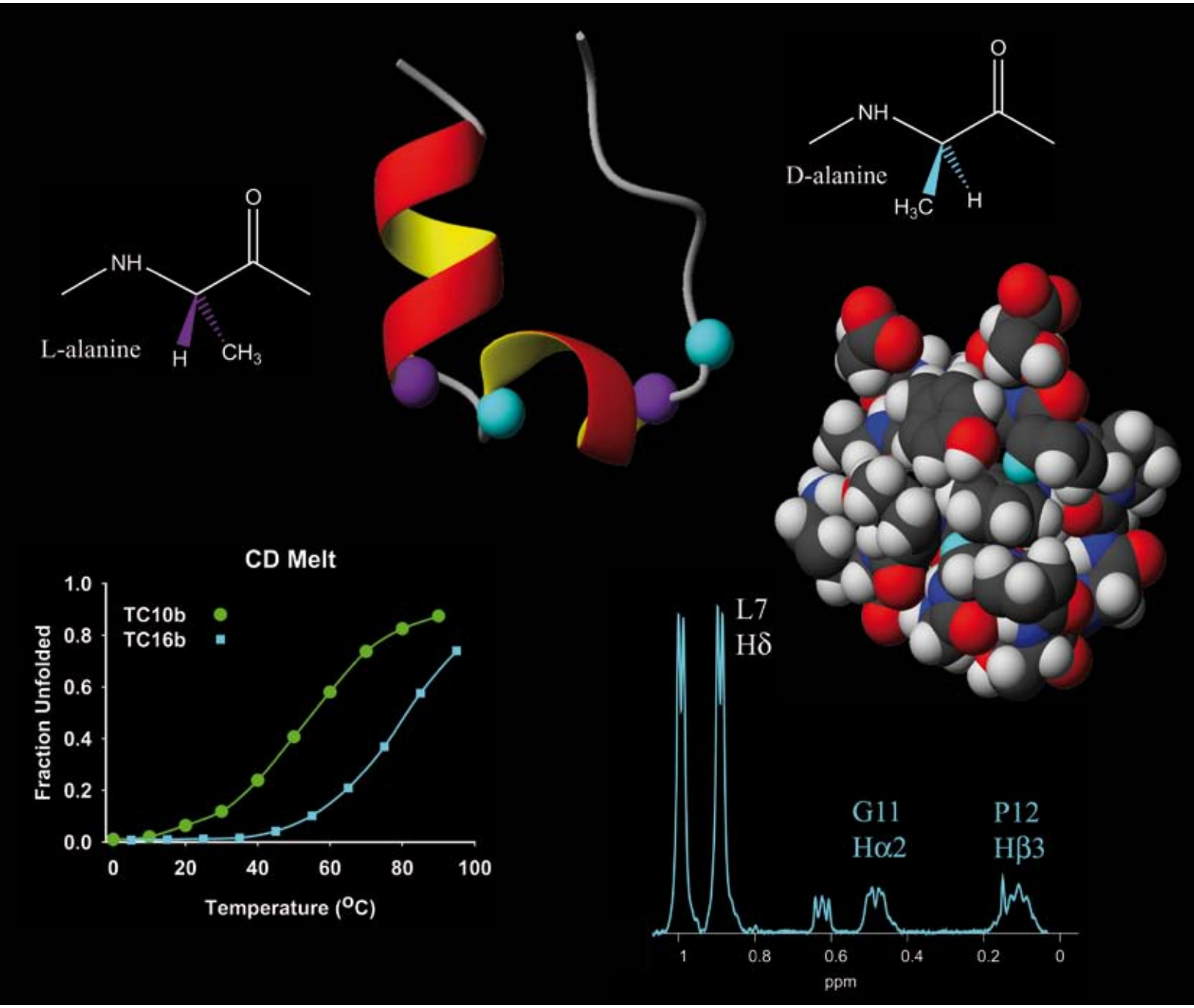
Organic & Biomolecular Chemistry

www.rsc.org/obc

Volume 6 | Number 23 | 7 December 2008 | Pages 4273–4468



ISSN 1477-0520

RSC Publishing

Hyperstable miniproteins: additive effects of D- and L-Ala mutations†

D. Victoria Williams, Bipasha Barua and Niels H. Andersen*

Received 18th August 2008, Accepted 25th September 2008
First published as an Advance Article on the web 15th October 2008
DOI: 10.1039/b814314e

The folding enantioselectivity for D-Ala *versus* L-Ala at one glycine site in the Trp-cage is 16 kJ mol^{-1}; judicious introductions of alanines of the correct chirality raises the melting temperature of this 20-residue fold to 83 °C.

