Published in final edited form as:

Chembiochem. 2009 July 6; 10(10): 1644–1647. doi:10.1002/cbic.200900207.

Bipartite tetracysteine display requires site flexibility for ReAsH coordination

Dr. Jessica L. Goodman^{[b],[c]}, Daniel B. Fried^{[b],[c]}, and Alanna Schepartz^{*,[a]}[Prof.]

^[a] Departments of Chemistry and Molecular, Cellular and Developmental Biology, Yale University 275 Prospect Street, New Haven, CT 06511 (USA)

^[b] Department of Chemistry, Yale University 275 Prospect Street, New Haven, CT 06511 (USA)

Abstract

Fluorescence is an essential tool for monitoring molecular interactions, cellular localization and pathway activation both *in vitro* and *in vivo*. The development of “pro-fluorescent” bisarsenical dyes in 1998 advanced this field in a significant way, as 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-EDT₂, Figure S1A) and 4,5-bis(1,3,2-dithiarsolan-2-yl)resorufin (ReAsH-EDT₂, Figure S1B) are non-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), permitting temporal fluorescent control.[1] Recently, we reported that the central Pro-Gly in the tetracysteine motif could be replaced with one or more folded proteins domains without a significant loss in bisarsenical affinity or brightness; we refer to this technique as ‘bipartite tetracysteine display’.[3] We showed, using a set of well characterized model systems, 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 evaluated in our initial report shared two features in common: 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 Cys₄ motif is introduced into loops and β -strands.

Keywords

fluorescence; protein engineering; ReAsH; secondary structure

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 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 whose thiols are located within flexible regions—as defined by crystallographic C α B-factors[9] greater than 20—bound ReAsH with high affinity and generated fluorescent complexes. Without exception, these regions are

Fax: (+1) 203-432-3486 alanna.schepartz@yale.edu.

^[c]These authors contributed equally to this work.

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

located within protein loops. By contrast, under the competitive binding conditions that characterize the intracellular milieu, no construct whose 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 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 C_{α} B-factor values in the crystal structures (PDB ID 1TUP[6,7] and 1EMA[10]). Using these criteria, five distinct p53 bipartite variants were designed (Figure 1A). Our first design, p53-1, mimics our initial bipartite designs with dicysteine motifs near the N and C-termini (P92C, L93C, P301C, G302C) and C_{α} B-factor values above 60. The other p53 variants incorporated dicysteine pairs within loops. These include p53-2 where the B-factors at the Cys-bearing positions (Y107C, G108C, D148C, S149C) range from 23.3 to 39.9, and p53-3 (S116C, G117C, T123C, C124C) with values between 10.2 and 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 to 17.3, whereas in p53-5 (A129C, L130C, R290C, K291C) they vary widely, 23.0 to 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 1B). 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 C_{α} 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 C_{α} B-factor values at Cys-bearing positions (D19C, D21C, K26C, S28C), that range between 15.6 and 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 to 18.1, 12.7 to 15.2, and 8.6 to 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 be expressed in *E. coli* and were purified to homogeneity (see 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), indicating 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, while 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, reasoning 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, providing 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 $90 \pm 80 \mu\text{M}$, p53-2 bound ReAsH with a K_d of $21 \pm 3 \mu\text{M}$ and p53-3 bound ReAsH with a K_d of $35 \pm 6 \mu\text{M}$ (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, $28 \pm 7 \mu\text{M}$ and $10 \pm 7 \mu\text{M}$, respectively.[3] We also quantified the affinity of these variants for ReAsH under permissive conditions (100 mM Tris, 75 mM NaCl, 3.5 mM 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 $2.0 \pm 0.7 \mu\text{M}$ for p53-1, $1.2 \pm 0.3 \mu\text{M}$ for p53-2, and $4.5 \pm 0.6 \mu\text{M}$ 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 $4.5 \pm 0.9 \mu\text{M}$, 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 α B-factor values ranging from 23.3 to 39.9, and binds ReAsH better than p53-3, whose relevant α B-factors range between 10.2 and 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 using EmGFP-1 (α B-factor values of 15.6 to 26.3) and EmGFP-2 (4.9 to 18.1), whose relevant B-factors vary significantly. EmGFP-1 only bound ReAsH under permissive conditions, while 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) using the same Cys-Xaa-Cys design. The authors showed successful coordination of FAsH to these variants under relatively permissive conditions (50 mM MOPS, pH 7.3, 150 mM NaCl, 1 mM TCEP, 0.1 mM EDT), with K_{dapp} values greater than 43 μM . 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 ($K_d > 500 \mu\text{M}$). We note that the reported α 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 FAsH.[15] It may be that our conclusions are restricted to use of ReAsH, or that the considerably less stringent binding conditions used in the experiments of Krishnan and Gierasch (10-fold 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 supports ideal coordination geometry[16] for ReAsH binding under stringent conditions. 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 loops which 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 in order to extend the utility of this technique in probing structural transitions and protein-protein interactions.

Experimental Section

Please refer to the supporting information for experimental details.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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 Zurich for the synthesis of ReAsH-EDT₂. 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.

