

Published in final edited form as:

*Chem Phys.* 2013 August 30; 422: . doi:10.1016/j.chemphys.2013.01.021.

## Using D-Amino Acids to Delineate the Mechanism of Protein Folding: Application to Trp-cage

Robert M. Culik<sup>1</sup>, Srinivas Annavarapu<sup>2</sup>, Vikas Nanda<sup>2,\*</sup>, and Feng Gai<sup>3,\*</sup>

<sup>1</sup>Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA 19104

<sup>2</sup>Department of Biochemistry, University of Medicine and Dentistry of New Jersey, Piscataway, NJ 08854

<sup>3</sup>Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104

### Abstract

Using the miniprotein Trp-cage as a model, we show that D-amino acids can be used to facilitate the delineation of protein folding mechanism. Specifically, we study the folding-unfolding kinetics of three Trp-cage mutants where the native glycine residue near the C-terminus of the  $\alpha$ -helix is replaced by a D-amino acid. A previous study showed that these mutations increase the Trp-cage stability, due to a terminal capping effect. Our results show that the stabilizing effect of D-asparagine and D-glutamine originates almost exclusively from a decrease in the unfolding rate, while the D-alanine mutation results in a similar decrease in the unfolding rate, but it also increases the folding rate. Together, these results support a folding mechanism wherein the  $\alpha$ -helix formation in the transition state is nucleated at the N-terminus, whereas those long-range native interactions stabilizing this helix are developed at the downhill side of the folding free energy barrier.

### 1. Introduction

Loops and turns in proteins not only ensure connectivity of the polypeptide chain, but they also, in many cases, play an important role in directing the process of folding. A survey of the available protein structures in the Protein Data Bank (PDB) indicated that the most frequently found amino acids in loops are glycine (Gly), serine (Ser), aspartate (Asp), asparagine (Asn), and proline (Pro).<sup>1</sup> In particular, Gly is frequently found to be at the C-terminus of  $\alpha$ -helices in proteins, serving as a secondary structure breaker and directing the main chain in a new direction. This is due to the fact that Gly can extensively populate backbone conformations that are normally forbidden for other L-amino acids, as seen in the Ramachandran plot. For the same reason, however, there is a relatively large conformational entropy penalty associated with restricting a Gly residue in a small region of phi and psi conformations. Therefore, Gly has recently become a popular target in protein design to improve the stability of proteins of interest via mutations that could alleviate this entropy penalty.<sup>2</sup> For example, it has been shown that Pro, which has low backbone conformational

© 2013 Elsevier B.V. All rights reserved.

\*To whom correspondence should be addressed: gai@sas.upenn.edu or nanda@cabm.rutgers.edu; Phone: 215-573-6256; Fax: 215-573-2112.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

entropy, can in some cases replace Gly and, as a result, further stabilize the native fold.<sup>3</sup> A different, but perhaps more elegant and more effective approach, is to replace Gly with various D-amino acids,<sup>2,4–6</sup> the enantiomeric relative of L-amino acids, which can occupy conformations on the Ramachandran plot that otherwise only glycine could. For example, D-alanine (D-Ala), which has no sidechain rotamer preferences and hence a smaller entropy penalty upon folding, has been used, in a number of cases, to increase the stability of proteins,<sup>4,6</sup> to reconstitute an active and specific ion channel,<sup>5</sup> and to create novel folds.<sup>7</sup>

