JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

ARTICLE

pubs.acs.org/JACS

Computational Design of Thermostabilizing D-Amino Acid Substitutions

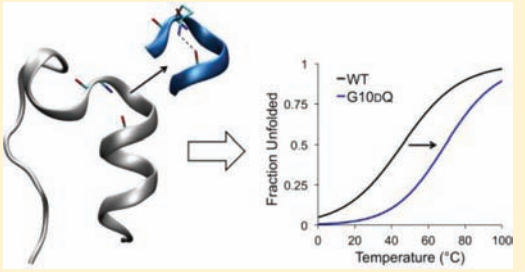
Agustina Rodriguez-Granillo,[†,§] Srinivas Annavarapu,[†,§] Lei Zhang,[‡] Ronald L. Koder,[*,‡] and Vikas Nanda[*,†]

[†]Department of Biochemistry, Robert Wood Johnson Medical School, University of Medicine and Dentistry of New Jersey (UMDNJ) and Center for Advanced Biotechnology and Medicine, Piscataway, New Jersey 08854, United States

[‡]Department of Physics, The City College of New York, New York, New York 10031, United States

Supporting Information

ABSTRACT: Judicious incorporation of D-amino acids in engineered proteins confers many advantages such as preventing degradation by endogenous proteases and promoting novel structures and functions not accessible to homochiral polypeptides. Glycine to D-alanine substitutions at the carboxy termini can stabilize α-helices by reducing conformational entropy. Beyond alanine, we propose additional side chain effects on the degree of stabilization conferred by D-amino acid substitutions. A detailed, molecular understanding of backbone and side chain interactions is important for developing rational, broadly applicable strategies in using D-amino acids to increase protein thermostability. Insight from structural bioinformatics combined with computational protein design can successfully guide the selection of stabilizing D-amino acid mutations. Substituting a key glycine in the Trp-cage miniprotein with D-Gln dramatically stabilizes the fold without altering the protein backbone. Stabilities of individual substitutions can be understood in terms of the balance of intramolecular forces both at the α-helix C-terminus and throughout the protein.



1. INTRODUCTION

We hypothesize that polar D-amino acids can stabilize the carboxy termini of α-helices through side chain to backbone hydrogen bonds. Computational methods are developed to predict the energetic consequences of such substitutions. A major drawback of peptide therapeutics that limits their biomedical application is low thermodynamic stability, which results in weak activity and rapid degradation by endogenous proteases. Globular proteins typically maintain stable native structures through extensive tertiary contacts within hydrophobic cores. Peptides have minimal tertiary contacts, resulting in highly dynamic ensembles that significantly populate partially or fully unfolded states. Strategies to stabilize peptides focus either on enhancing the scant tertiary structure or favoring secondary structure. Various chemical approaches have been developed to modulate protein stability, such as engineering disulfides[1] or metal binding sites[2] or incorporating fluorine-derivatized amino acids.[3] Constrained oligoproline derivatives have been used as templates for helix initiation, improving both folding kinetics and stability.[4] Various bio-orthogonal chemical approaches have been used to improve peptide stability by linking side chains[5] or replacing a backbone hydrogen bond with a carbon−carbon bond.[6]

Our understanding of the role of stereochemistry in protein structure and folding comes from seminal contributions by Ramachandran in the early 1960s and 1970s.[7,8] Models of allowed dipeptide conformations based on hard-sphere sterics of atomic interactions have provided a physical basis for understanding the complex structures of natural proteins. High-resolution structures of model peptides have provided some glimpse into the increased structural diversity accessible with non-natural amino acids.[9] Work on somatostatin demonstrated that incorporating D-amino acids into small peptides could dramatically enhance their biological activity.[10] Introducing one or at most a few D-amino acids has been used to stabilize β-turns,[11,12] novel folds,[13] and larger proteins.[14] We build on previous work and explore the use of computational design with D-amino acids to stereochemically favor native secondary structure as a strategy for improving peptide thermostability.

Typically, incorporation of D-amino acids into existing proteins and peptides will adversely impact structure and stability. Substituting an L-Ala for a D-Ala in the middle of an α-helix reduces its stability by ~1 kcal/mol due to steric clashes between the side chain methyl and the backbone carbonyls of proximal residues in the polypeptide chain.[15] In specific contexts, a Gly to D-Ala substitution at the carboxy terminus can improve its stability.[16−18] Glycine plays a unique structural role at the carboxy terminus of an α-helix, forming a network of hydrogen bonds that cap backbone carbonyls of adjacent residues while directing the chain to the next structural element. These structural "punctuation marks" have been frequently observed in natural proteins.[19] Glycine lacks a side chain, allowing it to easily assume an α_L conformation ($\phi \approx +65°$, $\psi \approx +42°$), which is energetically unfavorable for the other 19 L-amino acids.[8] By replacing Gly with D-Ala, it is possible to stably form the α_L capping structure without the entropic cost of fixing the conformation of glycine. This strategy has been applied to several proteins with mixed results. In a series of Gly to D-Ala

Received: June 17, 2011
Published: October 06, 2011

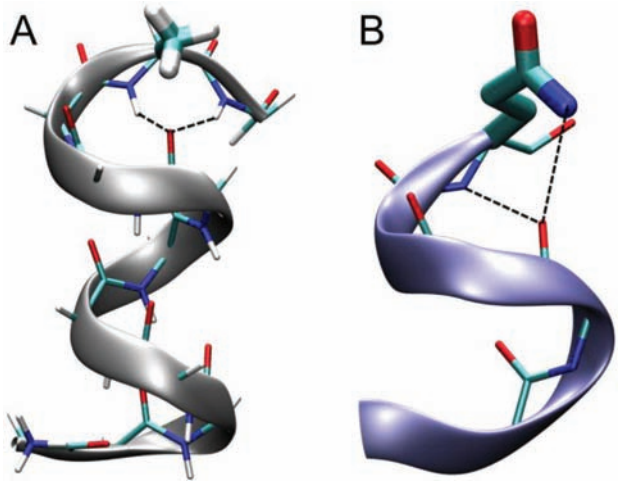


Figure 1. C-capping of an α-helix: (A) C-capping motif from a poly-Ala, $(\text{D})_9(\text{L})_2$, peptide obtained from Monte Carlo simulations of poly-LD-Ala peptides in a previous study (ref 21), (B) example of a natural L-Gln C′ capping motif in an α_L turn (residues 297−301 of PDB entry 2J6L).

substitutions in small proteins, it was found that stability was improved by anywhere from 0.6 to 2 kcal/mol.[16] Equivalent substitutions in ubiqutin showed no increase in stability.[17] In a monomeric α-helix, it was found that D-Ala had no effect, but D-Arg stabilized the helix by ∼1 kcal/mol.[20] These discrepancies indicate that the role of stereochemistry in modulating helix stability is still an open problem. A systematic and rigorous exploration of stereochemical rules will be critical in developing rational and computational methods for engineering enhanced peptide stability.

In this study, we investigate the potential for D-amino acids to stabilize the ends of α-helices through both reducing backbone flexibility and promoting novel side chain to backbone hydrogen bonds that are not accommodated by equivalent L-stereoisomers. The potential for such interactions was noted in Monte Carlo simulations of poly-LD-alanine peptides, sampling both backbone flexibility and side chain chirality.[21] In simulations of an 11-residue polyalanine peptide, (L-Ala)$_9$-D-Ala-L-Ala and (D-Ala)$_9$-L-Ala-D-Ala were frequently found. These corresponded to right- and left-handed α-helices, respectively. Interestingly, the change in chirality at the end of the sequence was due to the incorporation of C-capping motifs (Figure 1A). Note that the methyl group of alanine is proximal to exposed carbonyls of the helix terminus. We hypothesized other amino acids besides alanine might provide additional stability through additional side chain to backbone hydrogen bonds. An example of such an interaction was found in the high-resolution structure of the enzyme alanine racemase,[22] where the imidazole Nδ^1 nitrogen of an L-histidine at the C′ position made a hydrogen bond to the C1 carbonyl in a preceding α_L turn.

