separated domains can be monitored during one measurement. (iii) The impact of environmental conditions on folding pathways and on kinetics is a further subject of interest. A promising step in this direction is given by recently published work on protein folding in the presence of molecular chaperones during refolding [83, 84]. One of these studies indicates that only a certain subpopulation of the unfolded state ensemble, a partially structured folding intermediate, is bound to the chaperone GroEL. Finally, this subpopulation is released from the chaperone in a native state-like structure upon addition of ATP [83]. In similar studies the effect of artificial chaperones or the presence of crowding media can by studied. The goal would be to obtain information on how protein folding is influenced by pronounced intermolecular interactions or caging effects. It is assumed that such effects are present in the cellular interior, since 20–30 % of the total volume is occupied by macromolecules [78, 98].

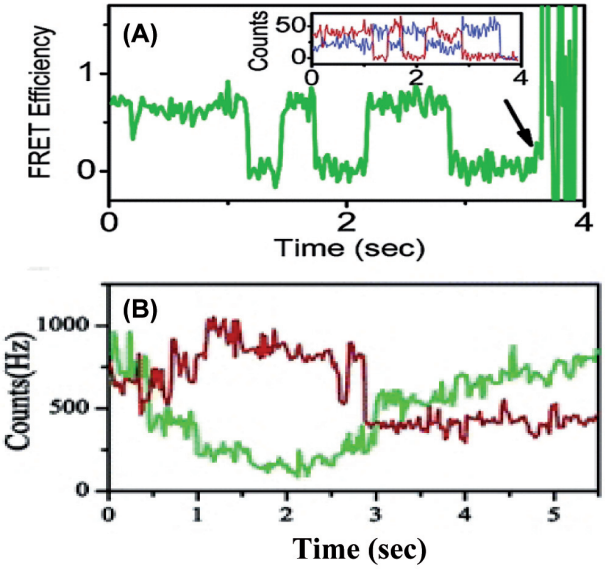

Figure 4. Typical FRET trajectories as observed for two different proteins (CspTm, adenylated kinase) showing unfolding and refolding transitions at their thermodynamic GndHCl midpoint concentrations. (*A*) CspTm shows fast transitions with single steps between folded and unfolded states (green line). The inset displays the corresponding donor signals (blue) and acceptor signals (red). (*B*) For the larger adenylate kinase, the donor signal (green) and the acceptor signal (red) show first a relative slow multi-step folding process (ending at ~2 s) and subsequently a fast unfolding process (at ~3 s). With permission from [87, 90].

Summary

Due to their abundant occurrence in the cell, multi-domain proteins and the folding of these proteins is an important issue. Although in the future it will still be possible, in principle, to answer many fundamental questions on protein folding through studies on smaller single-domain proteins, it is worth studying the folding of multi-domain proteins in more detail. Studies comparing classical refolding and co-translational folding of multi-domain proteins and the application of recently developed single-molecule techniques can contribute significantly to improve our knowledge on this topic. In particular, single-molecule fluorescence methods have been proven to be a versatile tool to study protein folding. As outlined in this review, this technique has reached a state of development at which it can now be applied to more complex reaction schemes that are assumed to occur in the folding process of multi-domain proteins.

Acknowledgements. The author thanks T. Rosenkranz for help with sample preparation and J. Enderlein and I. Gregor for many helpful and inspiring discussions. G. Büldt is acknowledged for comments on the manuscript and for continuous support in his institute.