In a more recent study,<sup>8</sup> Rodriguez-Granillo *et al.* have shown that D-Ala, D-Asn, and D-Gln all show a stabilizing effect, although to a different extent, on the folding free energy of the mini-protein Trp-cage (Figure 1), when replacing the Gly residue at position 10 (Specifically, the TC5b variant, sequence: NLYIQWLKDGGPSSGRPPPS, was used). Interestingly, both D-Ala and D-Gln were found to significantly stabilize the native fold, raising the thermal melting temperature ( $T_m$ ) of Trp-cage by about 20 °C, whereas D-Asn only increased the  $T_m$  by about 10 °C. This observation was interpreted as due to a combination of effects, including the difference in sidechain-backbone hydrogen-bonding interaction ability and conformational entropy of these D-amino acids. The hydrogen bonding ability of the D-Gln sidechain was found to be greater than that of D-Asn, while the sidechain conformational entropy penalties for both D-Gln and D-Asn are larger than for D-Ala. While this study demonstrated the effectiveness, especially when placed at the C-terminus of  $\alpha$ -helices, of several D-amino acids to alleviate the intrinsic conformational limitations of naturally occurring L-amino acids, and offered a thermodynamic rationalization of the stabilizing effect of these D-amino acids on the Trp-cage fold, it was unable to yield a mechanistic and dynamic view of how this is achieved, i.e., through an increase in the folding rate, a decrease in the unfolding rate, or even a change in the folding mechanism. Herein, we use a laser-induced temperature-jump (T-jump) infrared (IR) technique to study the conformational dynamics of several previously characterized D-amino acid mutants of TC5b, including 5b-G10DA, 5b-G10DN, and 5b-G10DQ, aiming to elucidate the kinetic role of these D-amino acids. In addition, Culik et al. have recently put forward a folding mechanism for Trp-cage, based on extensive spectroscopic studies and amino acid mutations,<sup>9</sup> in which the transition state of the main folding event, i.e., the formation of the cage structure, involves the consolidation of the  $\alpha$ -helix, from amino acids 2–9. Since the abovementioned D-amino acids are expected to provide further stabilization to the  $\alpha$ -helix,<sup>4,6</sup> studying the folding kinetics of these  $\alpha$ -D-amino acid mutants will help improve our understanding of the folding mechanism of this extremely popular protein folding model.<sup>10–48</sup>

## 2. Materials and methods

Peptides were synthesized using standard solid-phase Fmoc methods at the Tufts University Core Facility (<http://www.tucf.org>). All peptides were acetylated at the N-terminus and amidated at the C-terminus. Peptides were purified using reversed-phase high-performance liquid chromatography (HPLC), and peptide identities were verified using mass spectrometry. Trifluoroacetic acid (TFA) removal and H-D exchange were achieved by multiple rounds of lyophilization. Fourier transform infrared (FTIR) spectra were collected on a Magna-IR 860 spectrometer (Nicolet, WI) using a home made, two-compartment CaF<sub>2</sub> sample cell of 56  $\mu$ m. The details of the T-jump IR setup have been described elsewhere.<sup>49</sup> For both static and time-resolved IR measurements, the peptide sample was prepared by directly dissolving lyophilized solids in 20 mM phosphate D<sub>2</sub>O buffer (pH\* 7) and the final peptide concentration was between 1–2 mM.

### 3. Results and discussion

A number of studies have shown that the short  $3_{10}$ -helix in the Trp-cage fold is much less stable than the global cage structure<sup>9,22,48</sup> and its folding can proceed independently and with a faster rate.<sup>9</sup> Therefore, in the following discussion, we focus only on the effect of the corresponding D-amino acid mutations on the two-state folding and unfolding kinetics of the cage, which consists of several key structural elements: an  $\alpha$ -helix, a polyproline stretch, and a hydrophobic core that encloses a tryptophan residue (Figure 1). In addition, all the thermodynamic parameters used to analyze the T-jump-induced relaxation kinetics are taken from a previous study<sup>8</sup> and are summarized in Table 1.

#### 5b-10GDQ

As indicated (Table 1), substitution of Gly10 in TC5b with D-Gln significantly improves Trp-cage stability, increasing its  $T_m$  from 45.9 to 69.3 °C. It has been hypothesized and subsequently computationally verified that this stabilizing effect of D-Gln arises from, among other factors, its ability to form hydrogen bonds between the backbone amine of D-Gln and the backbone carbonyl of Leu7, and another between the backbone carbonyl of D-Gln and the backbone amine of Ser13. From a kinetic perspective, however, the higher stability of 5b-10GDQ, compared to that of the wild type, could result from either a faster folding rate or a slower unfolding rate, or both. Since the CD spectrum of 5b-10GDQ indicates that it has a higher helical content than the wild type,<sup>8</sup> it is reasonable, based on the notion that the  $\alpha$ -helix is populated in the folding transition state of the cage,<sup>9</sup> to assume that the  $\alpha$ -helix stabilizing effect of D-Gln would lower the free energy of the transition state and, as a result, speed up the rate of folding. As shown (Figure 2), however, the folding rate of 5b-10GDQ is, within our experimental uncertainties, identical to that of the wild type, indicating that the stabilizing interactions stemming from the D-Gln mutation only start to develop at the downhill side of the major folding free energy barrier. In other words, these interactions act primarily to decrease the free energy of the folded state, making the protein unfold slower as a result of an increased unfolding free energy barrier, as observed (Figure 2). This picture is entirely consistent with earlier studies<sup>9,13,17,20,24,25,30,48</sup> that indicated that the developments of several key structural elements that involve residues near this part of the sequence, e.g., the formation of the D9-R16 salt-bridge and the burial of the Trp6 sidechain, occurs after the folding transition state. Furthermore, this result is in accord with the observation that stabilizing sidechain-sidechain and/or sidechain-backbone interactions can decrease the unfolding rate of  $\alpha$ -helices in isolation.<sup>50</sup>

