the FEBS Journal

A knot in the protein structure – probing the near-infrared fluorescent protein iRFP designed from a bacterial phytochrome

Olesya V. Stepanenko[1], Grigory S. Bublikov[1], Olga V. Stepanenko[1], Daria M. Shcherbakova[2], Vladislav V. Verkhusha[2,3], Konstantin K. Turoverov[1,4] and Irina M. Kuznetsova[1,4]

1 Laboratory of Structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, St Petersburg, Russia
2 Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY, USA
3 Department of Biochemistry and Developmental Biology, Institute of Biomedicine, University of Helsinki, Finland
4 St Petersburg State Polytechnical University, Russia

Keywords
bacteriophytochrome; biliverdin; iRFP; knot; RpBphP2

Correspondence
V. V. Verkhusha, Department of Anatomy and Structural Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA
Fax: +1 718 430 8996
Tel: +1 718 430 8591
E-mail: vladislav.verkhusha@einstein.yu.edu
K. K. Turoverov, Laboratory of Structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, St Petersburg 194064, Russia
Fax: +7 812 2970341
Tel: +7 812 2971957
E-mail: kkt@incras.ru
I. M. Kuznetsova, Laboratory of Structural Dynamics, Stability and Folding of Proteins, Institute of Cytology, Russian Academy of Sciences, St Petersburg 194064, Russia
Fax: +7 812 2970341
Tel: +7 812 2971957
E-mail: imk@incras.ru

(Received 9 December 2013, revised 18 February 2014, accepted 11 March 2014)

doi:10.1111/febs.12781

The possibility of engineering near-infrared fluorescent proteins and biosensors from bacterial phytochrome photoreceptors (BphPs) has led to substantial interest in this family of proteins. The near-infrared fluorescent proteins have allowed non-invasive bio-imaging of deep tissues and whole organs in living animals. BphPs and derived near-infrared fluorescent proteins contain a structural element, called a knot, in their polypeptide chains. The formation of knot structures in proteins was refuted for a long time. Here, we studied the denaturation and renaturation processes of the near-infrared fluorescent probe iRFP, engineered from RpBphP2, which utilizes a heme-derived tetrapyrrole compound biliverdin as a chromophore. iRFP contains a unique figure-of-eight knot. The denaturation and renaturation curves of the iRFP apoform coincided well, suggesting efficient refolding. However, the iRFP holoform exhibited irreversible unfolding and aggregation associated with the bound chromophore. The knot structure in the apoform did not prevent subsequent binding of biliverdin, resulting in the functional iRFP holoform. We suggest that the irreversibility of protein unfolding is caused by post-translational protein modifications, such as chromophore binding, rather than the presence of the knot. These results are essential for future design of BphP-based near-infrared probes, and add important features to our knowledge of protein folding.

Abbreviations
BphPs, bacterial phytochrome photoreceptors; BV, biliverdin IXα; CBD, chromophore-binding domain; GdnHCl, guanidine hydrochloride; iRFP, near-infrared fluorescent protein from bacterial phytochrome RpBphP2 of R. palustris.

Introduction

The possibility of molecularly engineering near-infrared fluorescent proteins and biosensors from bacterial phytochrome photoreceptors (BphPs) for use in bioimaging of deep tissues and organs in whole living animals has increased interest in this family of proteins [1,2]. These probes may be used for numerous basic biological applications and biomedical diagnostics. In general, phytochromes are red/far-red photoreceptors found in plants [3], fungi [4–8] and several types of prokaryotes [9,10]. Light absorption by chromophores of phytochromes induces photoreactions that are transmitted to an effector domain, resulting in activation of signaling. BphPs utilize a heme-derived tetrapyrrole compound, called biliverdin IXα (BV), as a chromophore [11], and their output effector domains usually consist of histidine kinase domains [12].

BphPs exist in two stable states that absorb at 680–710 nm (the Pr state or red-absorbing state) and 740–760 nm (the Pfr state or far-red-absorbing state). Most BphPs adopt the Pr state as their ground state, but a few, namely the bathy BphPs, are found in the Pfr ground state [13]. The BphPs in the Pr and Pfr states differ in the conformation of BV: its D ring adopts either a *cis*- or a *trans*-configuration, respectively. The ground Pr state (or Pfr state) may be converted into the Pfr state (or Pr state) upon illumination with red (or far-red) light. This photo-induced state may be slowly reversed by means of dark reversion or rapidly reversed upon illumination with far-red (red) light [11,12].

BV binding occurs in the chromophore-binding domain (CBD) of the protein, which consists of the PAS and GAF domains in BphPs. The names of the PAS and GAF domains are based on the names of the protein in which these domains were first identified: PER/ARNT/Sim in the case of the PAS domain, and cGMP phosphodiesterase/adenylcyclase/FhlA transcriptional activator in the case of the GAF domain. The PAS domain in BphPs contains a conserved cysteine residue at the N-terminus that covalently attaches to the BV, while the chromophore itself fits into the cleft in the GAF domain.

The N-terminal region of 35 residues upstream of the PAS domain and the loop contributed by the GAF domain (residues 225–257) create a knot structure that bridges the PAS and GAF domains. Initially, this knot was mis-classified as a trefoil knot, in which the polypeptide chain has three crossings [14,15]. However, a high-resolution structure of the CBD from the *Deinococcus radiodurans* bacterial phytochrome DrBphP showed that this knot was a figure-of-eight knot with four crossings of the polypeptide chain [16]. The function of the knot in BphPs remains unclear. However, it has been shown for DrBphP that the knot does not confer an unusually high resistance to mechanical denaturation [17].

A number of possible functions have been attributed to the knot in BphPs, including stabilization of the contacts between the PAS and GAF domains, reduction of the undesirable energy losses upon chromophore isomerization during photoconversion associated with the flexibility of the chromophore-binding site, and ensuring correct positioning of the Cys residue for conjugation to the chromophore by restricting the mobility of the N-terminal region [14]. Another proposed function of knots is to prevent degradation of the proteins in the cell [15,18,19].

Recently, several BphPs have been developed into near-infrared fluorescent proteins [20–22], which are able to bind the BV present at low concentrations in mammalian tissues and thus acquire near-infrared fluorescence without supplementation with a large excess of exogenous BV. The first protein in this series, termed iRFP (recently renamed iRFP713), was engineered from the CBD of the *Rhodopseudomonas palustris* bacterial phytochrome RpBphP2 [20]. iRFP differs from the CBD of RpBphP2 by 13 amino acid substitutions (see Fig. 1). However, similar to the parental RpBphP2, it has a figure-of-eight knot structure. The palette of near-infrared fluorescent proteins has been expanded by the development of four spectrally distinct probes developed from the CBDs of the *R. palustris* bacterial phytochromes RpBphP2 and RpBphP6 [22]. The most blue- and red-shifted of them, iRFP670 and iRFP720, have been successfully used to visualize two tumors in a live mouse using photoacoustic tomography [23]. It was also shown that a bimolecular near-infrared fluorescence complementation probe, named iSplit [24], which was designed from iRFP by splitting the PAS and GAF domains, has decreased intracellular stability with respect to its ancestor. iSplit is formed by the complementation of separate PAS and GAF domains when they are brought into close proximity due to interaction between their fusion partners. As the PAS and GAF domains are synthesized and probably pre-fold separately, iSplit lacks the knot.

For a long time, it was believed that the presence of knots in a protein structure was impossible [25]. Our current knowledge of protein folding is summarized in the energy landscape model [26,27]. It is thought that, during evolution, protein sequences with cooperative and fast folding, i.e. proteins whose folding is described by the funnel-like minimally frustrated

