DOI: 10.1002/cbic.200900207

Bipartite Tetracysteine Display Requires Site Flexibility for ReAsH Coordination

Jessica L. Goodman,[b] Daniel B. Fried,[b] and Alanna Schepartz*[a]

Fluorescence is an essential tool for monitoring molecular interactions, cellular localization, and pathway activation both in vitro and in vivo. The development of "profluorescent" bisarsenical dyes in 1998 advanced this field in a significant way because these dyes provide an alternative labeling strategy that complements some of the limitations of labeling with fluorescent proteins.[1] These small, cell-permeable dyes, 4,5-bis(1,3,2-dithiarsolan-2-yl)fluorescein (FlAsH-EDT_2, Figure S1 A in the Supporting Information) and 4,5-bis(1,3,2-dithiarsolan-2-yl)-resorufin (ReAsH-EDT_2, Figure S1B) are not fluorescent in the 1,2-ethanedithiol (EDT)-bound form.[1,2] Fluorescence occurs upon thiol–arsenic ligand exchange of EDT for a tetracysteine motif (Cys-Cys-Pro-Gly-Cys-Cys), and this permits temporal fluorescent control.[1] Recently, we reported that the central Pro-Gly in the tetracysteine motif could be replaced with one or more folded protein domains without a significant loss in bisarsenical affinity or brightness; we refer to this technique as "bipartite tetracysteine display".[3] By using a set of well-characterized model systems, we showed that bipartite tetracysteine display supports fluorescence-based differentiation between folded and unfolded protein states and can detect protein–protein interactions in live cells. The model constructs that were evaluated in our initial report shared two features: the dicysteine motifs were located near the protein N and/or C termini, and the structures were primarily α-helical. Here we ask whether bipartite display also functions effectively within the context of two more complex protein architectures, p53 and EmGFP, in which the requisite tetracysteine motif is introduced into loops and β-strands.

We selected p53[4–6] to evaluate the feasibility of intramolecular bipartite display between adjacent protein loops, in large part because of the functional relevance of alternatively folded conformational states of this oncogenic protein.[4–7] To investigate bipartite display between adjacent intramolecular β-strands we selected emerald green fluorescent protein (EmGFP) because of its robust β-barrel fold.[8] We emphasize that for our binding experiments we monitored intramolecular bipartite display under stringent in vitro binding conditions that recapitulate a complex cellular environment replete with competing protein and small-molecule thiols. We find that under these conditions, only those constructs in which the thiols are located within flexible regions as defined by crystallographic Cα B-factors[9] greater than 20 bound ReAsH with high affinity and generated fluorescent complexes. Without exception, these regions are located within protein loops. By contrast, under the competitive binding conditions that characterize the intracellular milieu, no construct in which the thiols are located within β-strands bound ReAsH detectably. We conclude that site flexibility, as evaluated by crystallographic B-factors, is an excellent metric for the design of proteins that function effectively in bipartite tetracysteine display by using ReAsH.

Two criteria were applied to select positions at which Cys–Cys motifs could be introduced into loops of p53 (endogenous residues 94 to 300) or Cys-Xaa-Cys motifs could be introduced into the β-strands of EmGFP.[6] First, we identified those residues in each protein that are not conserved within their respective protein family and are demonstrably robust to mutation. Of this set, we selected sites that sampled a range of flexibilities, as judged by the Cα B-factor values in the crystal structures (PDB IDs: 1TUP[6,7] and 1EMA[10]). By using these criteria, five distinct p53 bipartite variants were designed (Figure 1 A). Our first design, p53-1, mimics our initial bipartite designs with dicysteine motifs near the N and C termini (P92C, L93C, P301C, G302C) and Cα B-factor values above 60. The other p53 variants incorporated dicysteine pairs within loops. These include p53-2 in which the B-factors at the Cys-bearing positions (Y107C, G108C, D148C, S149C) range from 23.3–39.9, and p53-3 (S116C, G117C, T123C, C124C), which has values between 10.2–21.2. In the case of p53-4, the B-factors at the Cys-bearing positions (S96C, V97C, M169C, T170C) span a smaller range, from 16.7–17.3, whereas in p53-5 (A129C, L130C, R290C, K291C) they vary widely, 23.0–68.5.

