CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Nature Precedings

A Hyperstable Miniprotein: Additive Effects of D- and L-Ala Substitutions

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

Department of Chemistry, University of Washington, Seattle, WA 98195

E-mail: andersen@chem.washington.edu

Replacing glycines or residues with long sidechains 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[1-4]. The $T\Delta S_U$ effect for Gly → Ala substitution is expected to be fold stabilizing by 2.6 kJ/mol[1,4]; this net stabilization presumably includes any destabilizing effect associated with decreased backbone solvation.[5] The potential for conformational strain effects is suggested by the observation of fold stabilization, in some[1,2,6,7] (but not all) instances, of the introduction of Gly in place of L-AA residues that have a positive Φ 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,[2,6,8] it is essential to consider the option of using Gly → D-Ala (rather than L-Ala) mutations for fold stabilization. Due to the difficulty of introducing unnatural D-AA's into proteins, there are very few examples of Gly → D-Ala mutations in proteins. Bang et al.[5] 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 a helix C-capping site with $\phi/\psi = +82/+10°$. Raleigh[9] has provided three examples in small protein domains that could be prepared by automated peptide synthesis, two of which were helix C-capping sites. At the three sites examined by Raleigh, net fold stabilization was observed for Gly → D-Ala mutations. 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' positions of the N-terminal α-helix and a short 3_{10} helix, respectively.[3,4] 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[4,10], 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[11]; this is, to our knowledge, the highest T_m reported for small native or truncated protein-like motifs of 36 or less residue length[12] and lacking an extensive web of disulfide linkages.

The relative stabilities of Trp-cage constructs can be readily measured[4] by NMR and CD melts (see **Supplementary Methods** online). Mutational stabilization of the two helical segments (underlined) of the original[3] Trp-cage sequence (TC5b, N**L**Y**I**QWL**K**D-GG-P**S**S-GRPPPS T_m = 42 °C) increases fold stability: an N1D mutation and L-Ala substitutions at each of the bold residue sites improve fold stability by 1.45 ± 0.35 kJ/mol per mutation[4,10]. 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. 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,[4] would appear to preclude the substitution of G11Hα2 by a methyl group without a significant alteration in the backbone geometry.

The effect of an L-Ala mutation at the first glycine was measured for the Ac-AYAQ WLKDG GPSSG RPPPS sequence; the remaining D- and L-Ala mutation were examined in TC10b (DAYAQ WLKDG GPSSG RPPPS, T_m = 57°C)[4]. 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 ring current shielding and diagnostic of formation of a complete Trp-cage fold. The chemical shift deviations (CSDs) observed for the G10A mutant indicated χ_F = 0.16 – 0.26 at 280K ($\Delta\Delta G_F \geq$ 12 kJ/mol). The G15A construct failed to display upfield shifts at P18α,β3 ($\chi_{full\ cage} \leq 0.11$), but did display an upfield shift at G11Hα2 (CSD = -1.86 ppm), diagnostic of the formation of a half-cage structure.[4] The G11A mutant also failed to display upfield shifts at P18α,β3 ($\chi_{full\ cage} \leq 0.04$); but the data was difficult to interpret due the formation of a folded state ($T_m \approx$ 5 °C) which differed (see Supplementary Information) 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 CD and by the decreased Hα CSDs in the N-terminal

helical span.[10] 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.

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. The observed NOEs support the general features of the full Trp-cage conformation, but a detailed examination, and preliminary NMR structure ensemble calculations, require conformational changes in the G10-R16 loop region that effect this residue 11 rotation.

D-Ala mutations at G10 and G15 provided easily measured fold stabilization in both CD and NMR melts (**Table 1**, **Fig. 2**). The melts as $[\theta]_{222}$ versus T plots appear in **Supporting Fig. 2** online. 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**A**D**a** GP**A**S**a** 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.

A T_m value (83 °C) for TC16b could be obtained from the CD melt (**Fig. 2**, see also **Supporting Fig. 1** online). 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,[3,4,10] G11H_N and the sidechain NH of the indole ring display the largest protection factors (ΔG_U = 18.6 and 17.2 kJ/mol) with the backbone NHs of W6 – D9 displaying

smaller values (ΔG_U = 14.9 ± 0.6 kJ/mol). 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 comparisons with the protection factors measured for TC10b[4,10], both sets of probes indicate an additional 5.7 kJ/mol stabilization by the alanines inserted in TC16b versus TC10b.