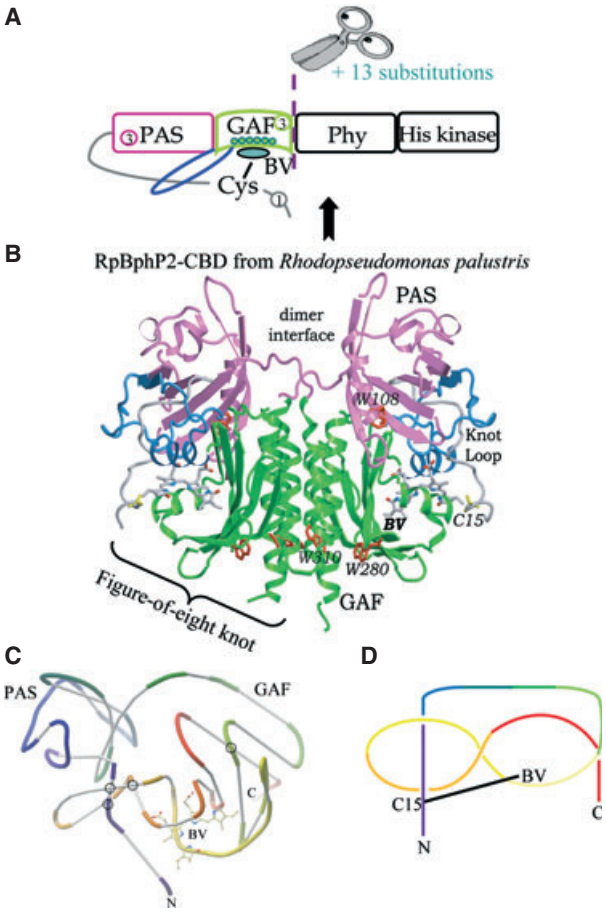

Fig. 1. Structure of the iRFP dimer. (A) iRFP is engineered from *R. palustris* bacterial phytochrome RpBphP2 by removal of the PHY and His kinase domains, and consists only of the CBD comprising the PAS and GAF domains, which differs from the CBD of wild-type RpBphP2 by 13 amino acid substitutions (one in the N-terminus, three in the PAS domain, and the others in the GAF domain, six of which are in direct contact with BV). (B) The 3D structure of the CBD of bacterial phytochrome RpBphP2. The location of Cys15, BV and the knotted loop are shown. The PAS and GAF domains are shown in pink and green, respectively. (C) The figure-of-eight topology of the knot in RpBphP2. The polypeptide chain of CBD of RpBphP2 is shown in various colors to facilitate comparison with the schematic representation of the knot. The bound BV is shown as the stick and ball union. The chain crossings in the protein are highlighted by circles. (D) The schematic representation of the knot. The drawing was made on the basis of X-ray data from the PDB (ID code4E04) [49] deposited in the Protein Data Bank [72].

energy landscape, were mainly selected [28]. However, during knot formation, the protein must overcome the high entropic barrier involved in threading one terminus through a loop in the polypeptide chain, and this process significantly slows down protein folding. Therefore, it is interesting that proteins with the knot topology have been preserved during evolution.

The existence of a knot formation in polypeptide chains was first suggested almost 40 years ago [29]. Nevertheless, this hypothesis was considered improbable for a long time. Moreover, protein structures with knotted topology were discarded from predictions of protein structure on the basis of crystallographic data and simulations of protein folding [30,31]. However, the existence of proteins with various types of knots has now been proven [32–35]. The first discovered knot structure was the trefoil-like knot of carbonic anhydrase [36]. Because deletion of two or three amino acids at the C-terminus of carbonic anhydrase results in disappearance of the knot, this protein is often not considered as a knotted protein. A deep trefoil knot has also been found in RNA methyltransferases [37,38], Rds3p from *Saccharomyces cerevisiae*, which contains a Zn-finger structure [39], and the nucleic acid binding protein of the ribbon-helix-helix (RHH) superfamily of DNA-binding proteins [40,41]. To date, the most complex knots have been detected in ubiquitin C-terminal hydrolases, which contain knots with five crossings of the polypeptide chain [15,42], and the haloacid dehalogenase DehI from *Pseudomonas putida*, which contains a knot with six crossings of the polypeptide chain [40,43].

Moreover, a high conservation of knot structure among members of specific protein families has been shown [44]. This suggests that knots confer some functional advantages to proteins despite the difficulties in folding proteins with such a topologically complex element. Theoretical studies of the folding of knotted proteins shed some light on the knotting mechanism [45]. These investigations suggest that knotting is a rate-limiting step during protein folding, and that knot formation occurs via twisted loop and slipknot intermediates [46–48]. Formation of the most complex knot (in DehI) has been proposed to involve flipping [40].

Here, we studied the guanidine hydrochloride (GdnHCl)-induced denaturation and renaturation processes of iRFP [20] in its holoform state (in complex with the BV chromophore) and its apoform state (chromophore-free). The effects that the BV chromophore and the knot in the iRFP structure have on protein folding are discussed.

Results

Spectral features of iRFP in the holoform

iRFP in the holoform state in buffer solution had an absorption band with a maximum at 280 nm, which may be attributed to aromatic residues, and two bands with

maxima at 390 and 690 nm, caused by chromophore absorption (Fig. 2A). The absorption band of BV in the holoprotein with a maximum at 690 nm was approximately twice as intense as that with a maximum at 390 nm. The far-red absorption band also had a shoulder at 640 nm. The absorption spectrum of free BV in buffer solution is also shown in Fig. 2A. This spectrum contained two absorption bands that peaked at 376 and 680 nm, with the absorption peak at 376 nm being the most pronounced in this case. The iRFP holoprotein has characteristic near-infrared fluorescence with a maximum at 713 nm under excitation with light at a wavelength of 690 nm. The excitation spectrum of iRFP in the holoform state recorded at the maximum of chromophore emission included two bands at 390 and 280 nm, in addition to the band at 690 nm; all of the excitation bands matched the absorption ones (Fig. 3).

The iRFP holoprotein has a pronounced CD spectrum, with a two-peaked negative band in the far-UV region, which is typical of proteins with mixed α/β secondary structures (Fig. 2B). The near-UV CD spectrum of iRFP in the holoform state was characterized by a marked negative band (Fig. 2E). The visible CD spectrum of iRFP had a clearly distinguishable negative band at 690 nm, with a shoulder at 640 nm, and a positive band at 390 nm, which is typical of a bacterial phytochrome in the Pr state (Fig. 2F).

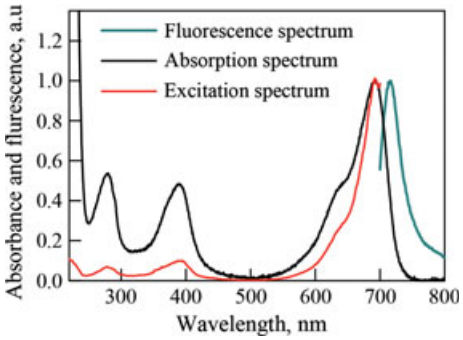

Fig. 3. Spectra of absorption, near-infrared fluorescence at the excitation wavelength of 690 nm, and excitation at the emission wavelength of 713 nm for iRFP in the holoform state.

The tryptophan fluorescence spectrum of iRFP in the holoform state was blue-shifted, with a maximum at 330 nm (Fig. 2C and Table 1). The holoprotein was characterized by a high value of fluorescence anisotropy (r = 0.15; λ_{ex} = 297 nm, λ_{em} = 365 nm), and a short lifetime of tryptophan fluorescence (1.2 ns) (Table 1). The fluorescence of the tryptophan residues of the iRFP holoprotein in its native state was

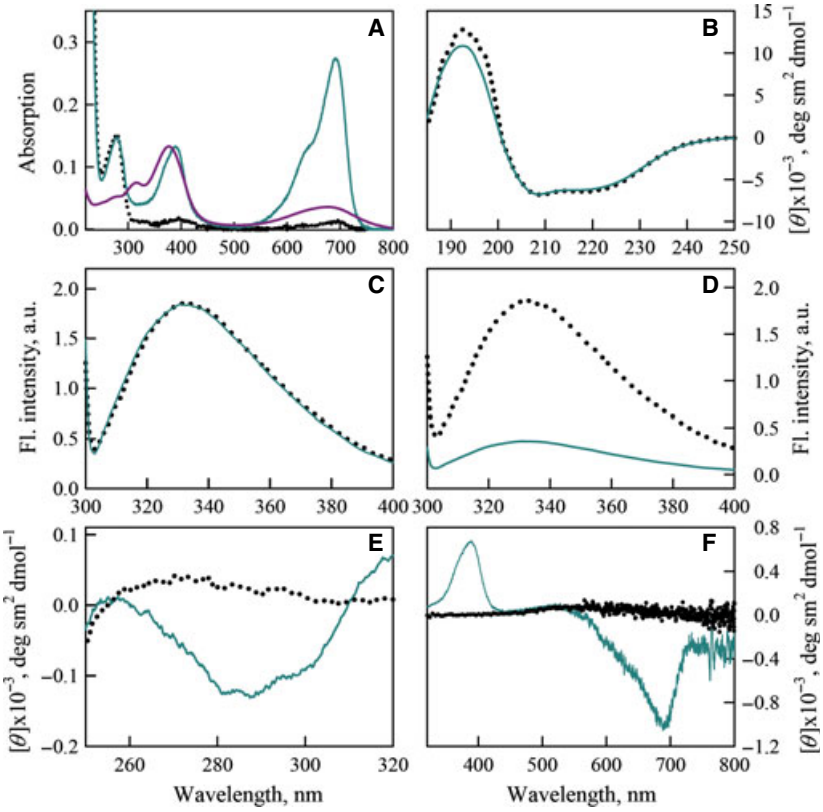

Fig. 2. Spectral features of iRFP in the holoform state (green line) and in the apoform state (black line). (A) Absorption spectra in the UV and visible spectra regions; the absorption spectrum of free BV is also shown (purple line). (B) CD in the far-UV region. (C) Intrinsic fluorescence spectra at the excitation wavelength of 295 nm, normalized to 1, and at the emission wavelength of 365 nm. (D) Intrinsic fluorescence spectra at the excitation wavelength of 295 nm, taking into account the difference in emission of the holoprotein and apoprotein. (E,F) CD in the near-UV and visible regions, respectively.