The analogous criteria were applied to select positions at which Cys-Xaa-Cys motifs could be introduced into adjacent β-strands of EmGFP[8] (Figure 1 B). We elected to incorporate Cys-Xaa-Cys motifs in the β-strands because of the alternate residue presentation in pleated β-sheets. Not surprisingly, the Cα B-factor values for the residues located within the β-strands of EmGFP are significantly lower than those characterizing the loops in p53. EmGFP-1 has Cα B-factor values at Cys-bearing positions (D19C, D21C, K26C, S28C) that range between 15.6–26.3. In greater contrast to p53, EmGFP-2 (E34C, D36C, K41C, T43C), EmGFP-3 (R109C, E111C, R122C, E124C), and EmGFP-4 (N149C, Y151C, Y200C, S202C) have considerably smaller B-factor ranges, 4.9–18.1, 12.7–15.2, and 8.6–18.3, respectively.[11]

Six of the eight p53 and EmGFP variants (p53-1, p53-2, p53-3, EmGFP-1, EmGFP-2, and EmGFP-4) described above could

[a] *Prof. A. Schepartz*
Departments of Chemistry and Molecular, Cellular and Developmental Biology
Yale University
275 Prospect Street, New Haven, CT 06511 (USA)
Fax: (+1) 203-432-3486
E-mail: alanna.schepartz@yale.edu

[b] *Dr. J. L. Goodman,[+] D. B. Fried[+]*
Department of Chemistry, Yale University
275 Prospect Street, New Haven, CT 06511 (USA)

[+] *These authors contributed equally to this work.*

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900207.

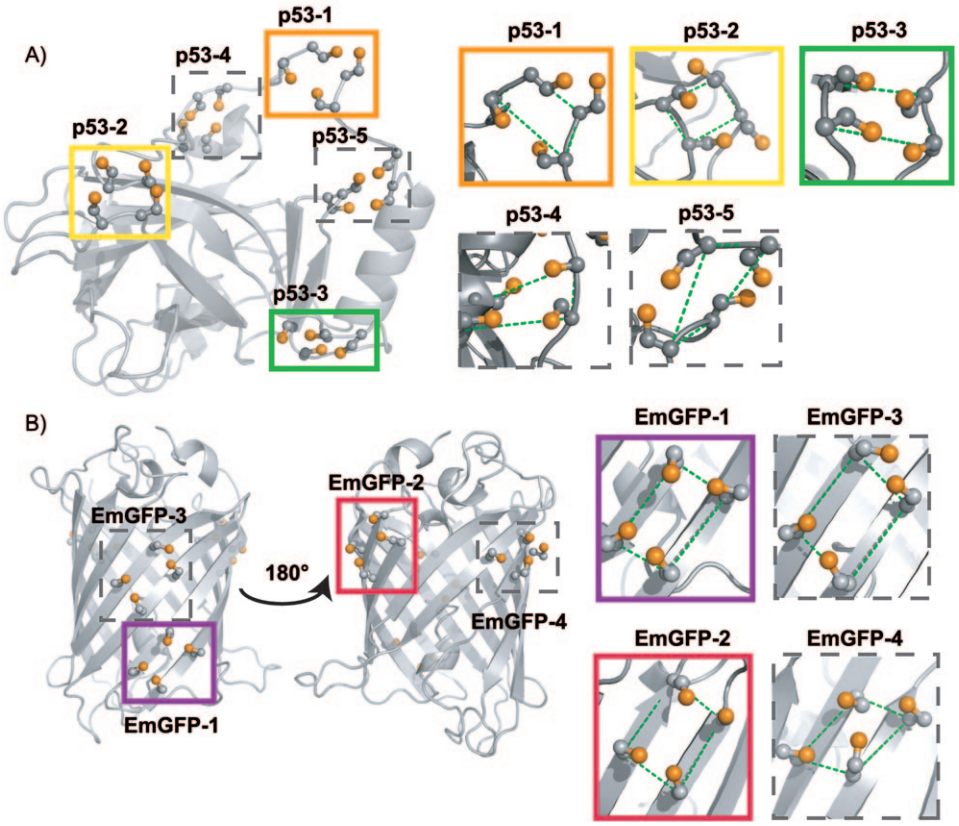

Figure 1. Bipartite display designs for p53 (with ten native cysteine residues)[6,7] and EmGFP (with two native cysteine residues). A) p53 bipartite display variants (PDB ID: 1TUP[7]), p53-1 (P92C, L93C, P301C, G302C), p53-2 (Y107C, G108C, D148C, S149C), p53-3 (S116C, G117C, T123C, C124C), p53-4 (S96C, V97C, M169C, T170C), and p53-5 (A129C, L130C, R290C, K291C). p53-4 and p53-5 are shown in a dotted box because they did not express well and therefore were not examined further. B) EmGFP bipartite display variants (PDB ID: 1EMA[10]), EmGFP-1 (D19C, D21C, K26C, S28C), EmGFP-2 (E34C, D36C, K41C, T43C), EmGFP-3 (R109C, E111C, R122C, E124C) and EmGFP-4 (N149C, Y151C, Y200C, S202C). EmGFP-3 and EmGFP-4 are shown in a dotted box because EmGFP-3 did not express well and EmGFP-4 was unstructured. These variants were not examined in the ReAsH binding trials.