In prior studies of Trp-cage species, ΔG_U^{280} estimates have been derived based on correlations [4,10] 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 ΔT_m / ΔG_U^{280} correlation; as a minimum criterion for meaningful T_m comparisons, the melting curves should be parallel in the χ_U = 0.20 – 0.80 span. This condition is met by the melting curves appearing in Fig. 2. 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 range (**Supplementary Methods** online); values from prior studies of Trp-cage species are – TC10b (-0.2 ± 0.1) and TC5b (+0.3 ± 0.1 kJ/mol-K), the latter[13] 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 4.4 kJ/mol at Gly10) are comparable to those observed previously in UBA and at the G24 site of NTL9 (2.5 and 5.4 kJ/mol)[9] but not as large as that observed for the G34 → D-Ala mutation (7.8 kJ/mol)[9] 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.[5] report a 'conformational preference', D-Ala over L-Ala, of 4.2 kJ/mol. At the Gly10 and Gly15 sites of TC10b the conformational preference is substantially larger, 16 and ≥12 kJ/mol, respectively. The value at Gly15 is somewhat compromised by the evidence for the formation of an alternative fold. Trp-cage Gly10 shares the greatest analogy with ubiquitin Gly35, both are fully solvent exposed sites at the C'

location of helices with similar local conformations: ϕ/ψ = +100/+10° *versus* +82/+10°. The difference in the D-Ala versus L-Ala conformational preference, 16 kJ/mol for the Trp-cage versus 4.2 kJ/mol for the ubiquitin site, beggars 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 sidechains so long as the required Cα chirality is maintained. Trp-cage constructs have already been used as a scaffold for pharmacophore display [14,15]: 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 sidechain/sidechain interactions in the Trp-cage as well as serving as a stereospecific scaffold for drug discovery against numerous targets.

ACKNOWLEDGMENTS.

We wish to thank Brandon Kier for developing the programs for performing thermodynamic fits to CD data. The reported research was supported by NIH grant GM59658.

Supplementary information available -- Full descriptions of peptide synthesis and characterization methods including the use of NMR and CD melts to obtain fold population estimates, validation of sigmoidal fitting procedures and the extraction of thermodynamic quantities from CD melts, four supporting figures, tabulated chemical shifts of all protons in residues 7, 10, 11, 12, and 16 – 19 for TC10b and the newly reported D-Ala and L-Ala substitution species and comments concerning implied structural differences.

REFERENCES

1. Matthews, B.W., Nicholson, H., Becktel W.J. *Proc. Natl. Acad. Sci. USA* **84**, 6663-6667 (1987).

2. Nicholson, H., Söderlind, E., Tronrud, D. E., Matthews, B.W. *J. Mol. Biol.* **210**, 181-193 (1989).

3. Neidigh, J.W., Fesinmeyer, R.M., Andersen, N.H. *Nature Struc. Biol*. **9**, 425-430 (2002).

4. Barua, B. *et al. PEDS* **21**, 171-185 (2008).

5. Bang, D. *et al*. *Nature Chem. Bio.* **2**, 139-143 (2006).

6. Kim, J., Brych, S.R., Lee, J., Logan, T.M., Blaber, M. J. *Mol. Biol.* **328**, 951-961 (2003).

7. Karplus, P.A. *Protein Sci.* **5**, 1406-1420 (1996).

8. Hovmöller, S., Zhou, T., Ohlson, T. *Acta Cryst. D* **58**, 768-776 (2002).

9. Anil, B., Song, B., Tang, Y., Raleigh, D.P. *J. Am. Chem. Soc.* **126**, 13194-13195 (2004); Anil, B., Craig-Schapiro, R., Raleigh, D.P. *J. Am. Chem. Soc.* **128**, 3144-3145 (2006).