Previous work on helix capping with natural amino acids often utilized the Protein Data Bank (PDB) as a resource for identifying common interactions at the ends of helices.[23] Finding similar motifs involving D-amino acids has proved challenging, because only a minute fraction of the structures deposited in the PDB contain D-amino acids.[24] Of these cases, many are short cyclic peptides that do not contain helical elements. To address this issue, we implemented a "mirror-image template" search of the PDB.[25] The examination of α_L-containing structures in the PDB has traditionally been useful in exploring the structural role of side chain stereochemistry.[26] It was reasoned that amino acids making side chain hydrogen bonds to the backbone of left-handed turns would be the mirror-image equivalent of D-amino acids interacting with right-handed turns. By normalizing the frequencies of observed amino acids flanking left-handed turns to random expectation, several over-represented and under-represented residues at both the N- and C-termini of helices were identified. Glutamine was a particularly promising candidate, in terms of both its propensity (3.7-fold over-represented) and the presence of C′ side chain to helix backbone hydrogen bonds in multiple high-resolution protein structures (an example is shown in Figure 1B), suggesting that D-Gln should be stabilizing at the C-terminus of an α_R-helix. Conversely, asparagine had a propensity of 0.6, suggesting it was unfavorable at this position. Less than a hundred α_L turns were identified in the previous study, resulting in large statistical uncertainties. It is crucial to test the validity of these motifs in an experimental system as is being done here. Furthermore, the C-termini of α-helices are often at the protein surface, making extensive interactions with solvent.[27] Thus, the expected contribution of such hydrogen bonds is debatable, given that formation of the interaction would come at the expense of desolvation.[17,28]

To test the prediction that certain polar D-amino acid motifs at the C′ position can stabilize an α_R-helix, we describe the design and characterization of D-amino acid substitutions to the Trp-cage miniprotein TC5b (Trp-cage version 5b).[29] The Trp-cage was chosen as a model system for several reasons. It has been the subject of intense computational and biophysical study, making it a well-understood model system for protein folding.[30−32] At 20 amino acids long, the Trp-cage is readily accessible to modification with non-natural amino acids by solid-phase peptide synthesis. Its small size also makes it amenable to molecular simulation, and a significant number of molecular dynamics studies have been performed on it.[31,33] Unlike many peptides of this size, folding is cooperative and can be described as a two-state process,[34,35] simplifying thermodynamic interpretations of amino acid substitutions. Previous work on members of the Trp-cage family demonstrated that D-Ala substitutions could dramatically raise protein thermal stability.[18] This study will present the design, thermodynamic stability, solution structure, and molecular dynamics characterizations of several Trp-cage D-amino acid variants with the goal of developing computational tools for the design of thermostabilizing substitutions.

2. MATERIALS AND METHODS

2.1. Sample Preparation. Peptides were synthesized using solid-phase 9-fluorenylmethoxycarbonyl (FMOC) chemistry at the Tufts University Core Facility (http://www.tucf.org). N- and C-termini were acetylated and amidated, respectively. Peptides were purified to 90% purity by reversed-phase high-performance liquid chromatography (HPLC), and products were verified by mass spectrometry. Sequences in this study are as follows:

TC5b WT	NLYIQWLKD	G GPSSGRPPPS
TC5b G10A	NLYIQWLKD	A GPSSGRPPPS
TC5b G10Q	NLYIQWLKD	Q GPSSGRPPPS
TC5b G10DA	NLYIQWLKD	DA GPSSGRPPPS
TC5b G10DQ	NLYIQWLKD	DQ GPSSGRPPPS
TC5b G10DN	NLYIQWLKD	DN GPSSGRPPPS

2.2. Thermal and Chemical Denaturation. *2.2.1. Circular Dichroism (CD).* Peptides were dissolved in deionized water, and the concentration was measured at 280 nm using an extinction coefficient of 6990 M^{-1} cm^{-1}. Peptide solutions at 50 μM concentration were

prepared by diluting peptide stocks in 10 mM phosphate buffer, pH 7.0. CD experiments were performed on an Aviv model 400 spectrometer (Aviv Biologicals, Lakewood, NJ). Ellipticity spectra were collected from 190 to 260 nm in 0.5 nm steps at 25 °C. A 1.0 nm bandwidth, 1 s averaging time, and 1 s equilibration time were used. Molar ellipticity was calculated by multiplying raw values by the peptide concentration (50 μM), number of residues (20), and cell path length (1 mm).

For thermal denaturation, peptide solutions were diluted in 10 mM phosphate buffer (pH 7.0) to obtain a concentration of 50 μM. Ellipticity was measured at a wavelength of 222 nm with a bandwidth of 1.0 nm, temperature dead band of 0.15 °C, temperature equilibration time of 1.0 min, and averaging time of 2.0 s. Temperatures ranged from 0 to 80 °C in 2 °C steps. Temperature melts were conducted for six concentrations of guanidine HCl (GnHCl) for each peptide: 0.0, 1.0, 2.0, 3.0, 4.0, and 5.0 M.

Temperature melt plots obtained for multiple GnHCl concentrations were fit to folded and unfolded baselines. Heat capacity was assumed to remain constant with changes in GnHCl concentration for each Trp-cage variant. The fraction folded, $f(T)$, was calculated:

$$f(T) = [\theta(T) - \theta_U(T)]/[\theta_F(T) - \theta_U(T)] \quad (1)$$

where θ, θ_F, and θ_U are the observed ellipticity, ellipticity of the folded form, and ellipticity of the unfolded form, respectively.

Free energy (ΔG), enthalpy (ΔH), and heat capacity (ΔC_P) estimates were fit to the experimental data using the Gibbs−Helmholtz equation and the following relationships:

$$\Delta G = \Delta H(1 - T/T_M) - \Delta C_P((T_M - T) + T \ln(T/T_M)) \quad (2a)$$

$$K = \exp(-\Delta G/(RT)) \quad (2b)$$

$$\alpha = K/(1 + K) \quad (2c)$$

$$\theta_t = \alpha(\theta_F - \theta_U) + \theta_U \quad (2d)$$

T_M is the melting temperature, and α is the model fraction folded.

2.2.2. Differential Scanning Calorimetry (DSC). DSC experiments were carried out on an N DSC II (Applied Thermodynamics) calorimeter. The instrument was run using deionized water in sample and reference cells from 273 to 373 K for at least 10 scans to ensure stabilization of the readings. Trp-cage peptides were dialyzed as described. Solutions of Trp-cage variants were prepared in 10 mM phosphate buffer (pH 7.0) at a concentration of 2 mg/mL. Aliquots (800 μL) of sample and buffer solutions were loaded into their respective cells while the instrument temperature was between 293 and 313 K. The pressure was then slowly raised to 2 atm. The temperature was raised from 273 to 373 K at a rate of 1.5 K/min. An equilibration time of 10 min was allowed between scans.

2.3. Nuclear Magnetic Resonance (NMR) Spectroscopy and Structure Calculations. Spectra were collected on a Varian Inova 600 MHz spectrometer at 4 °C. Samples were 1.5−2 mM protein in 10 mM potassium phosphate buffer, pH 7.0.

The majority of experiments were conducted using an unlabeled Trp-cage 10DQ sample with 10% D_2O for frequency lock. $^1H-^1H$ NOESY (nuclear Overhauser effect spectroscopy)[36] and $^1H-^1H$ TOCSY (total correlation spectroscopy)[37] spectra were collected using WATERGATE solvent suppression and assigned by standard methods. A DIPSI-2[38] spin lock was employed for TOCSY, and mixing times of 60, 150, and 200 ms were used in the NOESY spectra. Spectral widths were 7000 Hz for both directions of ^1H.

To resolve ambiguous assignments for some of the D-glutamine atoms, a sample was created with the 10th position residue uniformly labeled by ^{15}N and ^{13}C. Due to the prohibitive cost of labeled D-amino acids, a peptide with inverted chirality was synthesized: this peptide was synthesized with all D-amino acids except at position 10, where uniformly ^{15}N,^{13}C-labeled L-glutamine was used. A CD spectrum of the inverso-TC5b G10DQ was consistent with a folded, mirror-image structure (Figure S1, Supporting Information). Sensitivity-enhanced ^{13}C and ^{15}N HSQC (heteronuclear single-quantum correlation)[39] spectra of this sample were collected in 100% D_2O for ^{13}C and 10% D_2O for ^{15}N. Spectral widths were 6000 Hz for ^1H, 24130 Hz for ^{13}C, and 2000 Hz for ^{15}N.

$^1H-^1H$ NOESY spectra with 150 ms mixing times were collected in 100% D_2O using double ^{13}C isotope filters, resulting in cross-peaks and distance constraints between ^{13}C-bound protons only. Similar spectra collected using a ^{15}N isotope filter for the first proton and a ^{13}C isotope filter for the second in 10% D_2O resulted in cross-peaks and distance constraints between ^{15}N-bound protons and ^{13}C-bound protons only. Spectral widths were 6000 Hz for ^1H.

Natural-abundance sensitivity-enhanced constant-time ^{13}C HSQC was performed to confirm the NOESY and TOCSY-based assignments. Spectral widths were 7000 Hz for ^1H and 24130 Hz for ^{13}C.

Nuclear Overhauser effect (NOE) intensities were obtained from NOESY spectra. The constraints were converted to distance and used in a simulated annealing protocol using the ARIA[40] and CNS[41] software packages. Assignments and the constraints list are included in the Supporting Information.