Table 1. Fluorescence characteristics of the iRFP holoform and apoform.

Parameter	iRFP in the holoform state	iRFP in the apoform state
Intrinsic fluorescence		
λ_{max}, nm (λ_{ex} = 295 nm)	331	332
Parameter A (λ_{ex} = 295 nm)	1.7	1.6
r (λ_{ex} = 297 nm, λ_{em} = 365 nm)	0.15	0.16
λ, ns (λ_{ex} = 297 nm, λ_{em} = 330 nm)	3.7	1.2
Chromophore fluorescence		
λ_{max}, nm (λ_{ex} = 690 nm)	713	–
I_{max},%	100	–

quenched. We analyzed the microenvironment of the tryptophan residues of iRFP using X-ray data for the CBD of bacterial phytochrome RpBphP2 [49], bearing in mind the substitutions that were made to develop iRFP. iRFP contains three tryptophan residues: Trp108, Trp280 and Trp310. All of these residues affect the total fluorescence of the protein, as possible fluorescence quenchers are not present in their microenvironment. The density of the microenvironment of the tryptophan residues is quite high; there 72, 75 and 78 atoms within a 7 Å radius of the indole ring of Trp108, Trp280 and Trp310, respectively. In comparison, the microenvironment of the single tryptophan residue Trp48 of azurin, which has a unique blue fluorescence spectrum with a maximum at 308 nm, comprises only 69 atoms [50]. The high value of fluorescence anisotropy is also indicative of a rigid microenvironment of the tryptophan residues of iRFP. Residue Trp108 is submerged in the inner part of the PAS domain and is inaccessible to the solvent. Residues Trp280 and Trp310 are located closer to the periphery of the GAF domain and are therefore partially exposed to the solvent. Both of these residues have five molecules of bound water in their microenvironment. Residue Trp310 belongs to the α-helix of the GAF domain, which constitutes the dimeric interface of iRFP. As a result, the microenvironment of Trp310 from one monomer of the iRFP dimer includes the Trp310 residue of the other monomer.

All of the tryptophan residues have several polar groups in their microenvironment. The microenvironment of Trp108 contains the polar groups of Arg110 and bound water. The polar groups of residues Arg173, Arg275 and Glu316 are found in the microenvironment of Trp280, and the polar groups of Gln144 and Gln311 are located in the vicinity of Trp310. Usually, blue-shifted spectra are typical of tryptophan residues that are present in microenvironments that are formed mainly by hydrophobic residues, as is the case of the single tryptophan residue Trp48 of azurin [50]. The same fluorescence properties may be observed for tryptophan residues with microenvironments that are predominantly composed of polar residues if the microenvironment is rather rigid [51]. Analysis of the location and characteristics of the tryptophan microenvironments in iRFP showed that its blue-shifted tryptophan fluorescence was caused by the rigidity of the microenvironment, even though these microenvironments contained polar groups.

Unfolding and refolding of iRFP in the holoform state

Quasi-equilibrium curves of the various characteristics of iRFP in the holoform state, including the intensity of tryptophan fluorescence at the emission wavelength of 320 nm, parameter A and fluorescence anisotropy, chromophore fluorescence intensity at an emission wavelength of 713 nm and the ellipticity at 222 nm, are shown in Fig. 4 as a function of the final GdnHCl concentration. Quasi-equilibrium curves of the GdnHCl-induced unfolding and refolding of the holoprotein were recorded after the protein was equilibrated in the presence of a desired denaturant concentration for 24 h. The prolonged equilibration time did not result in a noticeable change in the detected characteristics. As the drastic change in the red absorption band of iRFP in the holoform state was clearly observed over the entire range of GdnHCl concentrations (Fig. 4B, inset, and Fig. 5A), we corrected the chromophore fluorescence by the change in the total absorbance of the solution at the excitation wavelength as described previously [52–54]. It was shown that the denaturation of iRFP in the holoform state was irreversible. Indeed, none the recorded characteristics of the holoprotein recovered to the level of the native protein, even under strong refolding conditions where the final concentration of GdnHCl was < 0.4 M (Figs 4 and 5). Accumulation of partially folded protein in 2 M GdnHCl was observed during refolding of the iRFP holoprotein. This state was characterized by a less pronounced secondary structure compared to the native state (Fig. 4E). The chromophore fluorescence was still quenched in this state, indicating that the protein core was not organized around BV (Fig. 4B). Conversely, the structure around the tryptophan residues was already organized (Fig. 4A,C,D). The characteristics of tryptophan fluorescence in the partially folded state differed from those of native iRFP. As stated previously, iRFP contains three tryptophan residues; Trp108 is located in

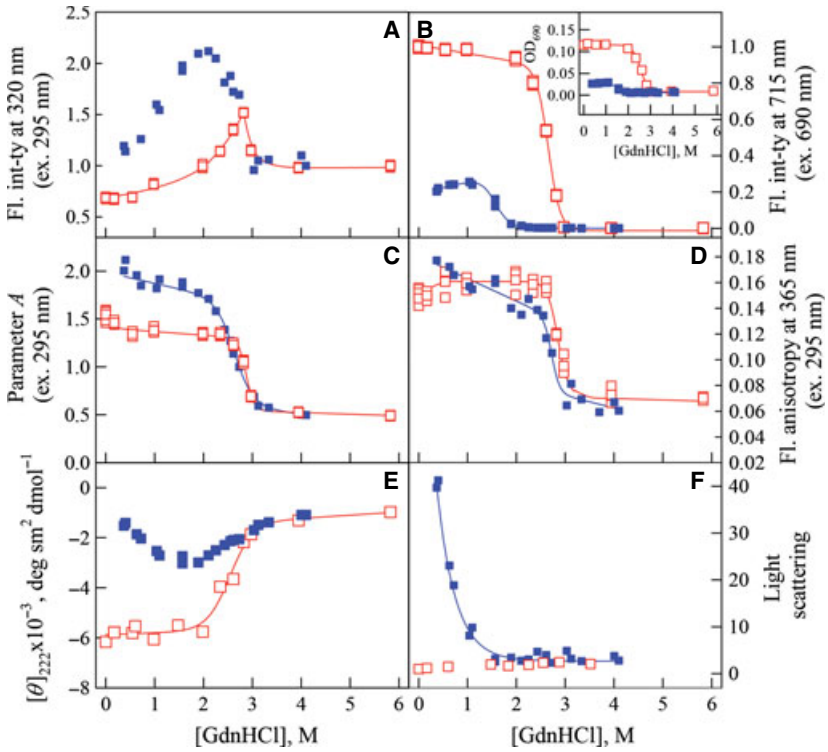

Fig. 4. Denaturation curves of GdnHCl-induced unfolding and refolding of iRFP in the holoform state. (A) Changes in tryptophan fluorescence intensity at the excitation wavelength of 295 nm and the emission wavelength of 320 nm. The data were normalized to the fluorescence signal of the protein in its unfolded state. (B) Change in chromophore fluorescence intensity at the excitation wavelength of 690 nm. As the visible absorption spectrum of iRFP changed significantly over the range of GdnHCl concentrations (inset), the chromophore fluorescence was corrected to the total absorbance of the solution at the excitation wavelength, as described in Experimental procedures. (C) Change in the parameter $A = I_{320}/I_{365}$ at the excitation wavelength of 295 nm. (D) Change in fluorescence anisotropy at the emission wavelength of 365 nm and the excitation wavelength of 297 nm. (E, F) Changes in ellipticity at 222 nm and in light scattering, respectively. Measurements were performed after 24 h incubation of the native or denatured protein in the presence of GdnHCl. Red symbols indicate unfolding of iRFP, and blue symbols represent refolding.

the PAS domain, and Trp280 and Trp310 are located in the GAF domain. In its partially folded state, the chromophore, which must be incorporated into the cleft of the GAF domain, has not yet been packed into the protein globule; therefore, it is tempting to speculate that, in the partially folded state, the PAS domain has collapsed, and the GAF domain remains unstructured. This assumption requires experimental confirmation.

Decreasing the GdnHCl concentration below 1.5 M led to aggregation of molecules in the partially folded state, as indicated by a significant increase in light scattering (Fig. 4F). It was previously shown that, under certain conditions, GdnHCl induce the aggregation of proteins in a partially folded state [55]. The decrease in CD in the far-UV region at GdnHCl concentrations < 1.5 M (Fig. 4E) was most likely due to protein aggregation, as light scattering distorts the CD spectra of macromolecules [56]. Therefore, we wished to determine which element of protein assembly, i.e. the chromophore or the knot, interfered with protein folding. To do this, we expressed and purified the iRFP apoprotein, which lacks the chromophore.