10. Lin, J.C., Barua, B., Andersen, N.H. *J. Am. Chem. Soc.* **126**, 13679-13684 (2004).

11. Fesinmeyer, R. M., Hudson F. M., Andersen, N. H. *J. Am. Chem. Soc.* **126**, 7238-7243 (2004).

12. McKnight, J.C.; Doering, D. S.; Matsudaira, P.T.; Kim, P. S. *J. Mol. Biol.*, **280**, 126-134 (1996).

13 Streicher, W.W. and Makhatadze, G.I. *Biochem.* **46,** 2876-2880 (2007).

14. Liu, Y., Liu, Z., Androphy, E., Chen, J., Baleja, J.D. *Biochem*. **43**, 7421-7433 (2004).

15. Herman, R.E. *et al. J. Biol. Chem.* **282**, 9813-9824 (2007).

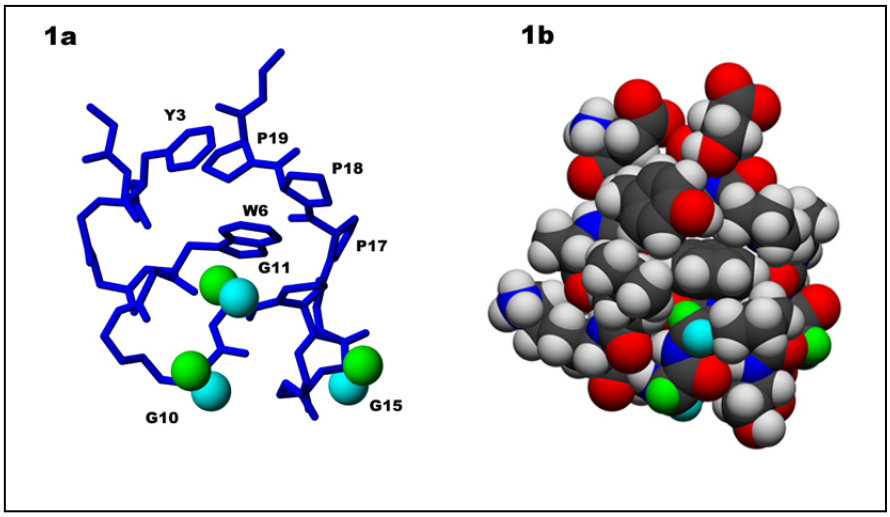


Figure 1. The NMR structure of TC10b from the coordinates reported in Barua *et al.*[4] In panel a, the CH_2 units of the three glycines are labeled with the proton (Hα2) that is replaced by the methyl group in a D-Ala mutation shown in green; the L-Ala Me position is shown in cyan blue; most of the hydrogens as well as some sidechains 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.

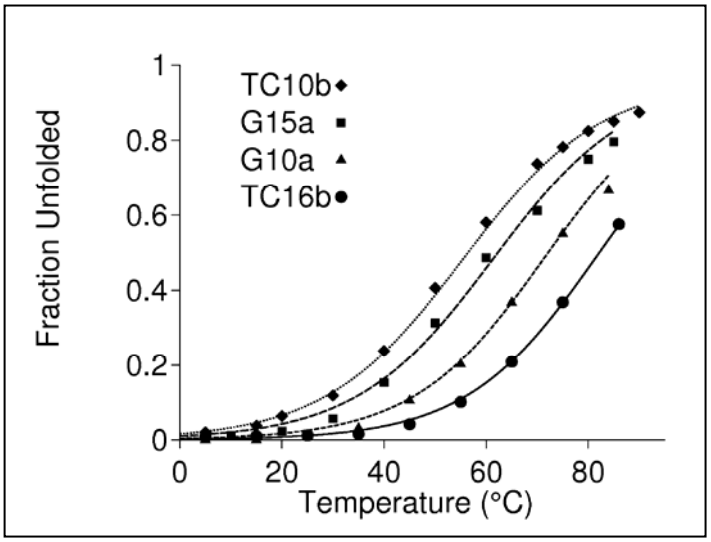


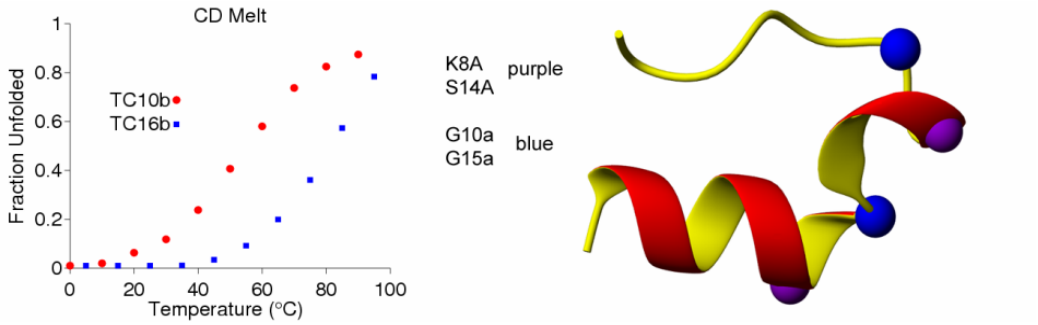
Figure 2. CD melts of a series of D-Ala mutations within a Trp-cage miniprotein. In addition to the Gly10 and Gly15 to D-Ala mutations examined individually in the TC10b construct, TC16b has L-Ala insertion for Lys8 and Ser13. The points are experimental data, the lines are "sigmoidal fits" and correspond to expectations for $\Delta C_{pU} = 0$ (**Supplementary Methods** online).

Table 1. Thermodynamic stability data for Trp-cage mutants [a]

protein	T_m cage (NMR) [b]	T_m (CD)	ΔT_m	ΔG_U^{280}(kJ/mol) [c]
TC5b	42	42	-15	9.0 [3,4]
TC10b	56	57	0	12.2 [4,10]
(K8A)	62	61	6, 4	
(S13A)	63	61	7, 4	
(G10a)	>70	72.5	-, 15.5	
(G15a)	61	62	5, 6	
			Σ = 29.5	
TC16b	>>70	83	26	17.9

[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.[4] [c] Derived from the NH exchange protection observed for the indole ring NH and the amide NH of Gly11.

Table of Contents artwork inserted here



Abstract

The effects of alanine substitutions in each helical segment of the structure, and Gly → D-Ala mutations at sites where glycines have positive phi angles in the Trp-cage miniprotein are reported. The effects of the stabilizing mutation were additive, yielding a 20-residue construct (T_m = 83°C). Gly → L-Ala substitutions were uniformly destabilizing ($\Delta\Delta G_F \geq 11$ kJ/mol): the preference for a D-Ala can be as large as 16 kJ/mol. Glycine to D-Ala mutations are validated as a strategy for the design of hyperstable miniprotein scaffolds suitable for stereospecific pharmacophore display.