A total of 50 structures were generated, out of which 20 structures were kept. The total energy was used for the sort criterion. A total of 425 restraints were used for structure calculation, with zero violations. Out of the used restraints, there were 282 unambiguous restraints and 143 ambiguous restraints. The lowest energy structure was deposited in the PDB (http://www.rcsb.org/pdb under accession ID 2LDJ).

2.4. Computational Methods. *2.4.1. Design of Trp-Cage Variants with protCAD.* Computational models of Trp-cage D-amino acid variants were created using protCAD (protein Computer Aided Design)[42] software. This protein design package is versatile and can be used for various design functions, such as energy calculations, mutations, changes to backbone and side chain dihedrals, and modeling simulations. Variants of the Trp-cage, where D-amino acids replaced left-handed glycines, were modeled using the structure of Trp-cage variant TC5b as a starting point (PDB entry 1L2Y).[29] All rotamers of D-Gln were modeled at position 10 of the Trp-cage using protCAD software, and the total van der Waals (vdW) energy was computed using nonbonding united-atom parameters from Amber 95.[43] Hydrogen bonds were calculated using a side chain Nε to backbone carbonyl oxygen distance cutoff of 3.5 Å (g−, gauche −; t, trans; g+, gauche +).

2.4.2. Molecular Dynamics (MD) Simulations. The atomic coordinates for the wild-type (WT) Trp-cage were obtained from the PDB, entry 1L2Y.[29] The following mutations were generated in the WT structure in silico with the *leap* module of Amber 10:[44] G10A, G10Q, G10DA, G10DN and G10DQ.

MD simulations were performed for the different protein variants using Amber 10.[44] The initial structures were immersed in a pre-equilibrated truncated octahedral cell of TIP3P explicit water molecules,[45] and counterions were added to neutralize the systems. Protein atoms were described with the parm99SB[43,46] force field parametrization. The protonation state of the titratable residues corresponded to the stable form at pH 7.0. Water molecules extended at least 12 Å from the surface of the proteins. Simulations were performed in the NPT ensemble (constant pressure of 1 atm and temperature of 300 K were maintained using the Berendsen coupling scheme[47]) employing periodic boundary conditions. The SHAKE algorithm was employed to keep bonds involving hydrogen atoms at their equilibrium length.[48]

The systems were optimized, heated to 300 K, and equilibrated for 200 ps. The structures were then simulated until the root-mean-square deviation (rmsd) as a function of time was stable for at least 50 ns (total simulation time was more than 100 ns for each). The converged last

50 ns were used for data analysis. rmsd's, root-mean-square fluctuations (rmsf's), average structures, and hydrogen bond frequencies were calculated for each of the systems using the *ptraj* module of Amber 10.[44] Cluster analysis was performed with the program *kclust* from MMTSB,[49] using rmsd as the similarity measure, with a cutoff of 2 Å.

2.4.3. Generating Potentials of Mean Force. To describe the thermodynamics of the side chain capping interaction in G10DN and G10DQ, we computed the Helmholtz free energy, $A(r)$, along the distance r between the side chain of residue 10 (Nδ in G10DN and Nε in G10DQ) and the backbone carbonyl of Leu7, Lys8, and Asp9, yielding a free energy profile or a potential of mean force (PMF):[50]

$$A(r) = -k_B T \ln \zeta(r) + \text{constant} \quad (3)$$

where k_B is the Boltzmann constant, T is the absolute temperature, and $\zeta(r)$ is the probability of the coordinate taking on a particular value of r. The constant was chosen so that the most probable distribution corresponded to a free energy of zero.

2.4.4. Estimating Side Chain Entropy. To estimate the change in entropy of side chain 10 between the folded and unfolded states in G10DN and G10DQ, two additional simulations that represented the unfolded states were performed. These consisted of the tripeptides GNG and GQG in explicit solvent, using the same parameters as above. Both peptides were simulated for ~160 ns, and the last 50 ns were used for data analysis.

The conformational entropy, S, of the residue 10 side chain in G10DN and G10DQ and the central amino acid in GNG and GQG was estimated using the classical definition:

$$S = -k_B \sum_i p_i \ln p_i \quad (4)$$

where the sum is taken over all conformational states (side chain rotamers) of the system and p_i is the probability of being at a particular state i. The change in conformational entropy was calculated for different side chain rotamer definitions, including two binning offsets (from 0° to 360° and from 30° to 390°) and three bin sizes (30°, 60°, and 90°) for χ_1 and χ_2 in Asn and χ_1, χ_2, and χ_3 in Gln (Table S2, Supporting Information). The averages of these values are reported in Table 3.

The vibrational entropy of the side chain in G10DN and G10DQ and the central amino acid in GNG and GQG was estimated by calculating the quasi-harmonic entropy[51] using the *ptraj* module of Amber 10. In this approach the atomic fluctuation matrix is calculated as the mass-weighted covariance matrix obtained from the snapshots of the MD simulation. The structures obtained from the simulations were spatially superimposed to a common reference structure to exclude all translational and rotational motions exhibited by the molecule.[52] The entropy was estimated every 250 ps from the last 50 ns of the MD simulations. The calculation was performed for the side chain heavy atoms of residue 10.

3. RESULTS AND DISCUSSION

3.1. Stereochemistry and Structure. The TC5b sequence[29] was selected as the starting point for all molecular engineering efforts. On the basis of the solution NMR structure, Gly10 is in the α_L conformation at the end of a nine-residue α-helix, making it an ideal candidate for introducing D-amino acid substitutions. As a negative control, two L-amino acid variants were tested: G10A and G10Q. Although it was expected that such substitutions would be inherently destabilizing, we were motivated to test them experimentally for two reasons: to explore the importance of correct α_L favoring stereochemistry at position 10 for protein stability, and previous structural bioinformatics work uncovered several examples of glutamine in an α_L position following a right-handed helix making side chain−backbone hydrogen bonds which could presumably provide stability.

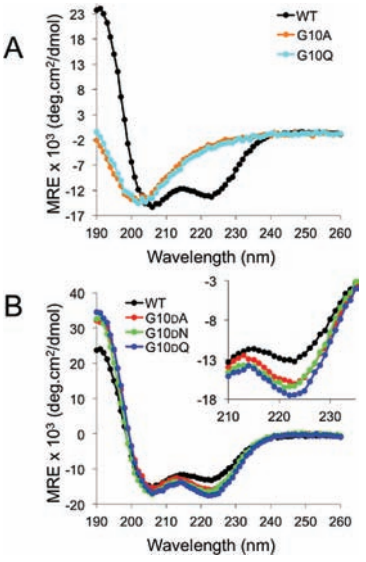


Figure 2. Far-UV CD spectra of the Trp-cage variants at 25 °C: (A) WT, G10A, and G10Q, (B) WT, G10DA, G10DN, and G10DQ. MRE = mean residue ellipticity.

Both L-variants showed no helical secondary structure as measured by CD at 25 °C, instead adopting a fully random-coil state (Figure 2A). No thermal unfolding transition was observed by CD. In contrast, WT TC5b was clearly α-helical, with characteristic minima in the CD spectrum at 208 and 222 nm. A residue capable of easily accommodating an α_L conformation is clearly important for folding, and while L-Gln can contribute an α-helix-capping hydrogen bond when introduced in TC5b, it is not sufficient to retain structure.

MD simulations on L-amino acid-substituted variants were consistent with experimental observations: the structure significantly deviated from the WT TC5b throughout the simulations (Figure 3A), and backbone fluctuations significantly increased in both L-variants with respect to the WT (Figure 3C). Although the average structure of G10A was similar to that of the WT in the production run (Figure 3E), averaging can mask dynamics throughout the entire simulation. Snapshots during this run significantly deviated from a WT conformation (Figure S2, Supporting Information). In the case of G10Q, the average structure was also significantly different from that of the WT in the production run (Figure 3E), and snapshots from the entire simulation show that key hydrophobic core interactions between the α-helix and the polyproline stretch were compromised (Figure S3, Supporting Information), resulting in partial unfolding.