Replacing glycines or residues with long side chains with alanine, or increasing the proline content of a sequence, are known strategies for increasing protein fold stability by decreasing the entropic advantage of unfolding.[1] The fold stabilizing effect of a Gly→Ala substitution at a site with no conformational strain effects has been measured as 2.6 kJ mol^{-1};[2] this net stabilization presumably includes the stabilizing $T\Delta S_U$ effect as well as any destabilizing effect associated with decreased backbone solvation.[3] That there are conformational strain effects is suggested by the observation of fold stabilization for L-AA to Gly mutations. These have been observed for some[1a,b,4] (but not all) instances of the introduction of Gly in place of L-AA residues that have a positive ϕ value in a native fold. Since glycines in native protein folds frequently occur in turn or loop positions where a positive ϕ torsion angle is required,[1b,4a,5] it is essential to consider the option of using Gly→D-Ala (rather than L-Ala) mutations for protein fold stabilization. Due to the difficulty of introducing unnatural D-AAs into proteins, there are very few examples of Gly→D-Ala mutations in proteins. Bang *et al.*[3] prepared Gly35→D-AA ubiquitin mutants by chemical ligation of synthetic fragments; no case of net stabilization was observed but the D-Ala mutant was more stable than the corresponding L-Ala form at the helix C-capping site with $\phi/\psi = +82/+10°$. Raleigh *et al.* [6] have provided three examples of this stabilizing Gly→D-Ala mutation in small protein domains that could be prepared by automated peptide synthesis, two of which were at helix C-capping sites.

Designed miniproteins present a testing ground for this stabilization strategy. The Trp-cage appeared ideally suited as there are three glycines with positive ϕ values—Gly10 $\phi/\psi = +100/+10$, Gly11 +57/−120, and Gly15 +80/+6°—with Gly10 and 15 appearing as the capping C′ position of the N-terminal α-helix and a short 3_{10} helix, respectively.[2] We now report the effects of D-Ala and L-Ala substitutions at the Gly sites in the Trp-cage: significant fold stabilization is observed with D-Ala at two sites and these mutations, in combination with stabilizing L-AA→L-Ala mutations at helical sites,[2,7] produce a 20-residue construct with a melting point (T_m) of 83 °C. Although β hairpins with T_m's in this range have been reported,[8] this is, to our knowledge, the highest T_m reported for small native or truncated protein-like motifs with a residue length of 36 or less[9] and lacking an extensive web of disulfide linkages.

Mutational stabilization of the two helical segments (in italics, *NLYIQWLKD*-GG-*PSS*-GRPPPS) of the original[1c] Trp-cage sequence (TC5b) increases fold stability. For example, an N1D mutation and L-Ala substitutions at each of the bold residue sites improve fold stability by 1.45 ± 0.35 kJ mol^{-1} per mutation.[2,7] The environments of the Gly residues that are the subject of the present investigation are illustrated in Fig. 1. At Gly10 and Gly15 (see Fig. 1a) the structure suggests that no van der Waals repulsion terms would result for a D-Ala substitution (Hα2→CH_3); however, Gly11 presents a more crowded circumstance.