1 Orengo, C. A. and Thornton, J. M. (2005) Protein families and their evolution – A structural perspective. Annu. Rev. Biochem. 74, 867–900
2 Jaenicke, R. (1999) Stability and folding of domain proteins. Prog. Biophys. Mol. Biol. 71, 155–241
3 Han, J. H., Batey, S., Nickson, A. A., Teichmann, S. A. and Clarke, J. (2007) The folding and evolution of multidomain proteins. Nat. Rev. Mol. Cell Biol. 8, 319–330
4 Lilie, H. and Seckler, R. (2005) Folding and association of multi-domain and oligomeric proteins. In: Protein Folding Handbook, pp. 32–72, Buchner, J. and Kiefhaber, T. (eds.), Wiley-VCH, Weinheim
5 Raman, E. P., Barsegov, V. and Klimov, D. K. (2007) Folding of tandem-linked domains. Proteins 67, 795–810
6 Privalov, P. L. (1996) Intermediate states in protein folding. J. Mol. Biol. 258, 707–725
7 Sudharshan, E. and Rao, A. G. (1999) Involvement of cysteine residues and domain interactions in the reversible unfolding of lipoxygenase-1. J. Biol. Chem. 274, 35351–35358
8 Frydman, J. (2001) Folding of newly translated proteins *in vivo:* The role of molecular chaperones. Annu. Rev. Biochem. 70, 603–647
9 Lange, C. and Rudolph, R. (2005) Production of recombinant proteins for therapy, diagnostic, and industrial research by *in vitro* folding. In: Protein Folding Handbook, pp. 1245–1280, Buchner, J. and Kiefhaber, T. (eds.), Wiley-VCH, Weinheim
10 Strucksberg, K. H., Rosenkranz, T. and Fitter, J. (2007) Reversible and irreversible unfolding of multi-domain proteins. Biochim. Biophys. Acta 1774, 1591–1603
11 Frydman, J., Erdjument-Bromage, H., Tempst, P. and Hartl, F. U. (1999) Co-translational domain folding as the structural basis for the rapid *de novo* folding of firefly luciferase. Nat. Struct. Biol. 6, 697–705
12 Fersht, A. R. and Daggett, V. (2002) Protein folding and unfolding at atomic resolution. Cell 108, 573–582
13 Kolb, V. A., Makeyev, E. V. and Spirin, A. S. (1994) Folding of firefly luciferase during translation in a cell-free system. EMBO J. 13, 3631–3637
14 Makeyev, E. V., Kolb, V. A. and Spirin, A. S. (1996) Enzymatic activity of the ribosome-bound nascent polypeptide. FEBS Lett. 378, 166–170
15 Galisteo, M. L., Mateo, P. L. and Sanchez-Ruiz, J. M. (1991) Kinetic study on the irreversible thermal denaturation of yeast phosphoglycerate kinase. Biochemistry 30, 2061–2066
16 Duy, C. and Fitter, J. (2005) Thermostability of irreversible unfolding alpha-amylases analyzed by unfolding kinetics. J. Biol. Chem. 280, 37360–37365
17 Kelly, S. M. and Price, N. C. (2000) The use of circular dichroism in the investigation of protein structure and function. Curr. Protein Pept. Sci. 1, 349–384
18 Royer, C. A. (2006) Probing protein folding and conformational transitions with fluorescence. Chem. Rev. 106, 1769–1784
19 Buchner, J. and Kiefhaber, T. (2005) Protein Folding Handbook, Vols. 1–5. Wiley-VCH, Weinheim
20 Duy, C. and Fitter, J. (2006) How aggregation and conformational scrambling of unfolded states govern fluorescence emission spectra. Biophys. J. 90, 3704–3711
21 Bustamante, C., Tinoco, I. and Maestre, M. F. (1983) Circular differential scattering can be an important part of the circular dichroism of macromolecules. Proc. Natl. Acad. Sci. USA 80, 3568–3572
22 Gekko, K. and Timasheff, S. N. (1981) Mechanism of protein stabilization by glycerol: Preferential hydration in glycerol-water mixtures. Biochemistry 20, 4667–4676