References

1. Griffin BA, Adams SR, Tsien RY. Specific covalent labeling of recombinant protein molecules inside live cells. *Science*. 1998; 281(5374):269–72. [PubMed: 9657724]
2. Adams SR, Campbell RE, Gross LA, Martin BR, Walkup GK, Yao Y, Llopis J, Tsien RY. New biarsenical ligands and tetracysteine motifs for protein labeling in vitro and in vivo: synthesis and biological applications. *J Am Chem Soc*. 2002; 124(21):6063–76. [PubMed: 12022841]
3. Luedtke NW, Dexter RJ, Fried DB, Schepartz A. Surveying polypeptide and protein domain conformation and association with FAsH and ReAsH. *Nat Chem Biol*. 2007; 3(12):779–84. [PubMed: 17982447]
4. Levine AJ. p53, the cellular gatekeeper for growth and division. *Cell*. 1997; 88(3):323–31. [PubMed: 9039259]
5. Levine AJ, Momand J, Finlay CA. The p53 tumour suppressor gene. *Nature*. 1991; 351(6326):453–6. [PubMed: 2046748]
6. Canadillas JM, Tidow H, Freund SM, Rutherford TJ, Ang HC, Fersht AR. Solution structure of p53 core domain: structural basis for its instability. *Proc Natl Acad Sci U S A*. 2006; 103(7):2109–14. [PubMed: 16461916]
7. Cho Y, Gorina S, Jeffrey PD, Pavletich NP. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science*. 1994; 265(5170):346–55. [PubMed: 8023157]

8. Heck T, Limbach M, Geueke B, Zacharias M, Gardiner J, Kohler HP, Seebach D. Enzymatic degradation of beta- and mixed alpha,beta-oligopeptides. *Chem Biodivers*. 2006; 3(12):1325–48. [PubMed: 17193247]
9. Yuan Z, Bailey TL, Teasdale RD. Prediction of protein B-factor profiles. *Proteins*. 2005; 58(4):905–12. [PubMed: 15645415]
10. Ormo M, Cubitt AB, Kallio K, Gross LA, Tsien RY, Remington SJ. Crystal structure of the *Aequorea victoria* green fluorescent protein. *Science*. 1996; 273(5280):1392–5. [PubMed: 8703075]
11. Ramadan D, Cline DJ, Bai S, Thorpe C, Schneider JP. Effects of As(III) binding on beta-hairpin structure. *J Am Chem Soc*. 2007; 129(10):2981–8. [PubMed: 17311379]
12. Bell S, Klein C, Muller L, Hansen S, Buchner J. p53 contains large unstructured regions in its native state. *J Mol Biol*. 2002; 322(5):917–27. [PubMed: 12367518]
13. Visser NV, Hink MA, Borst JW, van der Krogt GNM, Visser A. Circular dichroism spectroscopy of fluorescent proteins. *Febs Letters*. 2002; 521(1-3):31–35. [PubMed: 12067720]
14. Krishnan B, Gierasch LM. Cross-Strand Split Tetra-Cys Motifs as Structure Sensors in a beta-Sheet Protein. *Chem Biol*. 2008; 15(10):1104–15. [PubMed: 18940670]
15. Folli C, Calderone V, Ramazzina I, Zanotti G, Berni R. Ligand binding and structural analysis of a human putative cellular retinol-binding protein. *J Biol Chem*. 2002; 277(44):41970–7. [PubMed: 12177003]
16. Madani F, Lind J, Damberg P, Adams SR, Tsien RY, Graslund AO. Hairpin structure of a biarsenical-tetracysteine motif determined by NMR spectroscopy. *J Am Chem Soc*. 2009; 131(13):4613–5. [PubMed: 19281235]
17. DeLano, W. The PyMOL Molecular Graphics System. DeLano Scientific; San Carlos: 2002.

Figure 1.

Bipartite display designs for p53 (with 10 native cysteine residues)[6,7] and EmGFP (with 2 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 further examined. 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.

Figure 2.

Circular dichroism (CD) spectra for p53 and EmGFP bipartite display variants. A) Wavelength-dependent CD spectra of 8.1 μ M p53-wt (wild type), 8.1 μ M p53-1, 8.1 μ M p53-2 and 8.1 μ M p53-3. B) Wavelength-dependent CD spectra of 25 μ M EmGFP-wt, 25 μ M EmGFP-1 and 25 μ M 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).

Figure 3.

ReAsH binding trials. A) Measured fluorescence intensity (I) at 150 minutes 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, 75mM NaCl, 3.5 mM TCEP, pH 7.8 binding buffer). B) Measured fluorescence intensity (I) at 150 minutes 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, 75mM 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.

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], where 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).

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 (μ M)	
	1 mM EDT	0 mM EDT
p53-1	90 $\pm$ 80	2.0 $\pm$ 0.7
p53-2	21 $\pm$ 3	1.2 $\pm$ 0.3
p53-3	35 $\pm$ 6	4.5 $\pm$ 0.6
EmGFP-1	> 500	5 $\pm$ 1
EmGFP-2	> 500	> 500