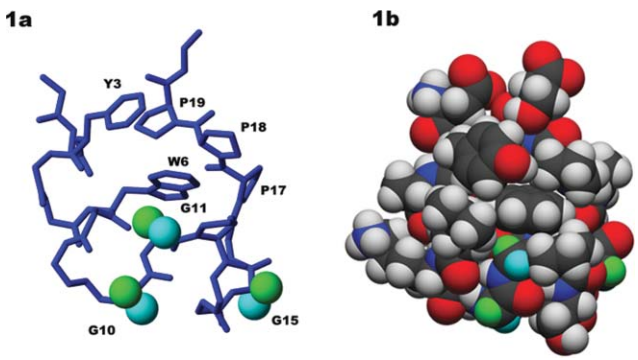


Fig. 1 The NMR structure of TC10b from the coordinates reported in Barua *et al.*[2] In panel a, the CH_2 units of the three glycines are labeled with the proton (Hα2) that is replaced by the methyl group in a D-Ala mutation shown in green; the L-Ala Me position is shown in cyan blue; most of the hydrogens as well as some side chains are deleted for clarity. Panel b, with the same Gly-CH_2 labeling scheme, is a CPK model that illustrates the crowding at Gly11 Hα2.

The close association of the G11-CH_2 with the indole ring (Fig. 1b), which yields a 3.5 ppm upfield shift for G11Hα2, with a much smaller shift (0.98 ppm) for the outwardly directed G11Hα3,[2] would appear to preclude the substitution of G11Hα2 by a methyl group without a significant alteration in the backbone geometry.

Both subtle changes in folded structure and the relative stabilities of Trp-cage constructs can be readily assessed[1c,2] by tracking NMR chemical shift deviations (CSDs) and their temperature dependence (see ESI†). Altered folded state geometries are indicated by changes in the CSD ratios; decreased fold stability is indicated by a greater, but uniform, loss in CSDs at increasing temperatures. The effects of L-Ala and D-Ala mutations were examined as shown in Table 1. NMR shift changes indicated that all three Gly→L-Ala mutations were highly destabilizing. Of these, the G10A mutant was the only one that displayed significant upfield shifts at G11Hα2 as well as P18α, β3. Upfield shifts at P18α, β3 (2.3 and 2.1 ppm, respectively) are the result of indole

Department of Chemistry, University of Washington, Seattle, WA 98195, USA. E-mail: andersen@chem.washington.edu

† Electronic supplementary information (ESI) available: Figures 3S–5S, experimental methods, methods for deriving melting thermodynamics, and chemical shift listings for analogs. See DOI: 10.1039/b814314e

Table 1 Sequences and thermodynamic stability data for Trp-cage mutants with D-AA in lower case[a], T_m values in °C

Peptide	Sequence	T_m cage (NMR)[b]	T_m (CD)	ΔT_m	ΔG_U^{280}/kJ mol^{-1}[c]
TC5b	NLYIQ WLKDG GPSSG RPPPS	42	42	−15	9.0[1c,2]
	Ac-AYAQ WLKDG GPSSG RPPPS	46	—	—	~10[7]
	Ac-AYAQ WLKD**A** GPSSG RPPPS	< 0	—	—	−2.4
TC10b	DAYAQ WLKDG GPSSG RPPPS	56	57	0	12.2[2,7]
(K8A)	DAYAQ WL**A**DG GPSSG RPPPS	62	61	6, 4	
(S13A)	DAYAQ WLKDG GP**A**SG RPPPS	63	61	7, 4	
(G10a)	DAYAQ WLKD**a** GPSSG RPPPS	> 70	72.5	–, 15.5	
(G15a)	DAYAQ WLKDG GPSS**a** RPPPS	61	62	5, 6	
				Σ = 29.5	
TC16b	DAYAQ WL**ADa** GP**ASa** RPPPS	>> 70	83	26	7.9

[a] All NMR shift and CD melts were recorded at pH 6.5–7 (20–50 mM potassium phosphate) with protein concentrations of *circa* 1 mM and 30 μM, respectively. TC10b is the reference for ΔT_m measures of relative stability. [b] The NMR T_m's given here are derived from a plot of T *versus* the sum of the CSDs for the L7α/G11Hα2/P18α,β3/P19δ2,δ3 sites.[2] [c] Derived from the NH exchange protection observed for the indole ring NH and the amide NH of Gly11.