#### 5b-10GDN

Similar to D-Gln, D-Asn was also found to form backbone-backbone hydrogen bonds with Leu7 and Ser13. However, its stabilizing effect is less pronounced compared to D-Gln (Table 1). For example, the  $T_m$  of 5b-10GDN was determined to be 56 °C, only 10 °C higher than the wild type. This improvement in stability has been attributed to the difference in backbone hydrogen bonding ability of D-Asn, in comparison to that of D-Gln,<sup>8</sup> due to the difference in sidechain-backbone sterics.<sup>3</sup> Despite this thermodynamic difference between D-Gln and D-Asn however, we expect that these amino acid substitutions would exert similar impacts on the folding/unfolding kinetics of the cage because of their structural similarity. Indeed, as shown (Figure 3), substitution of Gly10 with D-Asn does not, within our experimental uncertainty, change the folding rate of TC5b; instead it decreases the unfolding rate. Previous studies have shown that the  $\alpha$ -helix of TC5b in isolation is unfolded.<sup>37</sup> Thus, folding of this  $\alpha$ -helix in the absence of other native stabilizing interactions would place the system at a higher free energy position, as concluded previously.<sup>9,17,24,30,48</sup> However, such a free energy 'penalty' is warranted and necessary, as it prepares and ensures the polypeptide chain to rapidly evolve toward the native state. Since

D-Gln and D-Asn are able to act, in the current case, as a C-terminal helix cap, one could have assumed that both would be able to lower the folding free energy barrier, making the protein fold faster. The fact that this is not the case suggests that these mutations involve interactions that stabilize not only the  $\alpha$ -helix, through capping, but also other parts of the cage structure. This is consistent with previous MD simulations of these mutants, which found that backbone fluctuations in the D-amino acid mutants are reduced, especially from residues 9 to 15.<sup>8</sup> However, these ‘nonlocal’ interactions are expected to be intimately coupled to the structural evolution of the cage, not the  $\alpha$ -helix. Thus, the native rotameric state of these D-amino acids in the corresponding mutants would be expected to be significantly populated only after the major folding free energy barrier.

### 5b-10GDA

As shown (Table 1), D-Ala is almost as effective as D-Gln in stabilizing the Trp-cage fold at position 10. One of the factors this has been attributed to is that D-Ala has less sidechain conformational entropy than both D-Gln and D-Asn, which highlights the well-known importance of entropy-enthalpy compensation in protein folding and design. In light of the above discussion and the fact that D-Ala does not have a sidechain rotamer preference, one might expect that this mutation would lead to an increase the folding rate, in addition to a decrease in the unfolding rate. As shown (Figure 4), indeed, while the unfolding rate of 5b-10GDA is significantly slower than TC5b, similar to that observed for 5b-10GDQ and 5b-10GDN, its folding rate also shows a marginal, but detectable, increase. These results thus confirm the above hypothesis that, unlike D-Gln and D-Asn, D-Ala helps to stabilize, although only to a small extent, the folding transition state or the nascent  $\alpha$ -helix. A simple calculation indicates that at 30 °C the D-Ala mutation stabilizes the folding transition state by approximately 0.33 kcal/mol, suggesting that D-Ala already becomes partially native in the transition state, again owing to the relatively small number of configurations the residue can sample. Thus, taken together, the results obtained from these D-amino acids demonstrate their utility in helping reveal certain fine details of the folding mechanism of proteins of interest.