23 Gorovits, B. M. and Horowitz, P. M. (1998) High hydrostatic pressure can reverse aggregation of protein folding intermediates and facilitate acquisition of native structure. Biochemistry 37, 6132–6135
24 Panda, M., Gorovits, B. M. and Horowitz, P. M. (2000) Productive and nonproductive intermediates in the folding of denatured rhodanese. J. Biol. Chem. 275, 63–70
25 Mishra, R., Seckler, R. and Bhat, R. (2005) Efficient refolding of aggregation-prone citrate synthase by polyol osmolytes: How well are protein folding and stability aspects coupled? J. Biol. Chem. 280, 15553–15560
26 Hartl, F. U. and Hayer-Hartl, M. (2002) Molecular chaperones in the cytosol: From nascent chain to folded protein. Science 295, 1852–1858
27 Schroder, H., Langer, T., Hartl, F. U. and Bukau, B. (1993) DnaK, DnaJ and GrpE form a cellular chaperone machinery capable of repairing heat-induced protein damage. EMBO J. 12, 4137–4144
28 Kawata, Y., Hongo, K., Mizobata, T. and Nagai, J. (1998) Chaperonin GroE-facilitated refolding of disulfide-bonded and reduced Taka-amylase A from *Aspergillus oryzae.* Protein Eng. 11, 1293–1298
29 Weissman, J. S., Rye, H. S., Fenton, W. A., Beechem, J. M. and Horwich, A. L. (1996) Characterization of the active intermediate of a GroEL-GroES-mediated protein folding reaction. Cell 84, 481–490
30 Rozema, D. and Gellman, S. H. (1995) Artificial chaperones – Protein refolding *via* sequential use of detergent and cyclodextrin. J. Am. Chem. Soc. 117, 2373–2374
31 Daugherty, D. L., Rozema, D., Hanson, P. E. and Gellman, S. H. (1998) Artificial chaperone-assisted refolding of citrate synthase. J. Biol. Chem. 273, 33961–33971
32 Yazdanparast, R., Khodagholi, F. and Khodarahmi, R. (2005) Artificial chaperone-assisted refolding of chemically denatured alpha-amylase. Int. J Biol. Macromol. 35, 257–263
33 Nojima, H., Ikai, A., Oshima, T. and Noda, H. (1977) Reversible thermal unfolding of thermostable phosphoglycerate kinase. Thermostability associated with mean zero enthalpy change. J. Mol. Biol. 116, 429–442
34 Herbst, R., Schafer, U. and Seckler, R. (1997) Equilibrium intermediates in the reversible unfolding of firefly (*Photinus pyralis*) luciferase. J. Biol. Chem. 272, 7099–7105
35 Tomazic, S. J. and Klibanov, A. M. (1988) Mechanisms of irreversible thermal inactivation of *Bacillus* alpha-amylases. J. Biol. Chem. 263, 3086–3091

36 Feller, G., d'Amico, D. and Gerday, C. (1999) Thermodynamic stability of a cold-active alpha-amylase from the Antarctic bacterium *Alteromonas haloplanctis*. Biochemistry, 38 4613–4619
37 Fitter, J. (2005) Structural and dynamical features contributing to thermostability in alpha-amylases. Cell. Mol. Life Sci. 62, 1925–1937
38 Nazmi, A. R., Remisch, T. and Hinz, H. J. (2006) Ca-binding to *Bacillus licheniformis* alpha-amylase (BLA). Arch. Biochem. Biophys. 453, 18–25
39 Wright, C. F., Teichmann, S. A., Clarke, J. and Dobson, C. M. (2005) The importance of sequence diversity in the aggregation and evolution of proteins. Nature 438, 878–881
40 Yang, W. Y., Larios, E. and Gruebele, M. (2003) On the extended beta-conformation propensity of polypeptides at high temperature. J. Am. Chem. Soc. 125, 16220–16227