ring current shielding and are diagnostic of the formation of a complete Trp-cage fold. The chemical shift deviations (CSDs) observed for the G10A mutant indicated $\chi_F = 0.16$–0.26 at 280K ($\Delta\Delta G_F \geq 12$ kJ mol^{-1}). The G15A construct failed to display upfield shifts at P18α,β3 ($\chi_{\text{full cage}} \leq 0.11$), but did display an upfield shift at G11Hα2 (CSD = −1.86 ppm), diagnostic of the formation of a half-cage structure.[2] The G11A mutant also failed to display upfield shifts at P18α,β3 ($\chi_{\text{full cage}} \leq 0.04$); but the data was difficult to interpret due to the formation of a folded state ($T_m \approx 5$ °C) which differed (see ESI†) from the classic Trp-cage fold. All of the Gly→L-Ala mutants also displayed significant, but not complete, loss of helix stability both by circular dichroism (CD) and by the decreased Hα CSDs in the N-terminal helical span.[7] Independent of whether we employ the CD melt, or the estimates of the extent of formation of the alternative fold, the G11A effect corresponds to $\Delta\Delta G_F \geq 11$ kJ mol^{-1}.

The Gly11→D-Ala mutation was also destabilizing ($\Delta T_m \approx$ −23 °C). This likely represents the effect of non-ideal dihedral angles associated with the G11Hα2 to methyl group mutation in the crowded environment (Fig. 1b). The observed ring current shift for Hα of D-Ala11 was −2.2 ppm, much further upfield than G10Hα3 (−0.98 ppm), the corresponding site of TC10b. This implies a rotation of residue 11 such that the D-Ala α proton is directed toward the indole ring with the bulky methyl group shifting so as to have the Cα–Cβ bond parallel to, rather than directed toward, the indole ring plane. Although the observed NOEs support the general features of the full Trp-cage fold, a detailed examination indicates conformational changes in the G10–R16 loop region that effect this residue 11 rotation.

D-Ala mutations at Gly10 and Gly15 provided easily measured fold stabilization in both CD (*vide infra*) and NMR melts (Table 1). For the Gly10 and Gly15 to D-Ala mutants, the CSDs for P18α/β3 and the G11-CH_2 were identical (±0.10 ppm) to those observed for TC10b, indicating no change in the backbone geometry or the spatial relationship between the Trp/Tyr rings and the sequence-remote, monitoring proton sites. This prompted us to prepare a single mutant containing all the stabilizing alanine substitutions, TC16b (DAYAQ WL**ADa** GP**ASa** RPPPS). At 57 °C, the individual CSDs observed for L7α/G11α2/P18α/P18β3/P19δ2/P19δ3 of TC16b indicated $\chi_F = 0.87 \pm 0.08$ (versus $\chi_F = 0.58 \pm 0.06$ for TC10b at 52 °C), placing the $\Delta\Delta G_U$ associated with the mutations at +5.6 ± 2.4 kJ mol^{-1}. A T_m value (83 °C) for TC16b could be obtained from the CD melt (Fig. 2). The CD melts observed for the full series of D-Ala mutants appear in Fig. 3. The enhanced fold stability of TC16b indicated by the NMR and CD melts was confirmed by measuring NH H/D exchange protection factors at pD 6.9 (280 K). As in the previous H/D exchange studies of TC5b and TC10b,[1c,2,7] G11H_N and the side chain NH of the indole ring display the largest protection factors (ΔG_U = 18.6 and 17.2 kJ mol^{-1}) with the backbone NHs of W6–D9 displaying smaller values ($\Delta G_U = 14.9 \pm 0.6$ kJ mol^{-1}). The W6–D9 protection factors report on helix stability, while protection at W6Hε1 and G11H_N reflects H-bonds that are specific to the cage structure. In direct comparisons with the protection factors measured for TC10b,[2,7] both sets of probes indicate an additional 5.7 kJ mol^{-1} stabilization by the alanines inserted in TC16b *versus* TC10b.