Unfolding and refolding of iRFP in the apoform state

First, we studied the properties of iRFP in the apoform state (Fig. 2). The absorption spectrum of the apoprotein exhibited a dominant absorption band with a maximum at 280 nm, which may be attributed to aromatic residues, and two minor bands with maxima at 390 and 690 nm, caused by absorption of bound BV (Fig. 2A). This was evidence of slight contamination of the apoprotein solutions by the holoprotein. The tryptophan fluorescence spectra of the apoprotein and holoprotein overlap perfectly (Fig. 2C). The other characteristics of tryptophan fluorescence (anisotropy

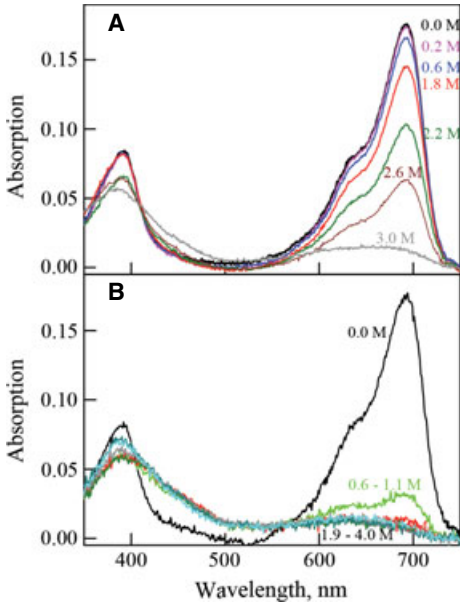

Fig. 5. Change in the visible absorption of iRFP in the holoform state during GdnHCl-induced unfolding (A) and refolding (B). The numbers on the curves indicate the final denaturant concentration in the protein sample. The black curve in (B) is the absorption spectrum of the holoprotein in buffer solution.

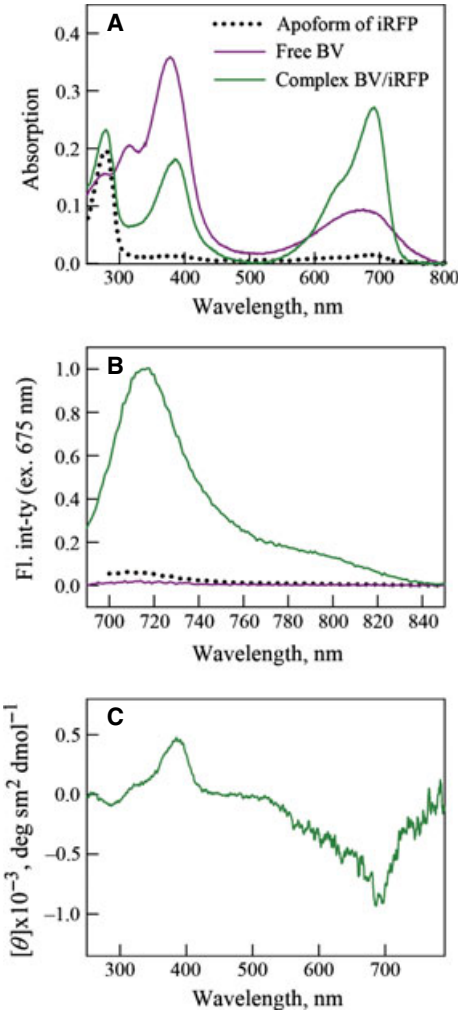

Fig. 6. Interaction of BV with iRFP in the apoform state, showing the absorption in the UV and visible regions (A) and near-infrared fluorescence at the excitation wavelength of 690 nm (B) of iRFP in the apoform state and of apo-iRFP in complex with BV. These spectra were recorded for the solutions obtained by equilibrium microdialysis (see Experimental procedures). This approach allowed us to evaluate the completeness of BV binding to iRFP in the apoform state. A protein concentration of 8 μM was used. According the experimental data, the concentration of bound BV was almost 8 μM. The absorption and fluorescence spectra of free BV at concentrations identical to the bound chromophore are also shown. (C) CD in the near-UV and visible regions for iRFP in the apoform state in complex with BV.

and parameter *A*) were similar for the two forms of the protein (Table 1). The tryptophan fluorescence of iRFP in the holoform state was strongly quenched (Fig. 2D). Furthermore, the lifetime of the tryptophan residues in the excited state of the apoprotein was three times higher compared with that of the holoprotein (3.7 and 1.2 ns, respectively). We assume that the dissipation of tryptophan fluorescence in the holoprotein occurred through an effective non-radiative energy transfer to the chromophore of iRFP. iRFP in the apo- and holoform states has almost identical secondary structures, as confirmed by the coincidence of their CD spectra in the far-UV region. As expected, the CD of the iRFP apoprotein in the visible region was not pronounced because CD in this region was determined by the absorption of the chromophore. Moreover, the apoprotein has no CD signal in the near-UV region. Apparently, the pronounced CD spectrum of the holoprotein in the near-UV region arose from the chromophore, which has minor absorption bands in the UV region in addition to the dominant absorption bands with maxima in the visible region (390 and 690 nm). Thus, the spectral features of iRFP in the apoform state suggest that it retains the secondary and tertiary structures inherent to the native protein. Next, we tested the ability of iRFP in the apoform state to bind BV (Fig. 6). These experiments were performed using equilibrium microdialysis analysis, which utilizes a device divided into two microdialysis cells that are connected by a membrane permeable to low-molecular-weight substances. In our experiment, one cell of the microdialysis device was filled with the apoprotein solution with a concentration of 8 μM. A molar excess of BV (25–70 μM) was added into the second cell of the microdialysis device. The concentration of free BV

was allowed to reach equilibrium in both cells. The binding of BV to iRFP included covalent attachment of the dye to the protein, meaning that the chromophore irreversibly interacts with iRFP, and this is non-equilibrium process. Consequently, the BV binding constant to iRFP cannot be calculated. However, this approach allowed us to establish the proportion of apoprotein molecules that were bound to the chromophore. BV was incorporated into iRFP at a molar ratio of 1 : 1, and full saturation of the apoprotein with BV was reached. This is probably indicative of the absence of mis-folded protein molecules. BV binding to apo-iRFP also resulted in a characteristic change in the visible absorption of the complex and the acquisition of near-infrared fluorescence, as well as pronounced CD in the visible region (Fig. 6). These data show that the apoprotein was able to properly bind BV, and this indicates that the conformation and microenvironment of BV in the apo-iRFP/BV complex is probably similar to that for native iRFP.

Various structural tests (tryptophan fluorescence and ellipticity at 222 nm) were used to examine GdnHCl-induced unfolding and refolding of iRFP in its apoform (Fig. 7). A comparison of the denaturation curves of tryptophan fluorescence intensity, parameter *A* and ellipticity at 222 nm recorded during unfolding of iRFP in the apo- and holoforms is shown in Fig. 7. The experimental data imply that, despite the same secondary and tertiary structures of the apo- and holoproteins, the chromophore significantly stabilizes iRFP and makes the denaturation transition more cooperative. It is important to note that we normalized the fluorescence intensity at the emission wavelengths of 320 and 365 nm for both proteins according to the fluorescence of its unfolded state. We assumed that, in the unfolded protein, there are no conditions that promote energy transfer from the tryptophan residues to the chromophore, and the tryptophan fluorescence of the apo- and holoproteins in the unfolded state may be considered equal. The normalization procedure described above allowed us to compare the denaturation curves of tryptophan fluorescence intensity obtained during unfolding of iRFP in the apo- and holoforms (Fig. 8A,B). It was obvious that the intrinsic fluorescence of the holoprotein was quenched in its native state, and an increase in fluorescence occurred at high denaturant concentrations after disruption of the protein structure.

Discussion

Over the past few decades, our understanding of protein folding has increased, but is based on studies that mainly address the folding of small globular proteins. Studies of proteins that possess complex structural elements will enrich our understanding of protein folding. Thus, studies of knotted proteins are important to understand protein folding in general. Most theoretical and experimental studies of the folding of proteins with knotted topology have been performed on YibK and YbeA, which are members of a class of RNA

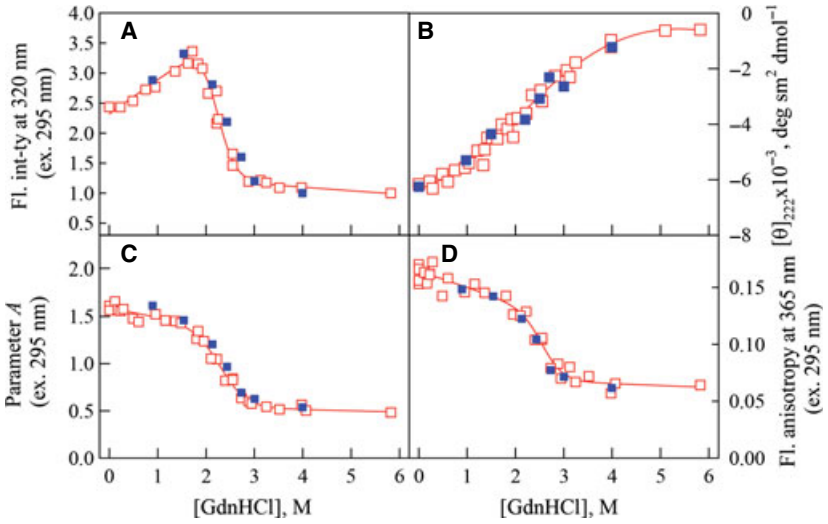

Fig. 7. Denaturation curves of GdnHCl-induced unfolding and refolding of iRFP in the apoform state. (A) Change in tryptophan fluorescence intensity at the excitation wavelength of 295 nm and emission wavelength of 320 nm. The data were normalized to the fluorescence signal of the protein in its unfolded state. (B) Change in ellipticity at 222 nm. (C) Change in parameter $A = I_{320}/I_{365}$ at the excitation wavelength of 295 nm. (D) Change in fluorescence anisotropy at the emission wavelength of 365 nm and the excitation wavelength of 297 nm. Measurements were performed after 24 h incubation of the native or denatured protein in the presence of GdnHCl. Red symbols indicate unfolding of the apoprotein, and blue symbols represent refolding.