41 Chiti, F. and Dobson, C. M. (2006) Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem. 75, 333–366
42 Kramer, G., Ramachandiran, V. and Hardesty, B. (2001) Cotranslational folding – omnia mea mecum porto? Int. J. Biochem Cell Biol. 33, 541–553
43 Fedorov, A. N. and Baldwin, T. O. (1997) Cotranslational protein folding. J. Biol. Chem. 272, 32715–32718
44 Elcock, A. H. (2006) Molecular simulations of cotranslational protein folding: Fragment stabilities, folding cooperativity, and trapping in the ribosome. PLoS Comput. Biol. 2, e98
45 Evans, M. S., Clarke, T. F. and Clark, P. L. (2005) Conformations of co-translational folding intermediates. Protein Pept. Lett. 12, 189–195
46 Fedorov, A. N. and Baldwin, T. O. (1999) Process of biosynthetic protein folding determines the rapid formation of native structure. J. Mol. Biol. 294, 579–586
47 Kudlicki, W., Chirgwin, J., Kramer, G. and Hardesty, B. (1995) Folding of an enzyme into an active conformation while bound as peptidyl-tRNA to the ribosome. Biochemistry 34, 14284–14287
48 Woolhead, C. A., McCormick, P. J. and Johnson, A. E. (2004) Nascent membrane and secretory proteins differ in FRET-detected folding far inside the ribosome and in their exposure to ribosomal proteins. Cell 116, 725–736
49 Ellis, J. P., Bakke, C. K., Kirchdoerfer, R. N., Jungbauer, L. M. and Cavagnero S (2008) Chain dynamics of nascent polypeptides emerging from the ribosome. ACS Chem. Biol. 3, 555–566
50 Hsu, S. T., Fucini, P., Cabrita, L. D., Launay, H., Dobson, C. M. and Christodoulou, J. (2007) Structure and dynamics of a ribosome-bound nascent chain by NMR spectroscopy. Proc. Natl. Acad. Sci. USA 104, 16516–16521
51 Uemura, S., Iizuka, R., Ueno, T., Shimizu, Y., Taguchi, H., Ueda, T., Puglisi, J. D. and Funatsu, T. (2008) Single-molecule imaging of full protein synthesis by immobilized ribosomes. Nucleic Acids Res. 36, e70
52 Katranidis, A., Atta, D., Schlesinger, R., Nierhaus, K. H., Choli-Papadopoulou, T., Gregor, I., Gerrits, M., Büldt, G. and Fitter, J. (2009) Fast biosynthesis of GFP molecules – A single molecule fluorescence study. Angew. Chem. Int. Ed. Engl. (in press), doi:10.1002/anie.200806070
53 Haas, E. (2008) Folding on the assembly line. ACS Chem. Biol. 3, 527–529
54 Zhuang, X. and Rief, M. (2003) Single-molecule folding. Curr. Opin. Struct. Biol. 13, 88–97
55 Borgia, A., Williams, P. M. and Clarke, J. (2008) Single-molecule studies of protein folding. Annu. Rev. Biochem. 77, 101–125
56 Rief, M., Gautel, M., Oesterhelt, F., Fernandez, J. M. and Gaub, H. E. (1997) Reversible unfolding of individual titin immunoglobulin domains by AFM. Science 276, 1109–1112
57 Schuler, B. (2005) Single-molecule fluorescence spectroscopy of protein folding. Chemphyschem 6, 1206–1220
58 Michalet, X., Weiss, S. and Jager, M. (2006) Single-molecule fluorescence studies of protein folding and conformational dynamics. Chem. Rev. 106, 1785–1813
59 Joo, C., Balci, H., Ishitsuka, Y., Buranachai, C. and Ha, T. (2008) Advances in single-molecule fluorescence methods for molecular biology. Annu. Rev. Biochem. 77, 51–76
60 Neuweiler, H., Doose, S. and Sauer, M. (2005) A microscopic view of miniprotein folding: Enhanced folding efficiency through formation of an intermediate. Proc. Natl. Acad. Sci. USA 102, 16650–16655