## 4. Conclusions

Many previous studies have elegantly demonstrated the versatility of D-amino acids in protein design, in particular to increase the conformational stability of the targeted folds. However, what is lacking in our current understanding of the effectiveness of specific D-amino acids in this regard is their kinetic mechanism of action. Herein, we use the miniprotein Trp-cage as a model to examine the role of three D-amino acids, i.e., D-Gln, D-Asn, and D-Ala in altering the folding/unfolding kinetics of this protein, based on an earlier study that showed that these D-amino acids can significantly increase the stability of Trp-cage, when replacing the Gly residue in the loop region at position 10. Our results show that these mutations almost exclusively reduce the unfolding rate of the Trp-cage, although the D-Ala mutation also shows a marginal increase in the folding rate. In light of the ability of these D-amino acids to stabilize the  $\alpha$ -helix in this miniprotein via a C-capping effect, these results indicate that this  $\alpha$ -helix, while being populated in the folding transition state as suggested by a previous study, is nucleated at its N-terminus when passing through the folding free energy barrier.

## Acknowledgments

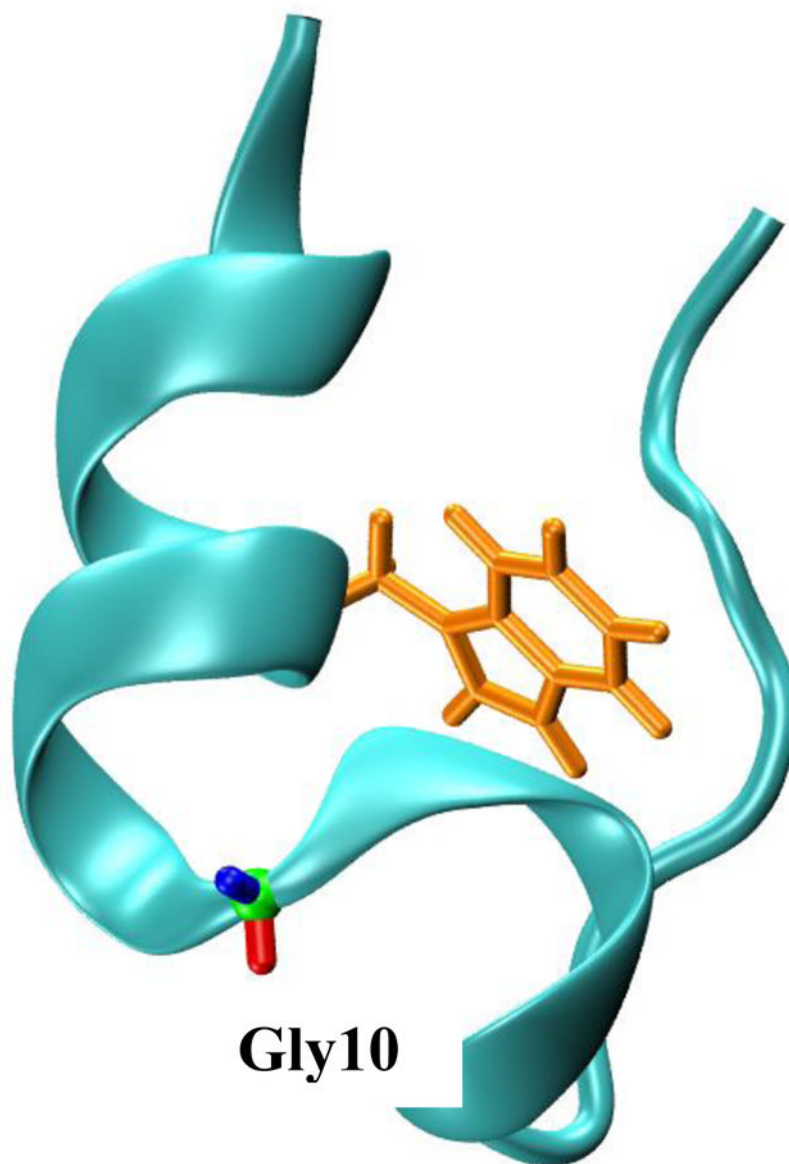
We thank the National Institutes of Health (GM-065978, RR01348, GM-008275, and GM-089949) for funding. R.M.C. is an NIH Ruth Kirschstein Predoctoral Fellow.