Given the lack of structure of G10A and G10Q substitutions, we proceeded to focus on D-amino acid substitutions at this position. G10DA was predicted to stabilize TC5b by stereochemically favoring the α_L conformation, reducing the entropic cost of folding. On the basis of previous analysis of the PDB, D-Gln is hypothesized to be the most favorable helix-capping residue.[25]

TC5b G10DA, G10DN, and G10DQ variants were synthesized and examined for structure and stability. CD spectra of all D-variants were consistent with significant α-helical structure and had stronger negative ellipticity bands at 208 and 222 nm relative to those of WT TC5b (Figure 2B). Tryptophan fluorescence emission spectra of WT and modified D-amino acid peptides were

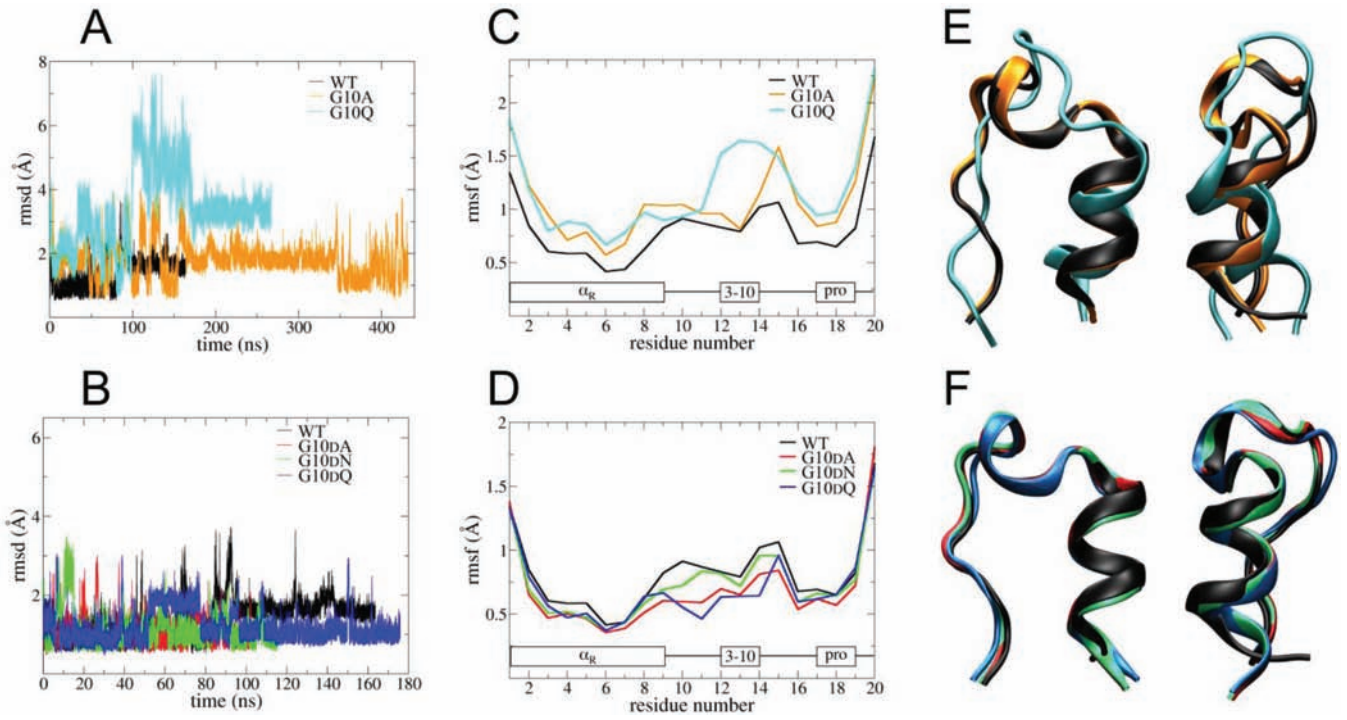


Figure 3. (A, B) rmsd's (Å, with respect to the first structure) of the backbone heavy atoms as a function of the simulation time (ns) for the WT and L-variants (A) and the WT and D-variants (B). The last 50 ns from each simulation were selected for data analysis (production run). (C, D) Average fluctuations (rmsf's, Å) of backbone heavy atoms as a function of the residue for the production run for the WT and L-variants (C) and the WT and D-variants (D). The secondary structure elements are indicated (α_R, right-handed α-helix; 3−10, right-handed 3_{10} helical turn; pro, polyproline stretch). (E, F) Superimposition of the average structures (front and side views) for the WT and L-variants (E) and the WT and D-variants (F). Key: WT (black), G10A (orange), G10Q (cyan), G10DA (red), G10DN (green), and G10DQ (blue).

nearly identical and consistent with a partially exposed tryptophan at position 6 (Figure S4, Supporting Information). MD simulations show that the rmsd's of the D-variants were comparable or even smaller than that of the WT (Figure 3B) and the average structures are closely aligned with TC5b (Figure 3F). The increase in helical content evidenced by CD agreed with the increase in helical content at the N-terminus of the α-helix in all D-variants, as shown in the average structures from the MD simulations (Figure 3F). This structural change was not observed in the L-variants.

3.2. Contributions of D-Amino Acids to Stability. Thermodynamic properties of the Trp-cage were determined by thermal denaturation, following loss of ellipticity at 222 nm as a function of increasing temperature. However, both G10DA and G10DQ were partially folded at temperatures greater than 85 °C, making it difficult to establish an unfolded-state baseline. To address this, we measured thermal denaturation curves at a series of GnHCl concentrations from 0 to 5 M, fitting the entire data set globally to the Gibbs−Helmholtz equation to extrapolate estimates of free energy of folding, enthalpy, heat capacity, and melting temperature (T_M) (Figure 4; Figure S5, Supporting Information).[53] On the basis of this analysis, the stability of TC5b was consistent with previous measurements of this peptide by other groups.[54] All D-variants were more stable than the WT. On the basis of the free energy of unfolding at 25 °C (Table 1), both G10DA and G10DQ were stabilized by ∼0.9 kcal/mol with respect to the WT, and G10DN by ∼0.6 kcal/mol. On the basis of T_M, the most stable variant was G10DQ, with a ΔT_m of ∼23 °C relative to the WT, followed by G10DA and G10DN. DSC experiments were carried out on the same variants, although the thermograms could not be fit to obtain thermodynamic

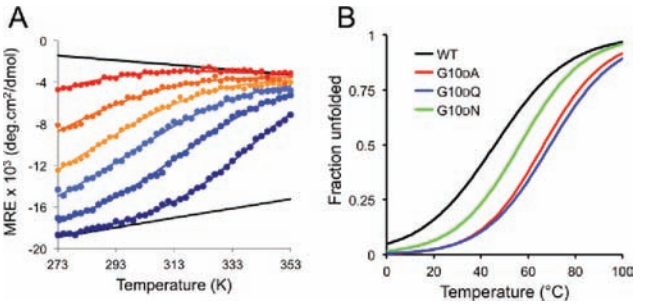


Figure 4. (A) Thermally induced unfolding of G10DQ Trp-cage at different GnHCl concentrations (from dark blue to red: 0.0, 1.0, 2.0, 3.0, 4.0, and 5.0 M) monitored by far-UV CD at 222 nm. (B) The melts for G10DQ (other variants in Figure S5, Supporting Information) were globally fit to obtain the fraction unfolded as a function of temperature (WT, black; G10DA, red; G10DN, green; G10DQ, blue). MRE = mean residue ellipticity.

Table 1. Thermodynamic Parameters of Trp-Cage Variants Obtained from DSC and CD Thermal Unfolding Experiments

	DSC	CD	
	T_M (°C)	T_M (°C)	ΔG_U(25 °C) (kcal/mol)
WT	36	46 ± 3	0.8 ± 0.2
G10DA	68	66.6 ± 0.2	1.7 ± 0.1
G10DN	63	56 ± 5	1.4 ± 0.4
G10DQ	69	69.3 ± 0.5	1.7 ± 0.2

dx.doi.org/10.1021/ja205609c |J. Am. Chem. Soc. 2011, 133, 18750–18759

Table 2. Time Fraction (%) of Backbone to Backbone Hydrogen Bond (HB) Capping Interactions Donated by the Backbone Amide (N) of Residue 10 to Positions 6−8 and the Backbone Carbonyl (O) of Residue 10 Accepted from Amides of Residues 13 and 14[a]

		WT	G10DA	G10DN	G10DQ	G10A	G10Q
X10.N	W6.O	36	24	30	24	0	99
	L7.O	63	73	67	78	40	0
	K8.O	9	7	7	4	42	0
X10.O	S13.N	66	86	82	89	74	0
	S14.N	12	6	18	3	25	0

[a] The HB distance and angle cutoffs were 3.5 Å and 120°, respectively.

parameters due to the broad, shallow transitions (Figure S6, Supporting Information). In the absence of distinct folded and unfolded baselines, T_M values were calculated using first-derivative plots and confirm the same trend as the one found by CD (Table 1). The difference between the two biophysical methods is notable for the lower stability WT and G10DN species. This is attributed to the broad unfolding transitions due to lower ΔH for these two species (Table S1, Supporting Information), resulting in a nonlinear dependence of the unfolded state heat capacity on temperature.[34,55]

MD simulations also indicated a similar stability trend. As opposed to the L-variants, all D-variants were less flexible than the WT (Figure 3D). The greatest changes were observed for residues 9−15 (which includes the 3_{10} helical turn). A decrease in backbone flexibility (rmsf's) may indicate increased stability as the reduced likelihood of random thermal motions away from the most compact, folded conformation. On the basis of rmsf trends, both G10DA and G10DQ are predicted to be more stable than G10DN, and G10DN is predicted to be more stable than the WT, as observed experimentally. Conversely, an increase of backbone flexibility, such as the one observed for G10A and G10Q, may increase the chances of random thermal motions away from the folded state, unfolding the overall structure.