61 Magde, D., Webb, W. W. and Elson, E. (1972) Thermodynamic fluctuations in a reacting system – Measurement by fluorescence correlation spectroscopy. Phys. Rev. Lett. 29, 705–708
62 Eigen, M. and Rigler, R. (1994) Sorting single molecules: Application to diagnostics and evolutionary biotechnology. Proc. Natl. Acad. Sci. USA 91, 5740–5747
63 Maiti, S., Haupts, U. and Webb, W. W. (1997) Fluorescence correlation spectroscopy: Diagnostics for sparse molecules. Proc. Natl. Acad. Sci. USA 94, 11753–11757
64 Haustein, E. and Schwille, P. (2003) Ultrasensitive investigations of biological systems by fluorescence correlation spectroscopy. Methods 29, 153–166
65 Dertinger, T., Pacheco, V., der Hocht, I., Hartmann, R., Gregor, I. and Enderlein, J. (2007) Two-focus fluorescence correlation spectroscopy: A new tool for accurate and absolute diffusion measurements. Chemphyschem 8, 433–443
66 Chattopadhyay, K., Saffarian, S., Elson, E. L. and Frieden, C. (2005) Measuring unfolding of proteins in the presence of denaturant using fluorescence correlation spectroscopy. Biophys. J. 88, 1413–1422
67 Sherman, E. and Haran, G. (2006) Coil-globule transition in the denatured state of a small protein. Proc. Natl. Acad. Sci. USA 103, 11539–11543
68 Sherman, E., Itkin, A., Kuttner, Y. Y., Rhoades, E., Amir, D., Haas, E. and Haran, G. (2008) Using fluorescence correlation spectroscopy to study conformational changes in denatured proteins. Biophys. J. 94, 4819–4827
69 Hillger, F., Nettels, D., Dorsch, S. and Schuler, B. (2007) Detection and analysis of protein aggregation with confocal single molecule fluorescence spectroscopy. J. Fluoresc. 17, 759–765
70 Shin, I., Wachtel, E., Roth, E., Bon, C., Silman, I. and Weiner, L. (2002) Thermal denaturation of *Bungarus fasciatus* acetylcholinesterase: Is aggregation a driving force in protein unfolding? Protein Sci. 11, 2022–2032
71 Eggers, D. K. and Valentine, J. S. (2001) Crowding and hydration effects on protein conformation: A study with sol-gel encapsulated proteins. J. Mol. Biol. 314, 911–922
72 Cannone, F., Bologna, S., Campanini, B., Diaspro, A., Bettati, S., Mozzarelli, A. and Chirico, G. (2005) Tracking unfolding and refolding of single GFPmut2 molecules. Biophys. J. 89, 2033–2045
73 Boukobza, E., Sonnenfeld, A. and Haran, G. (2001) Immobilization in surface-tethered lipid vesicles as a new tool for single biomolecule spectroscopy. J. Phys. Chem. B 105, 12165–12170
74 Rosenkranz, T., Katranidis, A., Atta, D., Gregor, I., Enderlein, J., Grzelakowski, M., Riegler, P., Meier, W. and Fitter, J. (2009) Observing proteins as single molecules encapsulated in surface-tethered polymeric nanocontainers. Chembiochem (in press), doi:10.1002/cbic.200800739
75 Dauty, E. and Verkman, A. S. (2004) Molecular crowding reduces to a similar extent the diffusion of small solutes and macromolecules: Measurement by fluorescence correlation spectroscopy. J. Mol. Recogn. 17, 441–447
76 Zorrilla, S., Hink, M. A., Visser, A. J. and Lillo, M. P. (2007) Translational and rotational motions of proteins in a protein crowded environment. Biophys. Chem. 125, 298–305
77 Kozer, N., Kuttner, Y. Y., Haran, G. and Schreiber, G. (2007) Protein-protein association in polymer solutions: From dilute to semidilute to concentrated. Biophys. J. 92, 2139–2149
78 Ellis, R. J. (2001) Macromolecular crowding: Obvious but underappreciated. Trends Biochem. Sci. 26, 597–604