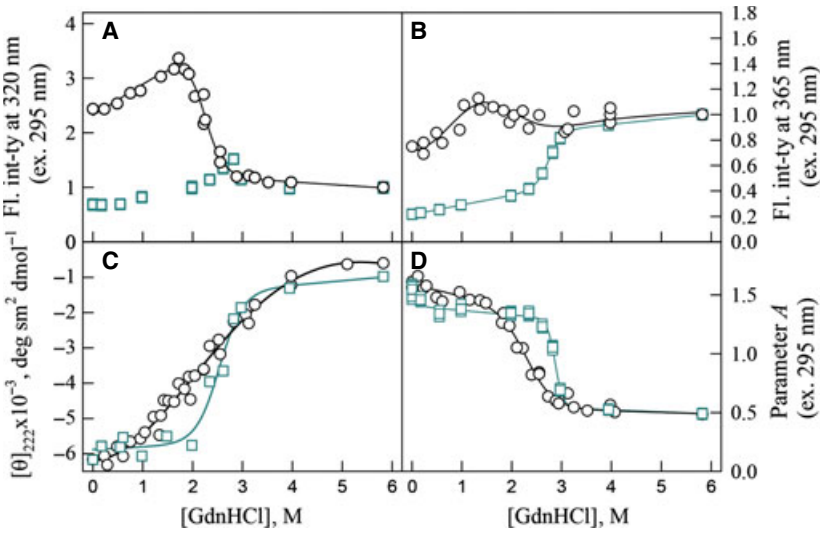

Fig. 8. Stability of iRFP in the holoform state (circles) and the apoform state (squares) against denaturation by GdnHCl.

methyltransferases that contain a deep trefoil knot [57–60]. These studies have shown that the *in vitro* unfolding of both YibK and YbeA induced by chemical denaturants is reversible [58,59], but refolding of these proteins is a complex process that involves accumulation of intermediate states [57]. These studies have also shown that the newly translated polypeptide chain spontaneously folds in a cell-free translation system, and that its correct folding does not require the assistance of molecular chaperones [61].

We studied the GdnHCl-induced denaturation and renaturation processes of the near-infrared fluorescent protein iRFP, which was developed from the truncated bacterial phytochrome RpBphP2 (Fig. 1). iRFP retains the PAS and GAF domains that interact in the figure-of-eight knot. It was shown that the figure-of-eight knot is a feature of all studied PAS-containing phytochromes [14, 16, 62-63]. The knot comprises the N-terminal region upstream of the PAS domain and the loop from the GAF domain [62,63]. The knot region contains two residues that are a highly conserved among PAS-containing phytochromes, i.e. Ile35 and Gln36 [64]. The Gln36 residue is involved in hydrogen bonding that is thought to strengthen the knot. However, the function of the knot is not well established. It was assumed that the knot of bacterial phytochromes minimizes energy losses upon chromophore isomerization during photoconversion, has an effect on productive conformational changes in the photosensory core module (formed by the PAS, GAF and PHY domains) during chromophore isomerization, and/or facilitates efficient binding of the chromophore [14,17].

It was revealed that iRFP denaturation was irreversible, and protein renaturation was complicated by the aggregation of partially folded protein molecules. At the same time, the denaturation of iRFP in the apoform state (BV-free) was fully reversible, meaning that the knot does not prevent efficient protein folding. Based on these data, we conclude that the irreversibility of iRFP denaturation and protein aggregation is associated with the bound chromophore instead of the knot structure.

We next tested an iRFP mutant with a substitution of Cys15 by Ser. This protein is not able to form a covalent bond with the chromophore but is able to bind it into the GAF cleft. Similar experiments to those described above regarding unfolding and refolding of this mutant showed the reversibility of its denaturation in the presence of BV. The integrity of protein structure was tested on the basis of the near-infrared fluorescence of the chromophore and parameter *A* (Fig. 9). We observed recovery of these characteristics at decreasing denaturant concentrations. It is interesting to note that the chromophore fluorescence of the mutant iRFP decreases significantly in the presence of low GdnHCl concentrations (Fig. 9B) under which the protein structure is preserved (Fig. 9A). This indicated that addition of a small amount of denaturant induces dissociation of non-covalently bound BV from the protein. Moreover, the stability of mutant iRFP is comparable to that of iRFP in the apoform state (Figs 8D and 9A). In contrast, covalently bound chromophore was shown to increase the protein stability and make the denaturation transition more cooperative (Fig. 8D). Thus, in the absence of covalently bound BV, the protein refolds into a native-like structure that is ready to bind the chromophore.

We propose that BV covalently bound to iRFP forms non-native interactions with a partially folded intermediate whose accumulation we observed during refolding of iRFP in the holoform state, and thus interferes with correct protein folding. The other possible reason for

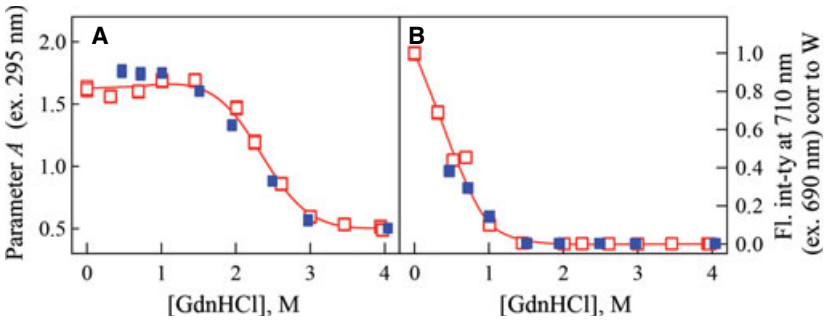

Fig. 9. Denaturation curves of GdnHCl-induced unfolding and refolding of an iRFP mutant with a Cys15 to Ser substitution in the presence of BV. (A) Change in parameter $A = I_{320}/I_{365}$ at the excitation wavelength of 295 nm. (B) Change in chromophore fluorescence intensity at the excitation wavelength of 690 nm. The chromophore fluorescence was corrected to the total absorbance of the solution at the excitation wavelength, as described in Experimental procedures. Measurements were performed after 24 h incubation of the native or denatured protein in the presence of GdnHCl. Red symbols indicate unfolding of iRFP, and blue symbols represent refolding.

the irreversibility of iRFP unfolding is that the covalently bound BV inhibits knot formation during refolding, during which the N-terminal region bearing BV must thread through a loop in the polypeptide chain.

Recent studies have suggested the existence of knotted regions in chemically denatured proteins [65,66]. Mallam *et al.* [65] developed a method to trap a knot in the unfolded state of protein. They constructed mutant variants of YibK and YbeA by inserting Cys residues at both the C- and N-termini; these constructs were circularized in their native state by disulfide bond formation. These oxidized YibK and YbeA variants were unable to untie in their unfolded states. Because the oxidized YibK and YbeA variants refolded with rate constants comparable to those of their reduced non-circularized counterparts, it was concluded that wild-type YibK and YbeA exist in a knotted conformation in their urea-denatured states [65]. Single-molecule Forster resonance energy transfer experiments of doubly labeled TrmD tRNA methyltransferase from *Escherichia coli* revealed that the knot is also preserved in this protein when denatured by GdnHCl [66]. The knotted conformation of the chemically denatured proteins is thought to be responsible for their successful refolding *in vitro*. Experimental data obtained for YibK and YbeA in the cell-free translation system, which is believed to accurately represent *in vivo* translation, demonstrated that knotting is an initial step after polypeptide chain synthesis, and this process is greatly accelerated by the presence of chaperones [61]. The knotted intermediate state is thought to be identical to the chemically denatured knotted structures observed for YibK, YbeA and TrmD. Further folding of the knotted intermediate is chaperone-independent [61].