MD may be used to interrogate the structure and dynamics of TC5b and variants for the molecular features of capping stability. In the WT Trp-cage, Gly10 participates in a network of backbone to backbone hydrogen bond capping interactions with both the α-helix and the right-handed 3_{10} helical turn (Table 2). These interactions include the backbone amide of Gly10 with the backbone carbonyls of Trp6, Leu7, and Lys8 in the α-helix, and the backbone carbonyl of Gly10 with the backbone amides of Ser13 and Ser14 in the 3_{10} helical turn. Because of the nature of the interactions and the proximity of the residues, more than one interaction can be present at the same time. The dominant, or the most frequent, interactions involved Leu7 and Ser13 (Table 2). In all D-variants, the frequency of these dominant hydrogen bonds increased, and the trend correlated with experimentally observed stability (G10DQ > G10DA > G10DN > WT). In the case of the Leu7 interaction, the increase ranges from ∼6% in G10DN to ∼24% in G10DQ. In the case of the Ser13 interaction, the increase is more dramatic and ranges from ∼24% in G10DN to ∼35% in G10DQ. We could speculate that the presence of a D-amino acid at position 10 drives the structure to a single state in which these dominant hydrogen bonds are present. The significant increase in the frequency of the hydrogen bond between position 10 and Ser13 may contribute to the increased rigidity of the entire helical turn in all D-variants.

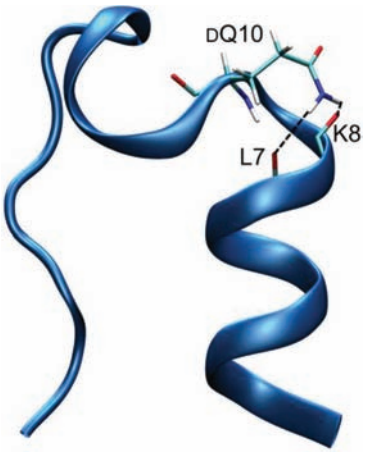


Figure 5. Lowest energy rotamer of D-Gln modeled onto the TC5b structure using protCAD.[29]

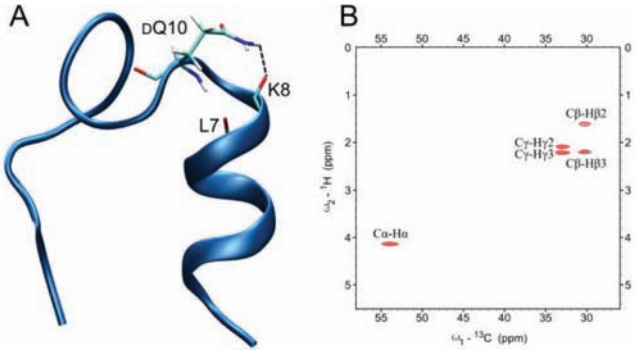


Figure 6. (A) One low-energy NMR structure of G10DQ showing a Gln side chain to backbone capping interaction with Lys8. (B) HSQC spectrum of DQ10 in G10DQ.

3.3. Structure of TC5b G10DQ. Using protCAD, D-Gln was modeled into the NMR structure of TC5b at position 10 (Figure 5). Over half of the 27 discrete rotamers sampled were sterically allowed, of which 4 formed a network of hydrogen bonds between D-Gln Nε and helix backbone carbonyls. One rotamer shown in Figure 5 participated in hydrogen bonds with carbonyls of Leu7 and Lys8, capping both C1 and C2 positions.

To solve the solution structure of the D-Gln variant, it was necessary to characterize the sequence enantiomer, as $^{15}N,^{13}C$-D-Gln was not readily available. The resulting molecule had the identical, but inverse, structure and stability as measured by CD (Figure S1, Supporting Information). The NMR structure of this variant showed the presence of similar capping interactions as modeled by protCAD. Low-energy structures in the final ensemble presented a D-Gln10 hydrogen bond with the backbone carbonyl from Lys8 (Figure 6A). Structures in which the D-Gln side chain faced away from the backbone were also consistent with experimental restraints but had higher energy scores. The HSQC spectrum of residue 10 in G10DQ (Figure 6B) presented two well-defined peaks for each of the pairs Cβ−Hβ and Cγ−Hγ, suggesting a unique environment for these atoms and limited mobility of the side chain.

Cα chemical shifts for the WT, D-Ala, and D-Asn variants were the same within experimental error, indicating that all peptides studied had the same fold.

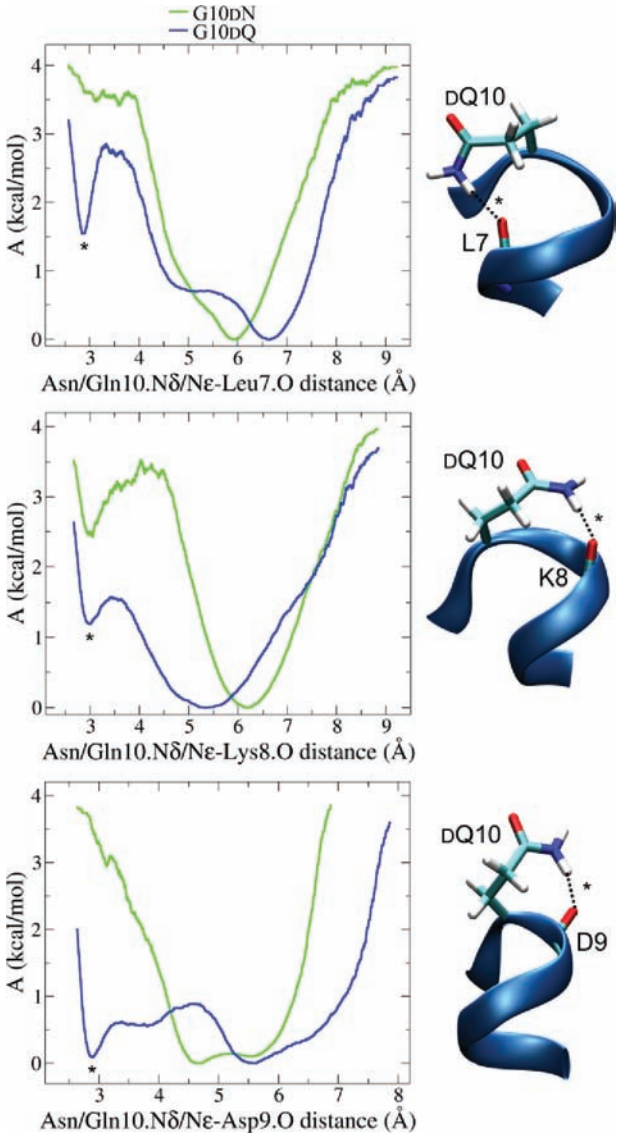


Figure 7. Helmholtz free energy (A, kcal/mol, eq 3) as a function of the distance between the residue 10 side chain (Nδ in G10DN, green; Nε in G10DQ, blue) and the backbone carbonyl of Leu7, Lys8, and Asp9 obtained from the MD simulations. Representative snapshots from the simulation of G10DQ in which the different capping hydrogen bond interactions are formed are shown.

3.4. Simulating Contributions of Side Chain Capping to Stability. The solution structure of G10DQ and molecular simulations support the model that D-Gln participates in side chain capping interactions. On this basis, we predict that the increase in stability in G10DQ relative to the WT is due to the presence of these side chain to backbone hydrogen bonds. In theory, D-Asn could participate in similar capping interactions, since Asn and Gln side chains have the same functional group, differing only by a methylene in the side chain, and modeling indicates that D-Asn can be incorporated and present side chain−backbone hydrogen bonds. If this is the case, why is G10DQ more stable than G10DN?

In agreement with the NMR data, MD simulations of TC5b G10DQ show that D-Gln can hydrogen bond with the backbone carbonyl of Leu7, Lys8, and Asp9 (insets, Figure 7). However, these interactions were not present in the average structure. Cluster analysis of the conformation of the D-Gln side chain identified 16 clusters. Two clusters accounting for ∼8.2% and ∼1.9% of the total population presented strong hydrogen bonding to backbone carbonyls (the heavy atom distance and angle in each cluster centroid structure were 3.1 Å and 144° to Asp9 and 2.8 Å and 172° to Leu7). Clusters that represented ∼19.5% and ∼2.3% of the structures hydrogen bond weakly (3.2 Å and 115° to Asp9 and 3.5 Å and 98° to Lys8).