be expressed in *E. coli* and were purified to homogeneity (see the Supporting Information). We first used circular dichroism (CD) spectroscopy to compare the secondary structures of the expressed p53 and EmGFP variants to their respective wild-type (wt) counterparts. The three p53 variants studied (p53-1, p53-2, and p53-3) have CD spectra comparable to that of p53-wt (Figure 2A); this indicates that cysteine incorporation did not dramatically alter protein secondary structure.[12] In contrast, only two (EmGFP-1 and EmGFP-2) of the three EmGFP variants studied displayed CD spectra comparable to that of EmGFP-wt[13] (Figure 2B). EmGFP-4, although it could be expressed, showed significantly reduced green fluorescence relative to the other variants and was unstructured as judged by CD. We note that the CD spectra of EmGFP-1 and EmGFP-2 showed a small reduction in β-sheet structure when compared to EmGFP-wt. To characterize the differences between these EmGFP variants more carefully, we recorded their fluorescence spectra and compared them to EmGFP-wt; we reasoned that chromophore formation was perhaps the most sensitive metric of secondary structure in this system. This comparison revealed that the fluorescence of EmGFP-1 and EmGFP-2 is comparable to that of EmGFP-wt, and this provides additional evidence that cysteine incorporation did not significantly perturb the EmGFP fold.

Next we set out to quantify the affinity of ReAsH for each of these variants in vitro under conditions necessary to monitor fluorescence in live cells. These conditions (100 mM Tris, 75 mM NaCl, 3.5 mM TCEP, 1 mM EDT, pH 7.8) are distinguished by a relatively high concentration of EDT, which is essential to minimize the binding of ReAsH to numerous thiol-containing intracellular proteins and small molecules. Under these stringent binding conditions, p53-1 bound ReAsH with a K_d of 90±80 μM, p53-2 bound ReAsH with a K_d of 21±3 μM, and p53-3 bound ReAsH with a K_d of 35±6 μM (Table 1, Figure 3A, stringent). As expected, p53-wt showed no affinity for ReAsH under these conditions. These K_d values are comparable to those previously observed for our GCN4 and Jun bipartite display variants, 28±7 and 10±7 μM, respectively.[3] We also quantified the affinity of these variants for ReAsH under permissive conditions (100 mM Tris, 75 mM NaCl, 3.5 mM tris(2-carboxyethyl)phosphine (TCEP), pH 7.8), which could be employed, for example, in an in vitro screen for molecules that rescue p53 folding mutants. Under permissive conditions (Figure 3A, permissive) the three p53 variants showed improved affinity for ReAsH with K_d values of

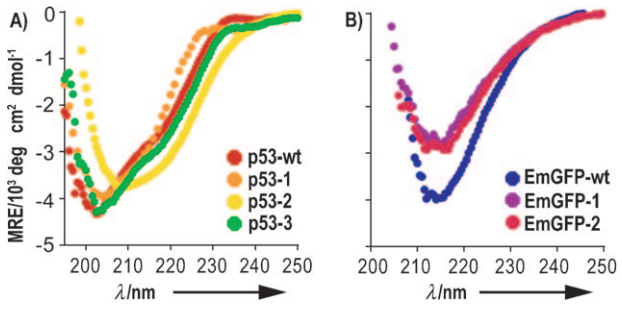

Figure 2. Circular dichroism (CD) spectra for p53 and EmGFP bipartite display variants. A) Wavelength-dependent CD spectra of 8.1 μM p53-wt (wild type), 8.1 μM p53-1, 8.1 μM p53-2, and 8.1 μM p53-3. B) Wavelength-dependent CD spectra of 25 μM EmGFP-wt, 25 μM EmGFP-1 and 25 μM EmGFP-2. CD data were acquired at 25 °C in 100 mM Tris·HCl, 75 mM NaCl, and 3.5 mM TCEP (pH 7.8) in the absence of competing dithiols (EDT). MRE = mean residue ellipticities.