## References

1. Costantini S, Colonna G, Facchiano AM. *Biochem Biophys Res Commun.* 2006; 342:441. [PubMed: 16487481]
2. Bang D, Gribenko AV, Tereshko V, Kossiakov AA, Kent SB, Makhatazde GI. *Nature Chem Biol.* 2006; 2:139. [PubMed: 16446709]
3. Stites WE, Pranata J. *Proteins.* 1995; 22:132. [PubMed: 7567961]
4. Anil B, Song BB, Tang YF, Raleigh DP. *J Am Chem Soc.* 2004; 126:13194. [PubMed: 15479052]
5. Valiyaveetil FI, Sekedat M, MacKinnon R, Muir TW. *Proc Natl Acad Sci USA.* 2004; 101:17045. [PubMed: 15563591]
6. Williams DV, Barua B, Andersen NH. *Org Biomol Chem.* 2008; 6:4287. [PubMed: 19005584]
7. Struthers MD, Cheng RP, Imperiali B. *Science.* 1996; 271:342. [PubMed: 8553067]
8. Rodriguez-Granillo A, Annavarapu S, Zhang L, Koder RL, Nanda V. *J Am Chem Soc.* 2011; 133:18750. [PubMed: 21978298]
9. Culik RM, Serrano AL, Bunagan MR, Gai F. *Angew Chem Int Ed.* 2011; 50:10884.
10. Neidigh JW, Fesinmeyer RM, Andersen NH. *Nat Struct Biol.* 2002; 9:425. [PubMed: 11979279]
11. Qiu LL, Pabit SA, Roitberg AE, Hagen SJ. *J Am Chem Soc.* 2002; 124:12952. [PubMed: 12405814]
12. Snow CD, Zagrovic B, Pande VS. *J Am Chem Soc.* 2002; 124:14548. [PubMed: 12465960]
13. Simmerling C, Strockbine B, Roitberg AE. *J Am Chem Soc.* 2002; 124:11258. [PubMed: 12236726]
14. Zhou RH. *Proc Natl Acad Sci USA.* 2003; 100:13280. [PubMed: 14581616]
15. Chowdhury S, Lee MC, Xiong GM, Duan Y. *J Mol Biol.* 2003; 327:711. [PubMed: 12634063]
16. Pitera JW, Swope W. *Proc Natl Acad Sci USA.* 2003; 100:7587. [PubMed: 12808142]
17. Nikiforovich GV, Andersen NH, Fesinmeyer RM, Frieden C. *Proteins.* 2003; 52:292. [PubMed: 12833552]
18. Schug A, Herges T, Wenzel W. *Phys Rev Lett.* 2003; 91:158102. [PubMed: 14611501]
19. Carnevali P, Toth G, Toubassi G, Meshkat SN. *J Am Chem Soc.* 2003; 125:14244. [PubMed: 14624550]
20. Ota M, Ikeguchi M, Kidera A. *Proc Natl Acad Sci USA.* 2004; 101:17658. [PubMed: 15591340]
21. Steinbach PJ. *Proteins.* 2004; 57:665. [PubMed: 15390266]
22. Ahmed Z, Beta IA, Mikhonin AV, Asher SA. *J Am Chem Soc.* 2005; 127:10943. [PubMed: 16076200]
23. Neuweiler H, Doose S, Sauer M. *Proc Natl Acad Sci USA.* 2005; 102:16650. [PubMed: 16269542]
24. Ding F, Buldyrev SV, Dokholyan NV. *Biophys J.* 2005; 88:147. [PubMed: 15533926]
25. Linhananta A, Boer J, MacKay I. *J Chem Phys.* 2005; 122:114901. [PubMed: 15836251]
26. Irbach A, Mohanty S. *Biophys J.* 2005; 88:1560. [PubMed: 15613623]
27. Alonso JL, Echenique P. *Biophys Chem.* 2005; 115:159. [PubMed: 15752599]
28. Chen J, Im W, Brooks CL. *J Am Chem Soc.* 2006; 128:3728. [PubMed: 16536547]
29. Juraszek J, Bolhuis PG. *Proc Natl Acad Sci USA.* 2006; 103:15859. [PubMed: 17035504]
30. Yang L, Grubb MP, Gao YQ. *J Chem Phys.* 2007; 126:125102. [PubMed: 17411164]
31. Beck DAC, White GWN, Daggett V. *J Struct Biol.* 2007; 157:514. [PubMed: 17113307]
32. Zhan LX, Chen JZY, Liu WK. *Proteins.* 2007; 66:436. [PubMed: 17094111]
33. Piana S, Laio A. *J Phys Chem B.* 2007; 111:4553. [PubMed: 17419610]
34. Copps J, Murphy RF, Lovas S. *Biopolymers.* 2007; 88:427. [PubMed: 17326200]
35. Kentsis A, Gindin T, Mezei M, Osman R. *PLoS ONE.* 2007; 2:e446. [PubMed: 17505540]
36. Barua B, Lin JC, Williams VD, Kummeler P, Neidigh JW, Andersen NH. *Protein Eng, Des Sel.* 2008; 21:171. [PubMed: 18203802]
37. Stewart JM, Lin JC, Andersen NH. *Chem Commun.* 2008:4765.
38. Hu ZH, Tang YH, Wang HF, Zhang X, Lei M. *Arch Biochem Biophys.* 2008; 475:140. [PubMed: 18474213]