As opposed to G10DQ, cluster analysis of the conformation of the D-Asn10 side chain from structures obtained from the MD simulation of G10DN identified 11 clusters, of which only 1 cluster that represented ∼1% of the structures hydrogen bonds strongly (3.5 Å and 150° with Lys8).

The free energy surface along a given coordinate is known as a PMF. Figure 7 shows the PMF for the distance between the side chain of residue 10 and the backbone carbonyl of Leu7, Lys8, and Asp9 for G10DN and G10DQ. These profiles show that, in general, the global minimum does not correspond to the capping interaction (distance <3.5 Å) and is between ∼5 and ∼7 Å. This agrees with the lack of side chain capping interactions in the average structures.

In the case of the PMF along the distance between D-Gln10 and Asp9, the capping interaction is very close in energy to the global minimum ($\Delta E \approx 0.1$ kcal/mol), with a barrier (E_a) of less than 1 kcal/mol, and is present in ∼15% of the simulation. In the case of the distance between D-Gln10 and Lys8, the minima are separated by ∼1 kcal/mol, with a barrier of ∼1.5 kcal/mol, and the hydrogen bond is present in ∼4% of the simulation. These low barriers allow rapid interconversion between the different conformations. In the case of the distance between D-Gln10 and Leu7, the minima are separated by ∼1.5 kcal/mol with an E_a of forming the capping interaction greater than 2.5 kcal/mol, and the interaction is present in only ∼2.2% of the simulation.

In comparison, the barriers of forming the capping interactions in G10DN are significantly greater (E_a greater than 3.5 kcal/mol). In this case, the PMF of only Lys8 shows a minimum that corresponds to the capping interaction, with a $\Delta E \approx 2.5$ kcal/mol, and this interaction is only present in ∼0.8% of the simulation. On the other hand, the hydrogen bonds with Leu7 and Lys9 are present in 0.08 and 0.2% of the simulation, respectively. Therefore, D-Asn forms only one of the three possible capping interactions.

Overall, these results suggest that the side chain capping interactions might be important for stability, since G10DQ is more stable and can make this type of interaction as opposed to G10DN.

3.5. Capping Interactions versus Entropic and Solvation Costs. If the side chain capping interactions are important for stability as suggested above, why then is G10DA more stable than G10DN? Even if G10DN does not certainly engage in capping interactions as frequently as G10DQ, it does it more frequently than G10DA, which completely lacks this ability. Therefore, other factors should contribute to the destabilization of G10DN with respect to G10DA.

While providing a network of stabilizing hydrogen bonds, helix-capping interactions must overcome unfavorable loss of side chain entropy and desolvation of backbone carbonyls in the C3−C1 and Ccap positions. The existence of a barrier is explicitly evident in the PMF plots of side chain−backbone hydrogen bonds. Position 10 in the Trp-cage is located at the end of the helix and is surface exposed. Therefore, this residue will

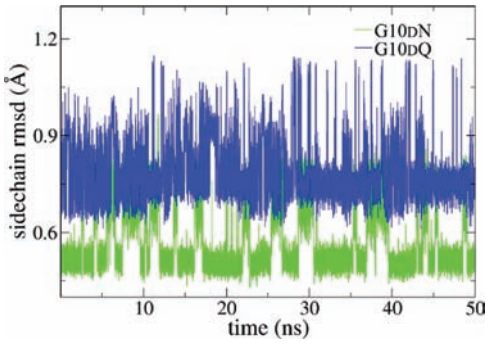


Figure 8. Mass-weighted rmsd's (Å, with respect to the average structure) of the side chain heavy atoms of residue 10 as a function of the production run (ns): G10DN variant (green) and G10DQ variant (blue).

Table 3. Entropic Cost (ΔS) of Fixing the Side Chain of D-Asn and D-Gln upon Folding of the Trp-Cage G10DN and G10DQ Variants[a]

	conformational	vibrational
D-Asn	0.29	0.01
D-Gln	0.32	−0.04

[a] The conformational entropy is given by the side chain rotamers, and the vibrational entropy is given by the quasi-harmonic entropy. Values are expressed as $T\Delta S$, where T is 300 K, in kcal/mol. Reported values are the mean over four calculations (Table S2, Supporting Information).

potentially be able to "freely" move and also to interact with the solvent. The backbone carbonyls of residues 7, 8, and 9 are also solvent exposed. For these reasons, the penalty of forming a capping interaction might include the entropic cost of fixing the side chain of residue 10 in a particular conformation and the desolvation of both the side chain and the backbone carbonyl of the residue that it is interacting with. The potential barriers of side chain entropy and solvation are treated separately.

Conformational restraints make capping interactions unfavorable in G10DN. Somewhat surprisingly, although the backbone fluctuation of residue 10 is smaller in G10DQ (Figure 3D), the side chain fluctuation (measured as the mass-weighted rmsd with respect to the average structure) is greater than in G10DN (Figure 8). Since D-Asn is in a relatively fixed conformation, it is more difficult for this residue to engage in capping interactions. On the other hand, the D-Gln10 side chain is more flexible, and thus, it is more probable for this residue to adopt conformations that would allow capping interactions.

To estimate the entropic cost (ΔS) of fixing the side chain in the folded state, we estimated the change in conformational and vibrational entropy between the unfolded and folded states. The folded state is considered as the ensemble of structures from the MD simulation of the corresponding variant, G10DN or G10DQ. Based on previous work,[56] the unfolded state was considered as the ensemble of structures from the MD simulation of a Gly-X-Gly peptide, where X = Asn or Gln. Although D-Asn is more rigid in G10DN than D-Gln in G10DQ (Figure 8), the ΔS upon folding is similar in both cases (Table 3). As suggested previously,[57] there is essentially no change in the vibrational entropy upon folding, and the entropic cost of folding is mainly due to a loss of available conformations or rotamers. In any case, the entropic cost is small, probably because both side chains are fully solvent exposed.

Table 4. Backbone Carbonyl (BB.O) Solvation of Leu7, Lys8, and Asp9 of the Different Trp-Cage Variants from the MD Simulations[a]

	number of waters			time fraction of HB (%)		
	L7	K8	D9	L7	K8	D9
WT	1.3 ± 0.5	2.2 ± 0.7	2.0 ± 0.8	116	178	153
G10DA	1.2 ± 0.4	2.0 ± 0.7	1.9 ± 0.7	107	170	149
G10DN	1.1 ± 0.4	1.9 ± 0.6 (1.4 ± 0.6)	2.0 ± 0.8	103	162	156
G10DQ	1.2 ± 0.5 (0.5 ± 0.6)	1.9 ± 0.7 (1.3 ± 0.6)	1.9 ± 0.7 (1.2 ± 0.6)	103	156	134

[a] The results from the first three columns are expressed as the mean number of waters (± standard deviation) within 3 Å of the corresponding atoms. When a side chain to backbone hydrogen bond (HB) capping interaction exists (G10DN with K8 and G10DQ with L7, K8, and D9), the values reported correspond to the cases in which there is no capping interaction, and in parentheses are the values when the interaction is formed. The capping interaction was defined to be present and absent when the distance between the residue 10 side chain (Nδ in G10DN and Nε in G10DQ) and the BB.O of the HB acceptor was smaller than 3.5 Å and greater than 4 Å, respectively. Time fraction (%) of HBs defined between solvent and BB.O; cutoffs were 3.5 Å and 120°. Values greater than 100 reflect multiple solvent donors.

Similarly, we can assume that the side chain ΔS upon folding equals zero for D-Ala in G10DA, since there are no side chain rotamers for Ala.

In Table 4, we can see that the presence of a D-amino acid in position 10 does not greatly reduce the backbone solvation of Leu7, Lys8, or Asp9, since G10DA has only slightly lower solvation than the WT. However, the presence of side chain to backbone capping interactions reduces backbone solvation of these three residues. For example, in all variants, including WT, there is on average more than one water molecule near (<3 Å) the carbonyl of Leu7 when there is no side chain capping interaction. When this interaction is formed in G10DQ, less than 1 water molecule solvates the carbonyl of Leu7. Similarly, backbone desolvation penalties are also observed for Lys8 and Asp9 in G10DQ, and for Lys8 in G10DN. In terms of the side chain of residue 10, there is only a desolvation penalty for D-Gln10 interaction with Leu7 in G10DQ (data not shown). To note, in G10DQ, the side chain capping interaction with Leu7 has the highest energy barrier in the free energy profiles (Figure 7). Therefore, a higher desolvation penalty of G10DQ when it interacts with Leu7 might explain the higher energy cost of this interaction.