Table 1. K_d values for bipartite design p53 and EmGFP variants under stringent (1 mM EDT) and permissive (0 mM EDT) conditions.

	K_d values [μM]	
	1 mM EDT	0 mM EDT
p53-1	90±80	2.0±0.7
p53-2	21±3	1.2±0.3
p53-3	35±6	4.5±0.6
EmGFP-1	>500	5±1
EmGFP-2	>500	>500

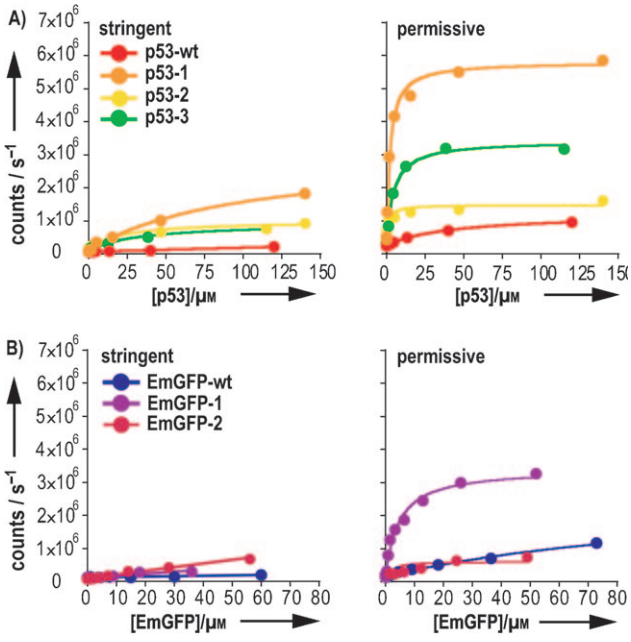

Figure 3. ReAsH binding trials. A) Measured fluorescence intensity (I) at 150 min for wild-type and mutant p53 variants under stringent (100 mM Tris·HCl, 75 mM NaCl, 3.5 mM TCEP, 1 mM EDT, pH 7.8 binding buffer) and permissive conditions (100 mM Tris·HCl, 75 mM NaCl, 3.5 mM TCEP, pH 7.8 binding buffer). B) Measured fluorescence intensity (I) at 150 min for wild-type and mutant EmGFP variants under stringent (100 mM Tris·HCl, 75 mM NaCl, 3.5 mM TCEP, 1 mM EDT, pH 7.8 binding buffer) and permissive conditions (100 mM Tris·HCl, 75 mM NaCl, 3.5 mM TCEP, pH 7.8 binding buffer). A concentration of 25 nM ReAsH was used for all trials. Three independent trials were performed for each bipartite display variant.

2.0±0.7 μM for p53-1, 1.2±0.3 μM for p53-2, and 4.5±0.6 μM for p53-3 (Table 1). These results represent the first example of intramolecular bipartite display within loop structures in a complex protein fold.

Although ReAsH bound with measurable affinity to both p53 loop variants and the terminal dithiol variant (p53-1) under stringent conditions, it did not bind detectably to either EmGFP-1 or EmGFP-2 under these same conditions (Figure 3B, stringent). Additionally, increasing the ReAsH concentration from 25 nM to 1 μM had no detectable effect under these conditions. We next sought to evaluate if the EmGFP variants could support ReAsH binding under permissive conditions (100 mM Tris, 75 mM NaCl, 3.5 mM TCEP, pH 7.8) suitable for in vitro assay development. Under these conditions, EmGFP-1 bound ReAsH (Figure 3B, permissive) with a K_d of 4.5±0.9 μM, although no ReAsH binding was detected for EmGFP-2 (Table 1). These results are consistent with our finding that bipartite designs require site flexibility for ReAsH coordination.

The utility of bipartite tetracysteine display will depend on how easily and predictably the split Cys-Cys motif can be introduced into proteins of complex structure. Our results suggest that intramolecular bipartite display is most successful when the Cys-Cys pairs are introduced into loops. In these constructs we observe a direct correlation between ReAsH affinity and site flexibility (Figure 4). In the case of p53, p53-2 has Cα B-factor values ranging from 23.3–39.9, and binds ReAsH better than p53-3, the relevant Cα B-factors of which range between 10.2–21.2. We note that the two p53 variants with Cys motifs within loops bound ReAsH with higher affinity than p53-1 in which the Cys motif is located at the protein N and C termini. We attribute this difference to the larger distance between the dicysteine motifs in p53-1. Our conclusion is further supported by the results obtained by using EmGFP-1 (Cα B-factor values of 15.6–26.3) and EmGFP-2 (4.9–18.1), whose relevant B-factors vary significantly. EmGFP-1 only bound ReAsH under permissive conditions, whereas EmGFP-2 showed no apparent affinity for ReAsH under either binding condition. These results are consistent with predictions based on relative bipartite site flexibilities.