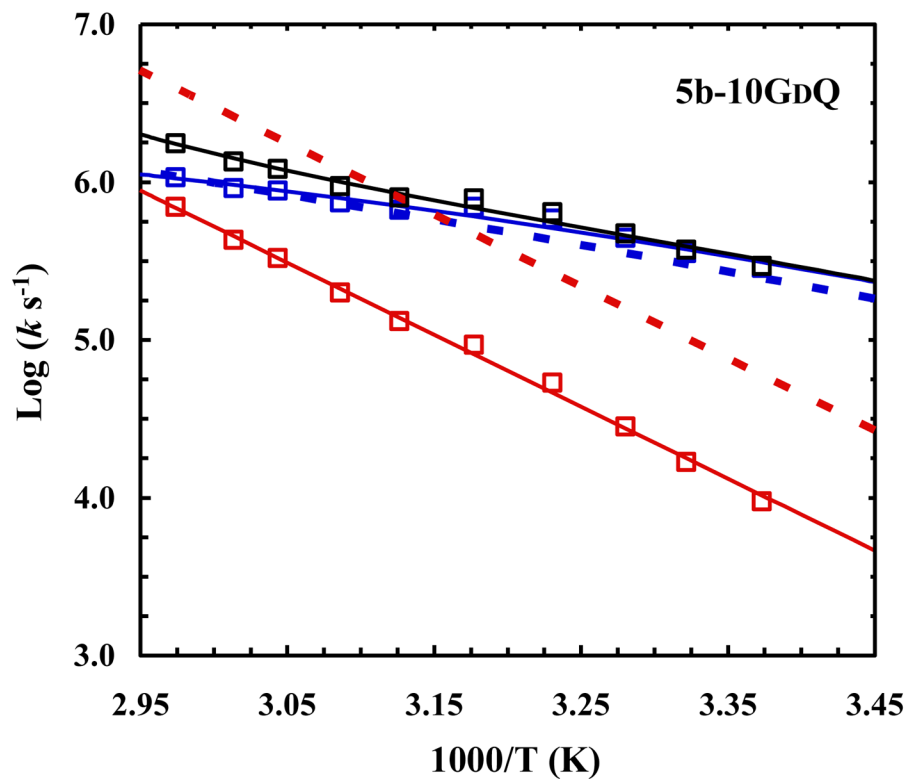
39. Wu S, Zhuravlev PI, Papoian GA. *Biophys J*. 2008; 95:5524. [PubMed: 18805918]
40. Hudaky P, Straner P, Farkas V, Varadi G, Toth G, Perczel A. *Biochemistry*. 2008; 47:1007. [PubMed: 18161949]
41. Xu WX, Mu YG. *Biophys Chem*. 2008; 137:116. [PubMed: 18775599]
42. Yao XQ, She ZS. *Biochem Biophys Res Commun*. 2008; 373:64. [PubMed: 18549806]
43. Kannan S, Zacharias M. *Proteins*. 2009; 76:448. [PubMed: 19173315]
44. Cerny J, Vondrasek J, Hobza P. *J Phys Chem B*. 2009; 113:5657. [PubMed: 19444987]
45. Chebaro Y, Dong X, Laghaei R, Derreumaux P, Mousseau N. *J Phys Chem B*. 2009; 113:267. [PubMed: 19067549]
46. Matthes D, de Groot BL. *Biophys J*. 2009; 97:599. [PubMed: 19619475]
47. Marinelli F, Pietrucci F, Laio A, Piana S. *PLoS Comput Biol*. 2009; 5:e1000452. [PubMed: 19662155]
48. Day R, Paschek D, Garcia AE. *Proteins*. 2010; 78:1889. [PubMed: 20408169]
49. Serrano AL, Waegle MM, Gai F. *Prot Sci*. 2012; 21:157.
50. Mukherjee S, Chowdhury P, Bunagan MR, Gai F. *J Phys Chem B*. 2008; 112:9146. [PubMed: 18610960]