Notably, the denaturation curves recorded during the unfolding and refolding of iRFP in the apoform state completely coincided. It is assumed that the functional regions and complex structural elements of the natural proteins complicate the energy landscape by creating deep traps, and may lead to hysteresis of the unfolding and refolding curves of the protein [67]. A well-known example of a protein with hysteretic behavior in the unfolding and refolding pathways is green fluorescent protein [68–70]. In the case of green fluorescent protein, an apparent hysteresis was presumed to arise from the green chromophore that is locked in the protein core in the correct active form; this process requires the precise isomerization of the chromophore. A hysteretic behavior of the unfolding–refolding curves was recently observed for 1O6D RNA methyltransferase. This behavior was explained by decoupling of the unfolding events from untying of the knotted polypeptide chain [67]. Thus, the non-hysteretic behavior of the unfolding–refolding curves of iRFP in the apoform state may be indicative of a lack of deep traps in the folding pathway of this protein. The unfolding and refolding curves of YibK, YbeA and TrmD did not exhibit hysteretic behavior either [57].

In conclusion, although denaturation of the iRFP apoform was fully reversible, denaturation of its holoform was irreversible. This suggests that the knot does not inhibit protein folding. Moreover, our data show that the irreversibility of iRFP denaturation and protein aggregation was associated with the bound BV chromophore rather than the presence of the knot in the protein. These results and those of studies of other proteins containing knots suggest that knotting potentially increases the efficiency of protein folding. Knotting of the polypeptide chain decreases the available conformations of the non-folded protein. The top of the energy landscape for the knotted protein is much narrower than that for the unknotted protein. It means

that formation of the knot will limit the folding traces that the unfolded knotted chain may follow. Thus, formation of the knot increases the probability of correct folding into the native state, and reduces the probability of forming mis-folded or aggregated states. In this way, the knots may be considered intrinsic 'chaperones' that assist in proper folding of the protein [61].

Experimental procedures

Protein expression and purification

Escherichia coli LMG194 host cells bearing two plasmids, pBAD/His-B-iRFP with ampicillin resistance marker, encoding infrared fluorescent protein iRFP, and pWA21 h with kanamycin resistance, encoding a heme oxygenase for BV synthesis in bacteria, were obtained as previously described [20].

LMG194 bacteria were grown overnight in LB medium supplemented with ampicillin, kanamycin and 2% glucose to repress synthesis of both proteins. The next morning, the bacterial cells were centrifuged for 20 min at 5000 rpm and 4°C, and resuspended in RM medium (1X M9 salts, 2% casamino acids, 1 mM $MgCl_2$, 1 mM thiamine) containing 0.02% rhamnose to induce heme oxygenase expression, and the cells were grown for an additional 6 h at 37 °C. iRFP expression was induced by addition of 0.002% arabinose. The cells were then grown for 8 h at 37 °C, followed by incubation at 18–20 °C for 24 h. Recombinant iRFP in the holoform state was purified using Ni^+-agarose packed in His-GraviTrap columns (GE Healthcare, Uppsala, Sweden) according to the manufacturer's instructions; however, the elution buffer contained 100 mM EDTA instead of 400 mM imidazole. Further purification was achieved using ion-exchange chromatography on a MonoQ column (GE Healthcare). The protocol described above was slightly modified to express iRFP in the apoform state. The overnight LMG194 culture was centrifuged, resuspended in RM medium, and grown for 2–3 h at 37 °C; then 0.002% arabinose was added to induce protein synthesis. The subsequent steps were the same as for expression of iRFP in the holoform state.

The purity of the protein was analyzed by SDS/PAGE using 15% polyacrylamide gels [71]. Protein concentration was determined by a theoretical extinction coefficient (λ_{280} = 25 350 $M^{-1}{\cdot}cm^{-1}$). The protein was concentrated and stored in 20 mM Tris/HCl buffer, pH 8.0. The absorbance of the protein samples did not exceed 0.2, and the measurements were performed in 20 mM Tris/HCl buffer, pH 8.0, containing 1 mM tris(2-carboxyethyl)phosphine (TCEP).

Spectrophotometric experiments

Absorption spectra were recorded using a U-3900H spectrophotometer (Hitachi, Tokyo, Japan). The experiments were performed in 101.016-QS microcells (5 × 5 mm, Hellma, Müllheim, Germany) with a path length of 5 mm at room temperature.

Analysis of 3D protein structure

The X-ray data for the CBD of bacterial phytochrome RpBphP2 (PDB ID code 4E04) [49] deposited in the Protein Data Bank [72] were used as input data for analysis of the microenvironment peculiarities of the tryptophan residues localized in the structure of iRFP. The analysis was performed as previously described [50,51,73,74]. Figure 10 shows the amino acid sequences of iRFP, its parental CBD of RpBphP2, and the crystallized variant of CBD of RpBphP2 (RpBphP-Cr).

Fluorescence spectroscopy

The fluorescence experiments were performed using a Cary Eclipse spectrofluorometer with FLR cells (10 × 10 × 4 mm (cell type 26.400-F); Starna Scientific, Hainault, England with a path length of 10 mm. Fluorescence anisotropy and the fluorescence lifetime were measured in 101.016-QS microcells using home-built spectrofluorometers with steady-state and time-resolved excitation [75].

The tryptophan fluorescence of the protein was excited at the long-wave absorption spectrum edge (λ_{ex} = 297 nm), where the contribution of the tyrosine residues in the bulk protein fluorescence is negligible. The position and form of the fluorescence spectra were characterized on the basis of parameter $A = I_{320}/I_{365}$, where I_{320} and I_{365} are the fluorescence intensities at the emission wavelengths of 320 and 365 nm, respectively [76,77]. The values for parameter A and fluorescence spectrum were corrected for the instrument sensitivity. The anisotropy of tryptophan fluorescence was calculated using the equation $r = (I_V^V - GI_H^V)/(I_V^V - 2GI_H^V)$, where I_V^V and I_H^V are the vertical and horizontal components of the fluorescence intensity when excited by vertically polarized light, respectively, and G is the relationship of the vertical and horizontal components of the fluorescence intensity when excited by horizontally polarized light; G was calculated using equation ($G = I_V^H/I_H^H$), λ_{em} =365 nm [75]. The specific near-infrared fluorescence of iRFP was excited at 690 nm, and emission was detected at 713 nm.

The unfolding (and refolding) of the protein was initiated by manually mixing 50 μL of the native protein (or protein pre-denatured for 8 h in 4 M GdnHCl) with 500 μL of a buffer solution containing the desired concentration of GdnHCl. GdnHCl (Nacalai Tesque, Kyoto, Japan) was used without further purification. The concentration of the stock GdnHCl solution was determined on the basis of the refraction coefficient, which was determined using an Abbe refractometer (LOMO, Saint Petersburg, Russia). The GdnHCl-dependent fluorescent characteristics of iRFP were recorded after protein incubation in a solution containing

Fig. 10. Alignment of the amino acid sequences of iRFP with RpBphP2 and its crystallized variant (RpBphP2-Cr). Mutations introduced into iRFP compared with RpBphP2 are highlighted in green. Mutations introduced into RpBphP2 to crystallize it are highlighted in blue.

the desired denaturant concentration for 24 h. Further increases in the equilibration time did not result in noticeable changes in the detected characteristics. The spectrofluorometer was equipped with a thermostat that constantly maintained the temperature at 23 °C.

The recorded fluorescence intensity was corrected to the total absorbance of the solution (D_Σ). The corrected fluorescence intensity was defined as I/W, where $W = (1 - 10^{-D\Sigma})/D_\Sigma$. The methods for correction and normalization of the fluorescence intensity have been described previously [52–54]. These studies showed that I/W is approximately equal to Dq, where D and q are the absorption and fluorescence quantum yield of the fluorophore, respectively. Only the corrected fluorescence intensity may be used to evaluate the fraction of molecules that are in different structural states.

Circular dichroism measurements

CD spectra were obtained using a Jasco-810 spectropolarimeter (Jasco, Easton, PA, USA). The far-UV CD spectra were recorded in a 1 mm path length cell from 260 to 190 nm, with a step size of 0.1 nm. The near-UV CD spectra were recorded in a 10 mm path length cell from 320 to 250 nm, with a step size of 0.1 nm. The visible CD spectra were scanned from 800 to 320 nm, with a step size of 0.1 nm, using a 10 mm path length cell. An average of three scans was obtained for all spectra. The CD spectra of the appropriate buffer solution were recorded and subtracted from the protein spectra.

Equilibrium microdialysis

Equilibrium microdialysis was performed using a Harvard Apparatus/Amika device (Holliston, MA, USA) that consists of two chambers (500 μL each) separated by a membrane with a molecular weight cut-off of 10 000. Equilibrium microdialysis involves allocation of two interacting agents, a ligand and its receptor, into two chambers (#2 and #1, respectively) that are divided by a membrane that allows the ligand to pass but is impermeable to the receptor. In our experiment, iRFP solution was placed in chamber #1, and the BV solution was placed in chamber #2. The protein and BV were dissolved in 20 mM Tris/HCl buffer, pH 8.6, for the microdialysis experiment. The iRFP concentration was 8 μM, and the BV concentration varied from 25 to 70 μM. BV hydrochloride (Frontier Scientific, West Logan, UT, USA) was used without further purification. The BV concentration was determined on the basis of an extinction coefficient (λ_{376}) of 45 500 $M^{-1}\cdot cm^{-1}$ [78].