4. CONCLUSIONS

The largest per-residue gains in Trp-cage stability have resulted from D-amino acid substitutions at α_L-glycines as compared to other molecular strategies: computational optimization of core interactions,[58] stabilizing proline by stereospecific fluorination, and adding helix favoring amino acids[32] and amino-terminus capping interactions.[54] Andersen and colleagues substituted D-Ala at position 10 in the TC10b variant, raising its T_M from 56 to 72 °C.[18] A similar magnitude of stabilization was found here for the TC5b, suggesting D-amino acid substitutions can be combined with other molecular strategies in an additive way, leading to hyperstable peptides.

Given the success of the Trp-cage designs, it is important to consider whether this approach can be generally applied. Placing D-Ala to the C-terminus of a monomeric α-helix did not enhance stability,[20] suggesting that much of the stability gain comes from reducing the flexibility of turns between secondary structure elements. Stereochemical stabilization of loops/turns between α-helices[59] and β-strands[12,60] is an established approach to improving fold stability. Gly10 is a hinge between the α-helical and polyproline segments of the Trp-cage fold. In all C′ L-Gln sites identified in the mirror-image template search of the PDB, the capping residue is located at the link between the α_L turn and another secondary structural element.[25] Much of the enhanced stability of D-amino acid substitutions may come from favoring an intermediate in folding that brings elements together.

In addition to location within the fold, a few studies have shown that the choice of D-amino acid influences the magnitude of stabilization.[17,20,61] A series of C-capping auxiliaries terminated by an amino or guanidino group varied in their contributions to the stability of a monomeric α-helix depending on the alkyl chain length.[20] It was suggested that both helix−dipole interactions and specific hydrogen bonds were responsible for enhancing stability. This was confirmed in the case of a D-Arg substitution, where a 0.4 kcal/mol increase in stability was observed when electrostatic interactions were screened by 2 M NaCl. Of the 10 L-Gln-capped α_L-helices in the mirror-image template PDB search, 6 participated in side chain to backbone hydrogen bonds, and the remaining side chains were involved in tertiary packing interactions. D-Gln may therefore be a generally applicable C-capping moiety. It may find particular utility in stabilizing monomeric helices where capping hydrogen bonds provide additional stability—an assertion that will require experimental validation.

One of the broader goals of this study was to develop protocols for predicting the stability of D-amino acid substitutions by MD and mirror-template bioinformatic methods. Overall, the experimental thermal stabilities agree with the stability trend predicted by the simulations: G10DQ > G10DA > G10DN > WT > G10A/Q. We observed that the magnitude of backbone fluctuations of amino acids surrounding the substitution correlates with stability: the most stable variants are also the most rigid. A reduction in backbone motion would result from a deeper energy minimum corresponding to the folded, native state.

On the other hand, some portion of the stability conferred by D-amino acid substitutions presumably results from promoting the formation of nativelike transition-state intermediates during folding, and destabilizing the unfolded state by reducing configurational entropy. To effectively model these phenomena, it will be necessary to thoroughly sample partially folded and unfolded states in the conformational ensemble. In G-X-G tripeptide simulations, the α_L conformation was significantly populated when X = L-Asn, but unoccupied for L-Gln (Figure S7, Supporting Information), consistent with helix-propensity studies.[62] The folding pathway plays an important role in addition to the native-state ensemble effects studied here.

Carrying out these simulations on larger peptides and small proteins is becoming tractable with replica-exchange methods[63] utilizing high-performance computing resources. Such intensive simulations are currently impractical as a direct step in sequence-space search algorithms utilized during computational protein design. Instead, they provide a better molecular understanding of the subtle contributions of solvation, side chain entropy, electrostatics, and hydrogen bonding. This can be used to refine design rules for incorporating stabilizing non-natural amino acid substitutions.

■ ASSOCIATED CONTENT

Supporting Information. Chemical shift ambiguity index value definitions, tables giving the enthalpy and heat capacity obtained from global fitting of thermal/GnHCl denaturation of Trp-cage variants to eqs 2a−2d and the conformational entropic cost of fixing the side chain of D-Asn and D-Gln upon folding of the Trp-cage G10DN and G10DQ variants, figures showing the far-UV CD spectra of the G10DQ Trp-cage and the labeled Trp-cage with inverse chirality, snapshots from the MD simulations of G10A and G10Q, normalized tryptophan emission spectra of Trp-cage variants, thermally induced unfolding of the Trp-cage variants, DSC thermograms of the different Trp-cage variants, and backbone torsion angle distribution for L-Asn and L-Gln in the GXG tripeptide MD simulations, and complete ref 44. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
nanda@cabm.rutgers.edu; rlkoder@gmail.com

Author Contributions
§These authors contributed equally.

■ ACKNOWLEDGMENT

We thank the office of Information Services and Technology/High Performance and Research Computing at UMDNJ for access to computational resources and Hsin Wang, of the Department of Chemistry, City College of New York (CCNY), for assistance with the NMR measurements. V.N., S.A., and A.R.G. were supported by a Seed Grant from the UMDNJ Foundation, NSF Grant MCB-0940187, and NIH Grant R01 GM089949-01. L.Z. and R.L.K. gratefully acknowledge support by NSF Grant MCB-0920448, infrastructure support from NIH Grant P41 GM-66354 to the New York Structural Biology Center, and a grant from the NIH National Center for Research Resources to CCNY (NIH 5G12 RR03060).

■ REFERENCES

(1) Zhou, N. E.; Kay, C. M.; Hodges, R. S. *Biochemistry* **1993**, *32*, 3178.
(2) Giblin, M. F.; Wang, N.; Hoffman, T. J.; Jurisson, S. S.; Quinn, T. P. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12814.
(3) Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. *Angew. Chem., Int. Ed.* **2001**, *40*, 1494. Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, I. W.; DeGrado, W. F.; Tirrell, D. A. *Biochemistry* **2001**, *40*, 2790. Naduthambi, D.; Zondlo, N. J. *J. Am. Chem. Soc.* **2006**, *128*, 12430.
(4) Kemp, D. S.; Boyd, J. G.; Muendel, C. C. *Nature* **1991**, *352*, 451.
(5) Bracken, C.; Gulyas, J.; Taylor, J. W.; Baum, J. *J. Am. Chem. Soc.* **1994**, *116*, 6431–6432. Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. *Science* **2004**, *305*, 1466.
(6) Chapman, R. N.; Dimartino, G.; Arora, P. S. *J. Am. Chem. Soc.* **2004**, *126*, 12252.
(7) Chandrasekaran, R.; Ramachandran, G. N. In *2nd American Peptide Symposium*, 1st ed.; Lande, S., Ed.; Gordon and Breach: Cleveland, OH, 1970; Vol. 2, p 195.