- The relaxation kinetics of three D-amino acid mutants of Trp-cage were measured.
- These mutations stabilize the fold.
- These mutations primarily increase the folding rate.
- The results suggest that the folding of the  $\alpha$ -helix in this miniprotein is nucleated at the N-terminus.

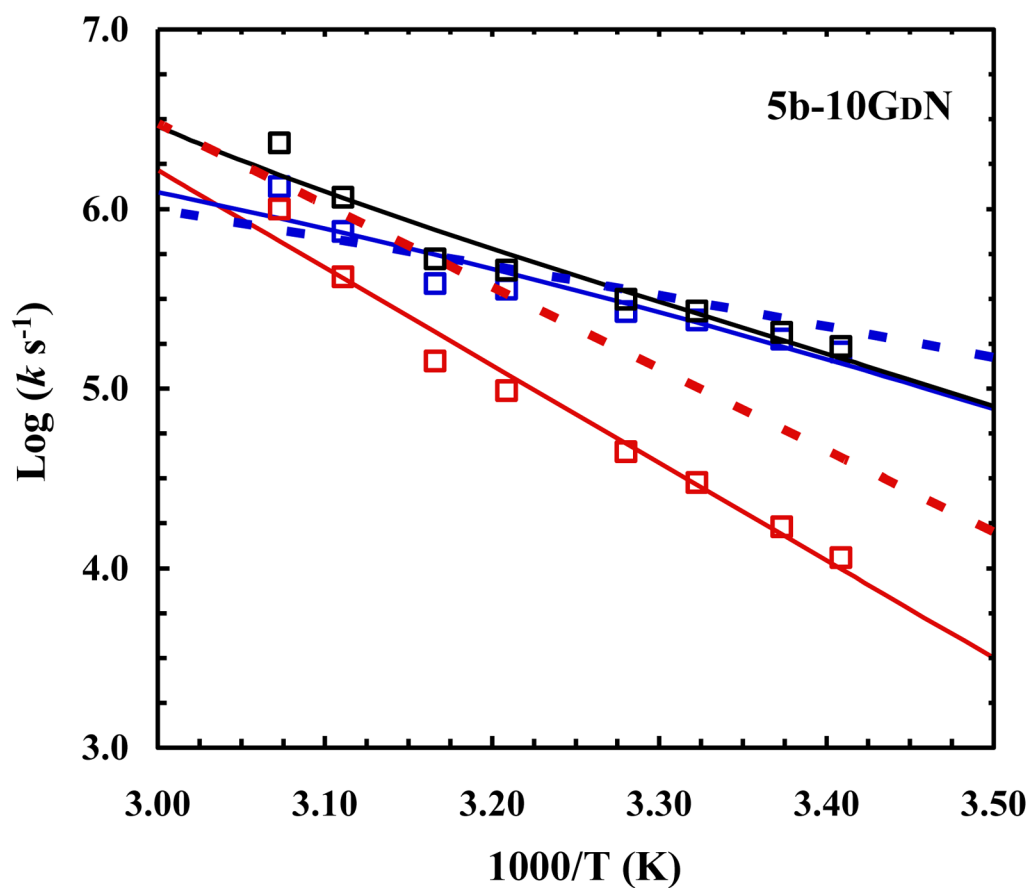


**Fig. 1.** The NMR structure of Trp-cage (PDB 1L2Y). The tryptophan sidechain and the glycine residue that was mutated to various D-amino acids in the current study are highlighted.