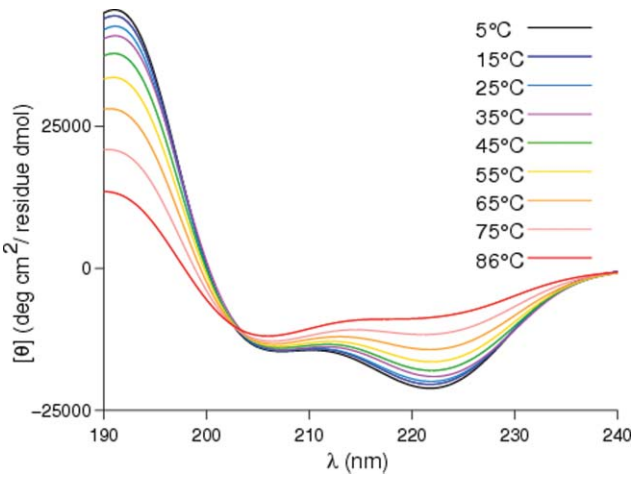


Fig. 2 CD spectra of TC16b at 5–86 °C; the slope of a plot of $[\theta]_{222}$ *versus* T (see Fig. 3) is still increasing at 80 °C. An additional melting study in 20 vol-% glycol increased the range of temperatures to −3→+95 °C, and displayed a maximal slope over the 80–95 °C increment.

In prior studies of Trp-cage species, ΔG_U^{280} estimates have been derived based on correlations[2,7] between ΔT_m values (based on NMR melting data) and ΔG_U^{280} from NH exchange protection factors. Significant changes in the relative magnitudes of ΔS_U and ΔC_{pU} in a series of mutants could preclude the use of a $\Delta T_m/\Delta G_U^{280}$ correlation; as a minimum criterion for meaningful

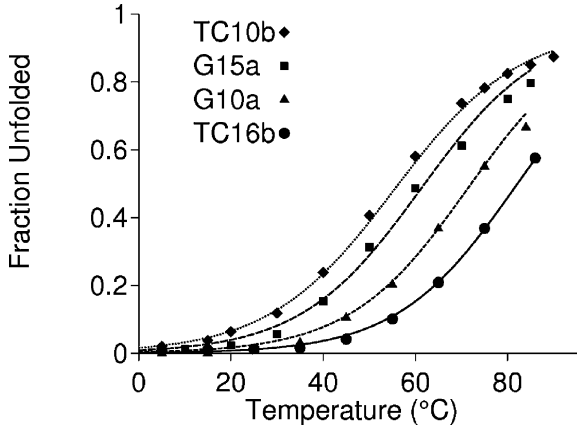


Fig. 3 CD melts of a series of D-Ala mutations within a Trp-cage miniprotein. In addition to the Gly10 and Gly15 to D-Ala mutations examined individually in the TC10b construct, TC16b has L-Ala insertion for Lys8 and Ser13. The points are experimental data, the lines are "sigmoidal fits" and correspond to expectations for $\Delta C_{pU} = 0$ (see ESI†). The melts as $[\theta]_{222}$ *versus* T plots appear in Fig. 3S (see ESI†).

T_m comparisons, the melting curves should be parallel in the $\chi_U = 0.20$–0.80 span. This condition is met by the melting curves appearing in Fig. 3. The ΔT_m data is summarized in Table 1. Full thermodynamic fits of the CD melts of the D-Ala mutants uniformly place ΔC_{pU} in the –0.30 to –0.74 kJ K^{-1} mol^{-1} range (ESI†); values from prior studies of Trp-cage species are TC10b = –0.2 ± 0.1 and TC5b = +0.3 ± 0.1 kJ mol^{-1} K^{-1}, the latter[10] from calorimetric as well as CD data. These results validate the use of ΔT_m data to address the additivity of mutational effects and cooperativity issues. In both instances of Gly10→D-Ala mutations, the thermodynamic analysis indicates a decrease in ΔS_U.

The fold-stabilizing Gly→D-Ala substitution effects observed in the Trp-cage ($\Delta\Delta G_U = 1.6$ at Gly15 and 4.0 kJ mol^{-1} at Gly10) are comparable to those observed previously in UBA and at the G24 site of NTL9 (2.5 and 5.4 kJ mol^{-1})[6] but not as large as that observed for the G34→D-Ala mutation (7.8 kJ mol^{-1})[6] in NTL9. Thus, in three systems, as long as there are no apparent steric problems introduced by the Hα2→methyl mutation, fold stabilization results for Gly→D-Ala mutations. In this regard, the absence of net fold stabilization at the helix C′ position (Gly35) of ubiquitin is puzzling: Bang *et al.*[3] report a 'conformational preference', D-Ala over L-Ala, of 4.2 kJ mol^{-1}. At the Gly10 and Gly15 sites of TC10b the conformational preference is substantially larger, 16 and ≥12 kJ mol^{-1}, respectively. The value at Gly15 is somewhat compromised by the evidence for the formation of an alternative fold. The Trp-cage Gly10 site shares the greatest analogy with ubiquitin Gly35, both are fully solvent exposed sites at the C′ location of helices with similar local conformations: $\phi/\psi = +100/+10°$ *versus* $+82/+10°$. The difference in the D-Ala *versus* L-Ala conformational preference, 16 kJ mol^{-1} for the Trp-cage *versus* 4.2 kJ mol^{-1} for the ubiquitin site, defies explanation. Apparently, there are additional factors that influence the D-/L-conformational preferences that have not been elucidated as yet.