It is interesting to consider our results in the context of related experiments by Krishnan and Gierasch, who recently described intramolecular bipartite display between adjacent β-strands in CRABP 1.[14] In these studies, several bipartite display mutants were constructed in the β-strands of CRABP 1 (St1-2, St1′-10, and St1-10) by using the same Cys-Xaa-Cys design. The authors showed successful coordination of FlAsH to these variants under relatively permissive conditions [50 mM 3-(*N*-morpholino)propanesulfonic acid (MOPS), pH 7.3, 150 mM NaCl, 1 mM TCEP, 0.1 mM ethanedithiol (EDT)], with $K_{d\,app}$ values greater than 43 μM. By using similar conditions (100 mM Tris, 75 mM NaCl, 3.5 mM TCEP, 0.1 mM EDT, pH 7.8), our EmGFP variants showed minimal affinity for ReAsH (K_d > 500 μM). We note that the reported Cα B-factors at the locations of Cys substitution in CRABP 1 (PDB: ID 1CBI) do not correlate directly with the affinity of these variants for FlAsH.[15] It might be that our conclusions are restricted to use of ReAsH, or that the considerably less-stringent binding conditions used in the experiments of experiments of Krishnan and Gierasch (tenfold lower EDT concentration) reduced the sensitivity for selective binding to native protein structure.

Herein, we have shown that successful bipartite tetracysteine display requires site flexibility and proximity to support high-affinity ReAsH binding under stringent conditions.[16] In the case of p53, we successfully designed intramolecular loop bipartite display variants that support ReAsH binding. Additionally, we found that despite our optimized design and geometry considerations, β-sheet bipartite display is not universally supportive of bisarsenical coordination. In our system, ReAsH binding was partially supported only in the absence of the dithiol competitor. From this study it is clear that ReAsH coordination preferentially favors proximal bipartite tetracysteine sites with flexible geometries, primarily unstructured