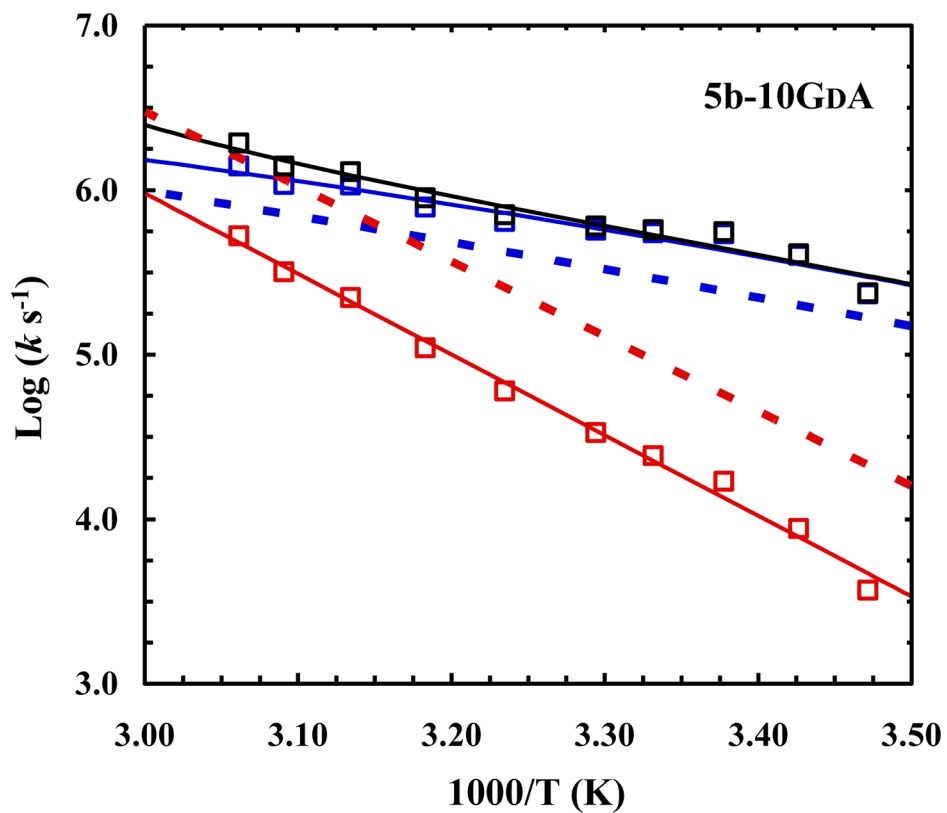




**Fig. 2.** Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GDQ. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type, adapted from ref 9.



**Fig. 3.** Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GdN. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type, adapted from ref 9.



**Fig. 4.** Arrhenius plots of the relaxation (black), folding (blue), and unfolding (red) rate constants of 5b-10GdA. The dotted lines correspond to the folding (blue) and unfolding (red) rate constants of the wild type, adapted from ref 9.

Unfolding thermodynamic parameters of TC5b and its D-amino acid mutants, adapted from ref 8. Also listed are the folding and unfolding times of these peptides at 30 °C.

**Table 1**

Peptide	$H_m$ (kcal mol <sup>-1</sup> )	$S_m$ (cal mol <sup>-1</sup> K <sup>-1</sup> )	$C_p$ (cal mol <sup>-1</sup> K <sup>-1</sup> )	$T_m$ (°C)	$\tau$ ( $\mu$ s)	$u$ ( $\mu$ s)
TC5b	12.9	40.4	72.1	45.9	3.0	7.6
5b-10GDA	17.1	50.4	53.6	66.6	1.7	33.4
5b-10GDQ	16.5	48.2	63.9	69.3	2.5	44.0
5b-10GDN	15.8	47.8	91.6	56.4	3.7	25.5