79 Förster, T. (1948) Zwischenmolekulare Energiewanderung und Fluoreszenz. Ann. Phys. 2, 55–75
80 Stryer, L. and Haugland, R. P. (1967) Energy transfer: A spectroscopic ruler. Proc. Natl. Acad. Sci. USA 58, 719–726
81 Schuler, B. and Eaton, W. A. (2008) Protein folding studied by single-molecule FRET. Curr. Opin. Struct. Biol. 18, 16–26
82 Haran, G. (2003) Single-molecule fluorescence spectroscopy of biomolecular folding. J. Phys. Condens. Matter 15, R1291-R1317
83 Hillger, F., Hanni, D., Nettels, D., Geister, S., Grandin, M., Textor, M. and Schuler, B. (2008) Probing protein-chaperone interactions with single-molecule fluorescence spectroscopy. Angew. Chem. Int. Ed. Engl. 47, 6184–6188
84 Sharma, S., Chakraborty, K., Muller, B. K., Astola, N., Tang, Y. C., Lamb, D. C., Hayer-Hartl. M. and Hartl, F. U. (2008) Monitoring protein conformation along the pathway of chaperonin-assisted folding. Cell 133, 142–153
85 Dill, K. A., Ozkan, S. B., Shell, M. S. and Weikl, T. R. (2008) The protein folding problem. Annu. Rev. Biophys. 37, 289–316
86 Lu, H. P., Xun, L. and Xie, X. S. (1998) Single-molecule enzymatic dynamics. Science 282, 1877–1882
87 Rhoades, E., Gussakovsky, E. and Haran, G. (2003) Watching proteins fold one molecule at a time. Proc. Natl. Acad. Sci. USA 100, 3197–3202
88 Okumus, B., Wilson, T. J., Lilley, D. M. and Ha, T. (2004) Vesicle encapsulation studies reveal that single molecule ribozyme heterogeneities are intrinsic. Biophys. J. 87, 2798–2806
89 Kuzmenkina, E. V., Heyes, C. D. and Nienhaus, G. U. (2005) Single-molecule Forster resonance energy transfer study of protein dynamics under denaturing conditions. Proc. Natl. Acad. Sci. USA 102, 15471–15476
90 Rhoades, E., Cohen, M., Schuler, B. and Haran, G. (2004) Two-state folding observed in individual protein molecules. J. Am. Chem. Soc. 126, 14686–14687
91 Schroder, G. F. and Grubmuller, H. (2003) Maximum likelihood trajectories from single molecule fluorescence resonance energy transfer experiments. J. Chem. Phys. 119, 9920–9924
92 Antonik, M., Felekyan, S., Gaiduk, A. and Seidel, C. A. M. (2006) Separating structural heterogeneities from stochastic variations in fluorescence resonance energy transfer distributions *via* photon distribution analysis. J. Phys. Chem. B 110, 6970–6978
93 McKinney, S. A., Joo, C. and Ha, T. (2006) Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophys. J. 91, 1941–1951
94 Gopich, I. V. and Szabo, A. (2007) Single-molecule FRET with diffusion and conformational dynamics. J. Phys. Chem. B 111, 12925–12932
95 Roy, R., Hohng, S. and Ha, T. (2008) A practical guide to single-molecule FRET. Nat. Methods 5, 507–516
96 Ross, J., Buschkamp, P., Fetting, D., Donnermeyer, A., Roth, C. M., and Tinnefeld, P. (2007) Multicolor single-molecule spectroscopy with alternating laser excitation for the investigation of interactions and dynamics. J. Phys. Chem. B 111, 321–326
97 Lee, N. K., Kapanidis, A. N., Koh, H. R., Korlann, Y., Ho, S. O., Kim, Y., Gassman, N., Kim, S. K. and Weiss, S. (2007) Three-color alternating-laser excitation of single molecules: Monitoring multiple interactions and distances. Biophys. J. 92, 303–312
98 Martin, J. (2004) Chaperonin function- effects of crowding and confinement. J. Mol. Recogn. 17, 465–472