The microdialysis device was left on a rocking platform for 5 days at 4 °C to allow equilibration of the free BV concentration in both chambers. The equilibration time was estimated from test experiments in which one chamber contained BV and the second chamber contained a buffer solution. After equilibration, the free BV concentration (C_f) in both chambers was equal, while the total BV concentration in chamber #1 was greater than that in chamber #2, due to the concentration of the bound chromophore (C_b), which was determined using the following equation: $C_b = C_0 - 2C_f$, where C_0 is the initial BV concentration in chamber #2.

The absorption spectrum of the solution in chamber #1 represents the superposition of the absorption spectra of free BV at concentration C_f and the iRFP/BV complex. The absorption spectrum of the latter was calculated by subtracting the absorption spectra in chamber #2 from that in chamber #1.

Acknowledgements

This work was in part supported by scholarships from the President of the Russian Federation (to Olesya V.

Stepanenko and Olga V. Stepanenko), and grants from the Russian Foundation of Basic Research (13-04-01842 to I.M.K.), the Russian Ministry of Education and Science (Agreement 8480 to K.K.T.), the Molecular and Cell Biology Program of the Russian Academy of Sciences (to K.K.T.), the Academy of Finland (263371 and 266992 to V.V.V.) and the US National Institutes of Health (GM073913 and CA164468 to V.V.V.).

Author contributions

K.K.T., V.V.V. and I.M.K. conceived the study. Olesya V. Stepanenko, G.S.B., Olga V. Stepanenko and D.M.S. performed the experiments and analyzed the data. All authors took part in discussion of the results and manuscript preparation. All authors read and approved the final manuscript.

References

1 Shu X, Royant A, Lin MZ, Aguilera TA, Lev-Ram V, Steinbach PA & Tsien RY (2009) Mammalian expression of infrared fluorescent proteins engineered from a bacterial phytochrome. *Science* **324**, 804–807.
2 Piatkevich KD, Subach FV & Verkhusha VV (2013) Engineering of bacterial phytochromes for near-infrared imaging, sensing, and light-control in mammals. *Chem Soc Rev* **42**, 3441–3452.
3 Butler WL, Norris KH, Siegelman HW & Hendricks SB (1959) Detection, assay, and preliminary purification of the pigment controlling photoresponsive development of plants. *Proc Natl Acad Sci USA* **45**, 1703–1708.
4 Blumenstein A, Vienken K, Tasler R, Purschwitz J, Veith D, Frankenberg-Dinkel N & Fischer R (2005) The *Aspergillus nidulans* phytochrome FphA represses sexual development in red light. *Curr Biol* **15**, 1833–1838.
5 Brandt S, von Stetten D, Gunther M, Hildebrandt P & Frankenberg-Dinkel N (2008) The fungal phytochrome FphA from *Aspergillus nidulans*. *J Biol Chem* **283**, 34605–34614.
6 De Riso V, Raniello R, Maumus F, Rogato A, Bowler C & Falciatore A (2009) Gene silencing in the marine diatom *Phaeodactylum tricornutum*. *Nucleic Acids Res* **37**, e96.
7 Froehlich AC, Noh B, Vierstra RD, Loros J & Dunlap JC (2005) Genetic and molecular analysis of phytochromes from the filamentous fungus *Neurospora crassa*. *Eukaryot Cell* **4**, 2140–2152.
8 Wu SH, McDowell MT & Lagarias JC (1997) Phycocyanobilin is the natural precursor of the phytochrome chromophore in the green alga *Mesotaenium caldariorum*. *J Biol Chem* **272**, 25700–25705.
9 Hughes J, Lamparter T, Mittmann F, Hartmann E, Gartner W, Wilde A & Borner T (1997) A prokaryotic phytochrome. *Nature* **386**, 663.
10 Montgomery BL & Lagarias JC (2002) Phytochrome ancestry: sensors of bilins and light. *Trends Plant Sci* **7**, 357–366.
11 Rockwell NC & Lagarias JC (2010) A brief history of phytochromes. *Chem Phys Chem* **11**, 1172–1180.
12 Auldridge ME & Forest KT (2011) Bacterial phytochromes: more than meets the light. *Crit Rev Biochem Mol Biol* **46**, 67–88.
13 Rottwinkel G, Oberpichler I & Lamparter T (2010) Bathy phytochromes in rhizobial soil bacteria. *J Bacteriol* **192**, 5124–5133.
14 Wagner JR, Brunzelle JS, Forest KT & Vierstra RD (2005) A light-sensing knot revealed by the structure of the chromophore-binding domain of phytochrome. *Nature* **438**, 325–331.
15 Virnau P, Mirny LA & Kardar M (2006) Intricate knots in proteins: function and evolution. *PLoS Comput Biol* **2**, e122.
16 Wagner JR, Zhang J, Brunzelle JS, Vierstra RD & Forest KT (2007) High resolution structure of *Deinococcus* bacteriophytochrome yields new insights into phytochrome architecture and evolution. *J Biol Chem* **282**, 12298–12309.
17 Bornschlogl T, Anstrom DM, Mey E, Dzubiella J, Rief M & Forest KT (2009) Tightening the knot in phytochrome by single-molecule atomic force microscopy. *Biophys J* **96**, 1508–1514.
18 Dzubiella J (2009) Sequence-specific size, structure, and stability of tight protein knots. *Biophys J* **96**, 831–839.
19 Huang L & Makarov DE (2008) Translocation of a knotted polypeptide through a pore. *J Chem Phys* **129**, 121107.
20 Filonov GS, Piatkevich KD, Ting LM, Zhang J, Kim K & Verkhusha VV (2011) Bright and stable near-infrared fluorescent protein for *in vivo* imaging. *Nat Biotechnol* **29**, 757–761.
21 Piatkevich KD, Subach FV & Verkhusha VV (2013) Far-red light photoactivatable near-infrared fluorescent proteins engineered from a bacterial phytochrome. *Nat Commun* **4**, 2153.
22 Shcherbakova DM & Verkhusha VV (2013) Near-infrared fluorescent proteins for multicolor *in vivo* imaging. *Nat Methods* **10**, 751–754.
23 Krumholz A, Shcherbakova DM, Xia J, Wang LV & Verkhusha VV (2014) Multicontrast photoacoustic *in vivo* imaging using near-infrared fluorescent proteins. *Sci Rep* **4**, 3939.
24 Filonov GS & Verkhusha VV (2013) A near-infrared BiFC reporter for *in vivo* imaging of protein–protein interactions. *Chem Biol* **20**, 1078–1086.
25 Mansfield ML (1994) Are there knots in proteins? *Nat Struct Biol* **1**, 213–214.