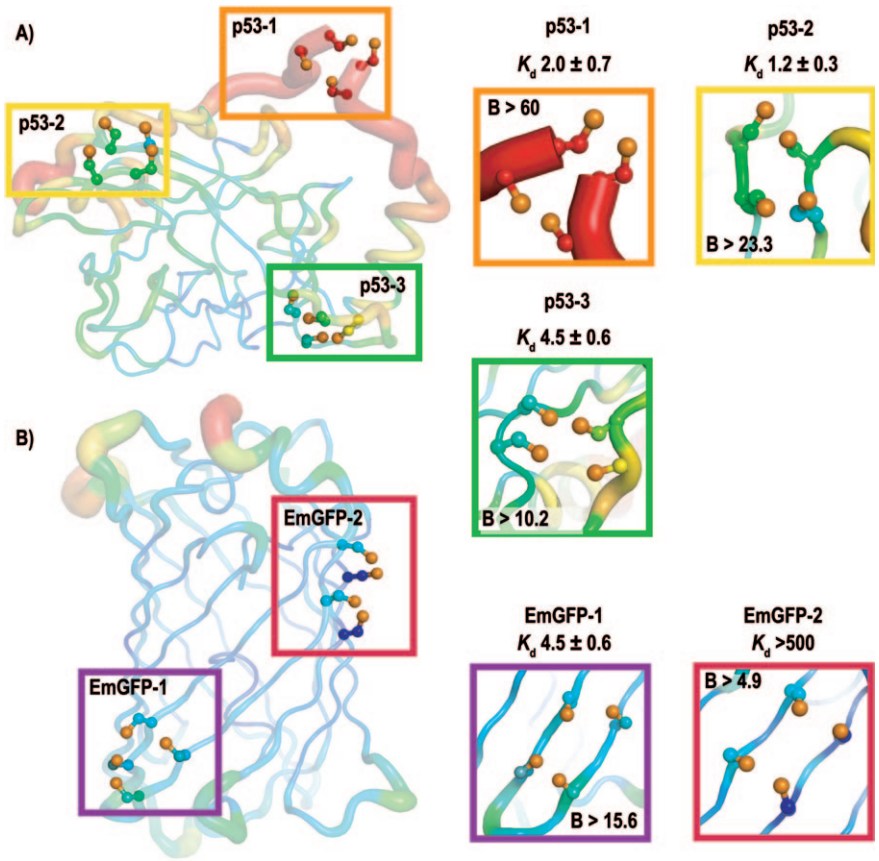

Figure 4. Site flexibility of bipartite display designs in A) p53 (1TUP[7]) and B) EmGFP (1EMA[10]) as analyzed by B-factor (B) values. The K_d values for the permissive conditions are listed for the respective bipartite display variants. B-Factors are represented by using the B-factor putty option in PyMol,[17] in which small B-factor values (in the range of 4 to 30) are defined by a narrow backbone with blue to green coloring, and larger B-factor values are defined by a wider backbone diameter with orange to red coloring (B-factor range of 40 to higher than 60).

loops that are adjacent in space. Furthermore, our findings suggest that β-sheet bipartite tetracysteine display is limited by the defined strand geometry and the rigidity of the secondary structure. Our results provide guiding principles for directing successful bipartite tetracysteine site designs to extend the utility of this technique in probing structural transitions and protein–protein interactions.

Experimental Section

Please see the Supporting Information for experimental details.

Acknowledgements

This work was supported by the NIH (GM 83257) and the National Foundation for Cancer Research. We thank Prof. Nathan W. Luedtke in the Department of Bioorganic Chemistry at the University of Zürich for the synthesis of ReAsH-EDT$_2$. We would also like to thank Sir Alan Fersht for providing the pRSET plasmid containing human p53 core domain with an N-terminal fusion of 6XHis/lipoamyl domain.

Keywords: fluorescence · protein engineering · ReAsH · secondary structure · tetracysteine

[1] B. A. Griffin, S. R. Adams, R. Y. Tsien, *Science* **1998**, *281*, 269–272.
[2] S. R. Adams, R. E. Campbell, L. A. Gross, B. R. Martin, G. K. Walkup, Y. Yao, J. Llopis, R. Y. Tsien, *J. Am. Chem. Soc.* **2002**, *124*, 6063–6076.
[3] N. W. Luedtke, R. J. Dexter, D. B. Fried, A. Schepartz, *Nat. Chem. Biol.* **2007**, *3*, 779–784.
[4] A. J. Levine, *Cell* **1997**, *88*, 323–331.
[5] A. J. Levine, J. Momand, C. A. Finlay, *Nature* **1991**, *351*, 453–456.
[6] J. M. Perez Canadillas, H. Tidow, S. M. V. Freund, T. J. Rutherford, H. C. Ang, A. R. Fersht, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 2109–2114.
[7] Y. Cho, S. Gorina, P. D. Jeffrey, N. P. Pavletich, *Science* **1994**, *265*, 346–355.
[8] T. Heck, M. Limbach, B. Geueke, M. Zacharias, J. Gardiner, H. P. Kohler, D. Seebach, *Chem. Biodiversity* **2006**, *3*, 1325–1348.
[9] Z. Yuan, T. L. Bailey, R. D. Teasdale, *Proteins Struct. Funct. Bioinf.* **2005**, *58*, 905–912.
[10] M. Ormö, A. B. Cubitt, K. Kallio, L. A. Gross, R. Y. Tsien, S. J. Remington, *Science* **1996**, *273*, 1392–1395.
[11] D. Ramadan, D. J. Cline, S. Bai, C. Thorpe, J. P. Schneider, *J. Am. Chem. Soc.* **2007**, *129*, 2981–2988.
[12] S. Bell, C. Klein, L. Muller, S. Hansen, J. Buchner, *J. Mol. Biol.* **2002**, *322*, 917–927.
[13] N. V. Visser, M. A. Hink, J. W. Borst, G. N. M. van der Krogt, A. Visser, *FEBS Lett.* **2002**, *521*, 31–35.
[14] B. Krishnan, L. M. Gierasch, *Chem. Biol.* **2008**, *15*, 1104–1115.
[15] C. Folli, V. Calderone, I. Ramazzina, G. Zanotti, R. Berni, *J. Biol. Chem.* **2002**, *277*, 41970–41977.
[16] F. Madani, J. Lind, P. Damberg, S. R. Adams, R. Y. Tsien, A. O. Graslund, *J. Am. Chem. Soc.* **2009**, *131*, 4613–4615.
[17] W. DeLano, *The PyMOL Molecular Graphics System*, DeLano Scientific, San Carlos, **2002**.

Received: April 3, 2009
Published online on June 16, 2009