An examination of the ΔT_m values associated with each L- or D-alanine insertion listed in Table 1 indicates that there is nearly complete additivity of fold stabilization over six sites. This provides additional evidence for a two-state folding scenario for the Trp-cage. In addition, these sites can now be used for the introduction of pharmacophore units from amino acids with either natural or altered side chains.[11] We expect that the Gly→D-Ala mutation strategy will be a powerful one for miniprotein fold optimization and that the hyperstable Trp-cage reported herein will be useful for evaluating fold stabilizations associated with side chain/side chain interactions in the Trp-cage as well as serving as a stereospecific scaffold for drug discovery[11] against numerous targets.

Acknowledgements

The reported research was supported by NIH grant GM59658.

Notes and references

1 (*a*) B. W. Matthews, H. Nicholson and W. J. Becktel, *Proc. Natl. Acad. Sci. U. S. A.*, 1987, **84**, 6663–6667; (*b*) H. Nicholson, E. Soderlind, D. E. Tronrud and B. W. Matthews, *J. Mol. Biol.*, 1989, **210**, 181–193; (*c*) J. W. Neidigh, R. M. Fesinmeyer and N. H. Andersen, *Nat. Struct. Biol.*, 2002, **9**, 425–430.
2 B. Barua, J. C. Lin, D. V. Williams, P. Kummler, J. Neidigh and N. H. Andersen, *Protein Eng., Des. Sel.*, 2008, **21**, 171–185.
3 D. Bang, A. V. Gribenko, V. Tereshko, A. A. Kossiakoff, S. B. Kent and G. I. Makhatadze, *Nat. Chem. Biol.*, 2006, **2**, 139–143.
4 (*a*) J. Kim, S. R. Brych, J. Lee, T. M. Logan and M. J. Blaber, *Mol. Biol.*, 2003, **328**, 951–961; (*b*) P. A. Karplus, *Protein Sci.*, 1996, **5**, 1406–1420.
5 S. Hovmöller, T. Zhou and T. Ohlson, *Acta Crystallogr., Sect. D*, 2002, **58**, 768–776.
6 B. Anil, B. Song, Y. Tang and D. P. Raleigh, *J. Am. Chem. Soc.*, 2004, **126**, 13194–13195; B. Anil, R. Craig-Schapiro and D. P. Raleigh, *J. Am. Chem. Soc.*, 2006, **128**, 3144–3145.
7 J. C. Lin, B. Barua and N. H. Andersen, *J. Am. Chem. Soc.*, 2004, **126**, 13679–13684.
8 R. M. Fesinmeyer, F. M. Hudson and N. H. Andersen, *J. Am. Chem. Soc.*, 2004, **126**, 7238–7243.
9 J. C. McKnight, D. S. Doering, P. T. Matsudaira and P. S. Kim, *J. Mol. Biol.*, 1996, **280**, 126–134.
10 W. W. Streicher and G. I. Makhatadze, *Biochemistry*, 2007, **46**, 2876–2880.
11 The Trp-cage has already been used as a scaffold for pharmacophore display: Y. Liu, Z. Liu, E. Androphy, J. Chen and J. D. Baleja, *Biochemistry*, 2004, **43**, 7421–7433; R. E. Herman, D. Badders, M. Fuller, E. G. Makienko, M. E. Houston, Jr., S. C. Quay and P. H. Johnson, *J. Biol. Chem.*, 2007, **282**, 9813–9824.