26 Onuchic JN & Wolynes PG (2004) Theory of protein folding. *Curr Opin Struct Biol* **14**, 70–75.
27 Turoverov KK, Kuznetsova IM & Uversky VN (2010) The protein kingdom extended: ordered and intrinsically disordered proteins, their folding, supramolecular complex formation, and aggregation. *Prog Biophys Mol Biol* **102**, 73–84.
28 Watters AL, Deka P, Corrent C, Callender D, Varani G, Sosnick T & Baker D (2007) The highly cooperative folding of small naturally occurring proteins is likely the result of natural selection. *Cell* **128**, 613–624.
29 Crippen GM (1974) Topology of globular proteins. *J Theor Biol* **45**, 327–338.
30 Khatib F, Weirauch MT & Rohl CA (2006) Rapid knot detection and application to protein structure prediction. *Bioinformatics* **22**, e252–e259.
31 Bryant TN, Watson HC & Wendell PL (1974) Structure of yeast phosphoglycerate kinase. *Nature* **247**, 14–17.
32 Virnau P, Mallam A & Jackson S (2011) Structures and folding pathways of topologically knotted proteins. *J Phys Condens Matter* **23**, 033101.
33 Lai YL, Yen SC, Yu SH & Hwang JK (2007) pKNOT: the protein KNOT web server. *Nucleic Acids Res* **35**, W420–W424.
34 Taylor WR (2007) Protein knots and fold complexity: some new twists. *Comput Biol Chem* **31**, 151–162.
35 Taylor WR (2000) A deeply knotted protein structure and how it might fold. *Nature* **406**, 916–919.
36 Richardson JS (1977) β-sheet topology and the relatedness of proteins. *Nature* **268**, 495–500.
37 Nureki O, Shirouzu M, Hashimoto K, Ishitani R, Terada T, Tamakoshi M, Oshima T, Chijimatsu M, Takio K, Vassylyev DG *et al.* (2002) An enzyme with a deep trefoil knot for the active-site architecture. *Acta Crystallogr D Biol Crystallogr* **58**, 1129–1137.
38 Michel G, Sauve V, Larocque R, Li Y, Matte A & Cygler M (2002) The structure of the RlmB 23S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot. *Structure* **10**, 1303–1315.
39 van Roon AM, Loening NM, Obayashi E, Yang JC, Newman AJ, Hernandez H, Nagai K & Neuhaus D (2008) Solution structure of the U2 snRNP protein Rds3p reveals a knotted zinc-finger motif. *Proc Natl Acad Sci USA* **105**, 9621–9626.
40 Bolinger D, Sulkowska JI, Hsu HP, Mirny LA, Kardar M, Onuchic JN & Virnau P (2010) A Stevedore's protein knot. *PLoS Comput Biol* **6**, e1000731.
41 Andreeva A & Murzin AG (2010) Structural classification of proteins and structural genomics: new insights into protein folding and evolution. *Acta Crystallogr F Struct Biol Cryst Commun* **66**, 1190–1197.
42 Lua RC & Grosberg AY (2006) Statistics of knots, geometry of conformations, and evolution of proteins. *PLoS Comput Biol* **2**, e45.
43 Schmidberger JW, Wilce JA, Weightman AJ, Whisstock JC & Wilce MC (2008) The crystal structure of DehI reveals a new α-haloacid dehalogenase fold and active-site mechanism. *J Mol Biol* **378**, 284–294.
44 Sulkowska JI, Rawdon EJ, Millett KC, Onuchic JN & Stasiak A (2012) Conservation of complex knotting and slipknotting patterns in proteins. *Proc Natl Acad Sci USA* **109**, e1715–e1723.
45 Sulkowska JI, Noel JK, Ramirez-Sarmiento CA, Rawdon EJ, Millett KC & Onuchic JN (2013) Knotting pathways in proteins. *Biochem Soc Trans* **41**, 523–527.
46 Sulkowska JI, Noel JK & Onuchic JN (2012) Energy landscape of knotted protein folding. *Proc Natl Acad Sci USA* **109**, 17783–17788.
47 Noel JK, Sulkowska JI & Onuchic JN (2010) Slipknotting upon native-like loop formation in a trefoil knot protein. *Proc Natl Acad Sci USA* **107**, 15403–15408.
48 Sulkowska JI, Sulkowski P & Onuchic J (2009) Dodging the crisis of folding proteins with knots. *Proc Natl Acad Sci USA* **106**, 3119–3124.
49 Bellini D & Papiz MZ (2012) Dimerization properties of the RpBphP2 chromophore-binding domain crystallized by homologue-directed mutagenesis. *Acta Crystallogr D Biol Crystallogr* **68**, 1058–1066.
50 Turoverov KK, Kuznetsova IM & Zaitsev VN (1985) The environment of the tryptophan residue in *Pseudomonas aeruginosa* azurin and its fluorescence properties. *Biophys Chem* **23**, 79–89.
51 Kuznetsova IM, Yakusheva TA & Turoverov KK (1999) Contribution of separate tryptophan residues to intrinsic fluorescence of actin. Analysis of 3D structure. *FEBS Lett* **452**, 205–210.
52 Kuznetsova IM, Sulatskaya AI, Povarova OI & Turoverov KK (2012) Reevaluation of ANS binding to human and bovine serum albumins: key role of equilibrium microdialysis in ligand–receptor binding characterization. *PLoS ONE* **7**, e40845.
53 Sulatskaya AI, Kuznetsova IM & Turoverov KK (2011) Interaction of thioflavin T with amyloid fibrils: stoichiometry and affinity of dye binding, absorption spectra of bound dye. *J Phys Chem B* **115**, 11519–11524.
54 Sulatskaya AI, Povarova OI, Kuznetsova IM, Uversky VN & Turoverov KK (2012) Binding stoichiometry and affinity of fluorescent dyes to proteins in different structural states. *Methods Mol Biol* **895**, 441–460.
55 Povarova OI, Kuznetsova IM & Turoverov KK (2010) Differences in the pathways of proteins unfolding induced by urea and guanidine hydrochloride: molten globule state and aggregates. *PLoS ONE* **5**, e15035.
56 Tinoco I Jr, Bustamante C & Maestre MF (1980) The optical activity of nucleic acids and their aggregates. *Annu Rev Biophys Bioeng* **9**, 107–141.

57 Mallam AL & Jackson SE (2007) A comparison of the folding of two knotted proteins: YbeA and YibK. *J Mol Biol* **366**, 650–665.
58 Mallam AL & Jackson SE (2006) Probing nature's knots: the folding pathway of a knotted homodimeric protein. *J Mol Biol* **359**, 1420–1436.
59 Mallam AL & Jackson SE (2005) Folding studies on a knotted protein. *J Mol Biol* **346**, 1409–1421.
60 Mallam AL, Morris ER & Jackson SE (2008) Exploring knotting mechanisms in protein folding. *Proc Natl Acad Sci USA* **105**, 18740–18745.
61 Mallam AL & Jackson SE (2011) Knot formation in newly translated proteins is spontaneous and accelerated by chaperonins. *Nat Chem Biol* **8**, 147–153.
62 Essen LO, Mailliet J & Hughes J (2008) The structure of a complete phytochrome sensory module in the Pr ground state. *Proc Natl Acad Sci USA* **105**, 14709–14714.
63 Yang X, Kuk J & Moffat K (2008) Crystal structure of *Pseudomonas aeruginosa* bacteriophytochrome: photoconversion and signal transduction. *Proc Natl Acad Sci USA* **105**, 14715–14720.
64 Wagner JR, Zhang J, von Stetten D, Gunther M, Murgida DH, Mroginski MA, Walker JM, Forest KT, Hildebrandt P & Vierstra RD (2008) Mutational analysis of *Deinococcus radiodurans* bacteriophytochrome reveals key amino acids necessary for the photochromicity and proton exchange cycle of phytochromes. *J Biol Chem* **283**, 12212–12226.
65 Mallam AL, Rogers JM & Jackson SE (2010) Experimental detection of knotted conformations in denatured proteins. *Proc Natl Acad Sci USA* **107**, 8189–8194.
66 Wang P, Yang L, Liu P, Gao YQ & Zhao XS (2013) Single-molecule detection reveals knot sliding in TrmD denaturation. *Chemistry* **19**, 5909–5916.
67 Andrews BT, Capraro DT, Sulkowska JI, Onuchic JN & Jennings PA (2012) Hysteresis as a marker for complex, overlapping landscapes in proteins. *J Phys Chem Lett* **4**, 180–188.
68 Andrews BT, Roy M & Jennings PA (2009) Chromophore packing leads to hysteresis in GFP. *J Mol Biol* **392**, 218–227.
69 Stepanenko OV, Stepanenko OV, Kuznetsova IM, Verkhusha VV & Turoverov KK (2013) β-barrel scaffold of fluorescent proteins: folding, stability and role in chromophore formation. *Int Rev Cell Mol Biol* **302**, 221–278.
70 Stepanenko OV, Stepanenko OV, Kuznetsova IM, Shcherbakova DM, Verkhusha VV & Turoverov KK (2012) Distinct effects of guanidine thiocyanate on the structure of superfolder GFP. *PLoS ONE* **7**, e48809.
71 Laemmli UK (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685.
72 Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN & Bourne PE (2000) The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242.
73 Kuznetsova IM, Stepanenko OV, Turoverov KK, Staiano M, Scognamiglio V, Rossi M & D'Auria S (2005) Fluorescence properties of glutamine-binding protein from *Escherichia coli* and its complex with glutamine. *J Proteome Res* **4**, 417–423.
74 Giordano A, Russo C, Raia CA, Kuznetsova IM, Stepanenko OV & Turoverov KK (2004) Highly UV-absorbing complex in selenomethionine-substituted alcohol dehydrogenase from *Sulfolobus solfataricus*. *J Proteome Res* **3**, 613–620.
75 Turoverov KK, Biktashev AG, Dorofeiuk AV & Kuznetsova IM (1998) A complex of apparatus and programs for the measurement of spectral, polarization and kinetic characteristics of fluorescence in solution. *Tsitologiia* **40**, 806–817. (in Russian with English abstract).
76 Turoverov KK & Kuznetsova IM (2003) Intrinsic fluorescence of actin. *J Fluoresc* **13**, 41–57.
77 Stepanenko OV, Kuznetsova IM, Turoverov KK, Huang C & Wang CC (2004) Conformational change of the dimeric DsbC molecule induced by GdnHCl A study by intrinsic fluorescence. *Biochemistry* **43**, 5296–5303.
78 McDonagh AF & Palma LA (1980) Preparation and properties of crystalline biliverdin IXα. Simple methods for preparing isomerically homogeneous biliverdin and [^{14}C]biliverdin by using 2,3-dichloro-5,6-dicyanobenzoquinone. *Biochem J* **189**, 193–208.