(8) Ramachandran, G. N.; Ramakrishnan, C.; Sasisekharan, V. *J. Mol. Biol.* **1963**, *7*, 95.
(9) Karle, I. L.; Gopi, H. N.; Balaram, P. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13946. Karle, I. L. *Biopolymers* **2001**, *60*, 351. Banerjee, A.; Raghothama, S. R.; Karle, I. L.; Balaram, P. *Biopolymers* **1996**, *39*, 279. Karle, I. L.; Flippenanderson, J. L.; Uma, K.; Balaram, P. *Biopolymers* **1993**, *33*, 401. Karle, I. L.; Flippenanderson, J. L.; Uma, K.; Balaram, P. *Biochemistry* **1989**, *28*, 6696.
(10) Veber, D. F.; Holly, F. W.; Nutt, R. F.; Bergstrand, S. J.; Brady, S. F.; Hirschmann, R.; Glitzer, M. S.; Saperstein, R. *Nature* **1979**, *280*, 512. Veber, D. F.; Holly, F. W.; Paleveda, W. J.; Nutt, R. F.; Bergstrand, S. J.; Torchiana, M.; Glitzer, M. S.; Saperstein, R.; Hirschmann, R. *Proc. Natl. Acad. Sci. U.S.A.* **1978**, *75*, 2636.
(11) Stanger, H. E.; Gellman, S. H. *J. Am. Chem. Soc.* **1998**, *120*, 4236. Imperiali, B.; Fisher, S. L.; Moats, R. A.; Prins, T. J. *J. Am. Chem. Soc.* **1992**, *114*, 3182. Yan, Y. B.; Erickson, B. W.; Tropsha, A. *J. Am. Chem. Soc.* **1995**, *117*, 7592.
(12) Haque, T. S.; Little, J. C.; Gellman, S. H. *J. Am. Chem. Soc.* **1996**, *118*, 6975.
(13) Rana, S.; Kundu, B.; Durani, S. *Chem. Commun.* **2005**, 207. Rana, S.; Kundu, B.; Durani, S. *Chem. Commun.* **2004**, 2462. Struthers, M. D.; Cheng, R. P.; Imperiali, B. *Science* **1996**, *271*, 342.
(14) Valiyaveetil, F. I.; Sekedat, M.; Mackinnon, R.; Muir, T. W. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17045.
(15) Hermans, J.; Anderson, A. G.; Yun, R. H. *Biochemistry* **1992**, *31*, 5646. Fairman, R.; Anthony-Cahill, S. J.; DeGrado, W. F. *J. Am. Chem. Soc.* **1992**, *114*, 5458.
(16) Anil, B.; Song, B.; Tang, Y.; Raleigh, D. P. *J. Am. Chem. Soc.* **2004**, *126*, 13194.
(17) Bang, D.; Gribenko, A. V.; Tereshko, V.; Kossiakoff, A. A.; Kent, S. B.; Makhatadze, G. I. *Nat. Chem. Biol.* **2006**, *2*, 139.
(18) Williams, D. V.; Barua, B.; Andersen, N. H. *Org. Biomol. Chem.* **2008**, *6*, 4287.
(19) Schellman, C. In *Protein Folding*; Jaenicke, R., Ed.; Elsevier/North-Holland: New York, 1980; p 53. Presta, L. G.; Rose, G. D. *Science* **1988**, *240*, 1632. Richardson, J. S.; Richardson, D. C. *Science* **1988**, *240*, 1648. Aurora, R.; Srinivasan, R.; Rose, G. D. *Science* **1994**, *264*, 1126.
(20) Schneider, J. P.; DeGrado, W. F. *J. Am. Chem. Soc.* **1998**, *120*, 2764.
(21) Nanda, V.; Degrado, W. F. *J. Am. Chem. Soc.* **2004**, *126*, 14459.
(22) Watanabe, A.; Yoshimura, T.; Mikami, B.; Hayashi, H.; Kagamiyama, H.; Esaki, N. *J. Biol. Chem.* **2002**, *277*, 19166.
(23) Serrano, L.; Fersht, A. R. *Nature* **1989**, *342*, 296. Aurora, R.; Rose, G. D. *Protein Sci.* **1998**, *7*, 21. Gunasekaran, K.; Nagarajaram, H. A.; Ramakrishnan, C.; Balaram, P. *J. Mol. Biol.* **1998**, *275*, 917. Lu, M.; Shu, W.; Ji, H.; Spek, E.; Wang, L.; Kallenbach, N. R. *J. Mol. Biol.* **1999**, *288*, 743. Altschuler, E. L. *Med. Hypotheses* **2001**, *56*, 478.
(24) Mitchell, J. B.; Smith, J. *Proteins* **2003**, *50*, 563.
(25) Annavarapu, S.; Nanda, V. *BMC Struct. Biol.* **2009**, *9*, 61.
(26) Kleywegt, G. J. *J. Mol. Biol.* **1999**, *285*, 1887. Watson, J. D.; Milner-White, E. J. *J. Mol. Biol.* **2002**, *315*, 171.
(27) Engel, D. E.; DeGrado, W. F. *J. Mol. Biol.* **2004**, *337*, 1195.
(28) Thomas, S. T.; Loladze, V. V.; Makhatadze, G. I. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 10670.
(29) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425.
(30) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. *J. Am. Chem. Soc.* **2002**, *124*, 12952. Nikiforovich, G. V.; Andersen, N. H.; Fesinmeyer, R. M.; Frieden, C. *Proteins* **2003**, *52*, 292. Chowdhury, S.; Lee, M. C.; Xiong, G.; Duan, Y. *J. Mol. Biol.* **2003**, *327*, 711. Ota, M.; Ikeguchi, M.; Kidera, A. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17658. Steinbach, P. J. *Proteins* **2004**, *57*, 665. Kaledin, M.; Brown, A.; Kaledin, A. L.; Bowman, J. M. *J. Chem. Phys.* **2004**, *121*, 5646.
(31) Zhou, R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 13280. Pitera, J. W.; Swope, W. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 7587.
(32) Lin, J. C.; Barua, B.; Andersen, N. H. *J. Am. Chem. Soc.* **2004**, *126*, 13679.
(33) Simmerling, C.; Strockbine, B.; Roitberg, A. E. *J. Am. Chem. Soc.* **2002**, *124*, 11258. Snow, C. D.; Zagrovic, B.; Pande, V. S. *J. Am. Chem. Soc.* **2002**, *124*, 14548. Kannan, S.; Zacharias, M. *Proteins* **2009**, *76*, 448. Day, R.; Paschek, D.; Garcia, A. E. *Proteins* **2010**, *78*, 1889.
(34) Streicher, W. W.; Makhatadze, G. I. *Biochemistry* **2007**, *46*, 2876.
(35) Wafer, L. N.; Streicher, W. W.; Makhatadze, G. I. *Proteins* **2010**, *78*, 1376.
(36) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *J. Chem. Phys.* **1979**, *71*, 4546.
(37) Braunschweiler, L.; Ernst, R. R. *J. Magn. Reson.* **1983**, *53*, 521.
(38) Rucker, S. P.; Shaka, A. J. *Mol. Phys.* **1989**, *68*, 509.
(39) Bodenhausen, G.; Ruben, D. J. *Chem. Phys. Lett.* **1980**, *69*, 185.
(40) Rieping, W.; Habeck, M.; Bardiaux, B.; Bernard, A.; Malliavin, T. E.; Nilges, M. *Bioinformatics* **2007**, *23*, 381.
(41) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr.* **1998**, *54*, 905.
(42) Summa, C. M. *protCAD*; University of Pennsylvania School of Medicine: Philadelphia, PA, 2002.
(43) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179.
(44) Case, D. A.; et al. *Amber 10*; University of California: San Francisco, CA, 2008.
(45) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
(46) Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. *Proteins* **2006**, *65*, 712.
(47) Berendsen, H. J.; Postma, J. P.; van Gunsteren, W. F.; Di Nola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684.
(48) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327.
(49) Feig, M.; Karanicolas, J.; Brooks, C. L., 3rd. *J. Mol. Graphics Modell.* **2004**, *22*, 377.
(50) Leach, A. R. *Molecular Modelling. Principles and Applications*; Addison-Wesley Longman, Pearson Education Ltd.: London, England, 2001.
(51) Schlitter, J. *Chem. Phys. Lett.* **1993**, *215*, 617. Andricioaei, I.; Karplus, M. *J. Chem. Phys.* **2001**, *115*, 6289. Brooks, B. R.; Janei, D.; Karplus, M. *J. Comput. Chem.* **2004**, *16*, 1522.
(52) Dixit, S. B.; Andrews, D. Q.; Beveridge, D. L. *Biophys. J.* **2005**, *88*, 3147.
(53) Walsh, S. T.; Sukharev, V. I.; Betz, S. F.; Vekshin, N. L.; DeGrado, W. F. *J. Mol. Biol.* **2001**, *305*, 361.
(54) Barua, B.; Lin, J. C.; Williams, V. D.; Kummler, P.; Neidigh, J. W.; Andersen, N. H. *Protein Eng. Des. Sel.* **2008**, *21*, 171.
(55) Makhatadze, G. I.; Privalov, P. L. *J. Mol. Biol.* **1993**, *232*, 639.
(56) Creamer, T. P. *Proteins* **2000**, *40*, 443.
(57) Karplus, M.; Ichiye, T.; Pettitt, B. M. *Biophys. J.* **1987**, *52*, 1083.
(58) Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. *J. Phys. Chem. B* **2006**, *110*, 3759.
(59) Lahr, S. J.; Engel, D. E.; Stayrook, S. E.; Maglio, O.; North, B.; Geremia, S.; Lombardi, A.; DeGrado, W. F. *J. Mol. Biol.* **2005**, *346*, 1441. Effimov, A. V. *Curr. Opin. Struct. Biol.* **1993**, *3*, 379.
(60) Sibanda, B. L.; Thornton, J. M. *Nature* **1985**, *316*, 170.
(61) Narayana, N.; Phillips, N. B.; Hua, Q. X.; Jia, W.; Weiss, M. A. *J. Mol. Biol.* **2006**, *362*, 414.
(62) O'Neil, K. T.; DeGrado, W. F. *Science* **1990**, *250*, 646. Hovmoller, S.; Zhou, T.; Ohlson, T. *Acta Crystallogr., D: Biol. Crystallogr.* **2002**, *58*, 768.
(63) Okamoto, Y.; Sugita, Y. *Chem. Phys. Lett.* **1999**, *314*, 141.