

RESEARCH ARTICLE

Effect of heterochiral inversions on the structure of a β -hairpin peptide

Gül H. Zerze¹ | Frank H. Stillinger² | Pablo G. Debenedetti¹

¹Department of Chemical and Biological Engineering, Princeton University, Princeton, New Jersey

²Department of Chemistry, Princeton University, Princeton, New Jersey

Correspondence

Pablo G. Debenedetti, Department of Chemical and Biological Engineering, Princeton University, Princeton, NJ 08544.
Email: pdebene@princeton.edu

Abstract

We study computationally a family of β -hairpin peptides with systematically introduced chiral inversions, in explicit water, and we investigate the extent to which the backbone structure is able to fold in the presence of heterochiral perturbations. In contrast to the recently investigated case of a helical peptide, we do not find a monotonic change in secondary structure content as a function of the number of L- to D-inversions. The effects of L- to D-inversions are instead found to be highly position-specific. Additionally, in contrast to the helical peptide, some inversions increase the stability of the folded peptide: in such cases, we compute an increase in β -sheet content in the aqueous solution equilibrium ensemble. However, the tertiary structures of the stable (folded) configurations for peptides for which inversions cause an increase in β -sheet content show differences from one another, as well as from the native fold of the non-chirally perturbed β -hairpin. Our results suggest that although some chiral perturbations can increase folding stability, chirally perturbed proteins may still underperform functionally, given the relationship between structure and function.

KEYWORDS

advanced sampling, all-atom simulations, chiral inversion, D-amino acids, GB1, β -hairpin folding

1 | INTRODUCTION

The asymmetric tetrahedral carbon atom of the backbones of all proteinogenic amino acids except glycine allows these essential building blocks in principle to exist in two nonsuperimposable forms of the same constitution, commonly referred to as L- and D-amino acids. Although D-amino acids exist in nature,¹⁻¹⁰ ribosomally synthesized proteins in all living organisms are exclusively composed of L-amino acids. Homochirality is essential to the proper functioning of proteins: molecular recognition and enzymatic specificity rely on it. Another aspect that protein function relies upon is folding into a native structure. The possibility of structural changes introduced by chiral perturbations has been inferred by reverse phase chromatography, circular dichroism, and nuclear magnetic resonance experiments, and by molecular dynamics and Monte Carlo simulations.¹¹⁻¹⁵ However, the relationship between the extent and location of monomer-level chiral perturbations and folding in aqueous solution has not been studied systematically until recently. In a recent publication, we investigated

computationally the effect of systematic backbone chiral perturbations on the structure of an α -helical peptide and found that, depending on its location along the backbone, even the smallest degree of heterochirality can significantly affect the folding ability of the peptide.¹

In the present work, we aim to broaden the scope of our numerical investigation by studying a mini-protein composed of a β -pleated sheet in an antiparallel arrangement, connected with a reverse turn. The 16-residue long GB1 peptide, the C-terminal hairpin (residues 41-56) of streptococcal Protein G's B1 domain,^{16,17} is a representative of a large family of peptides featuring the β -hairpin structural motif. The folding mechanism of β -hairpins has been studied in the literature, both experimentally and via simulations.¹⁸⁻⁴⁸ Two key events have been identified, namely formation of a stable turn and formation of cross-strand hydrogen bonds and hydrophobic contacts. Earlier experimental evidence interpreted by a statistical mechanical model suggests a "zipper" mechanism of hairpin folding, where the native backbone hydrogen bonds form sequentially starting from the turn,

and progressing toward the termini.^{18,19} It has been shown computationally that the GB1 hairpin can fold following the zipper mechanism, but it can also fold following a “pincer” mechanism, in which hydrogen bonds form sequentially from the termini to the turn.⁴⁶ Regardless of whether the hairpin folds following zipper or pincer mechanisms, the correct formation of the turn has been shown to be the initial folding step.⁴⁶ Formation of a turn relies on the adoption of certain torsion angles and local hydrogen bonds by the backbone. In addition to the angle requirements for the central turn residues, $i + 1$ and $i + 2$, where the first turn residue has the index i , the $i + 3$ position of GB1's turn is required to adopt an α_L configuration¹⁶ which is a rare event for an L-amino acid, as positive ϕ -angles are energetically unfavorable. For glycine, an achiral amino acid, positive and negative backbone torsion angles of the same magnitude are equally favorable, and it is therefore one of the preferred mutations in such a position when the goal is to increase the stability of the protein folds.^{49–51} The glycine mutation approach has been shown to work for the turn of GB1.^{34,52} The positive ϕ angles being more favorable for D-amino acids, they have also, accordingly, been suggested as possible candidates for mutations in such positions in order to increase the folding stability.^{53,54} One of the many chiral inversions that we investigate in this work addresses specifically the effect of mutated chirality at the $i + 3$ turn position.

Being amenable to atomistic simulations, the GB1 hairpin allows us to investigate systematically the effect of backbone chirality on a foldable fragment of a real protein. In this work, we address the effect of heterochiral perturbations on one of the most prevalent protein structural motifs after α -helices (which we recently investigated¹). Studying 32 different GB1 peptides with varied backbone chiralities, we find that, unlike a purely helical peptide, which shows a monotonic change in the helical fraction with respect to the number of L- to D-inversions, there is no analogous monotonic trend in the secondary structure content of GB1. In addition, we find that chirally perturbed peptides show a variety of different hairpin folds (when they are still able to fold) some of which are more stable, possessing increased β -sheet content, than the native fold of GB1. We clarify here that we associate increases in specific secondary structure content with increases in folding stability since a shift in the folding equilibrium toward the folded state increases secondary structure content. From the protein design perspective, our results confirm that certain chiral inversions can make highly stable protein folds. However, we also see concomitant changes in tertiary structure, suggesting that the expected function may not be fully conserved.

2 | METHODS

2.1 | Inversion pattern design and peptide initial configurations

The native-fold of GB1 is illustrated in Supporting Information Figure S1. There are two main structural elements: β -strands stabilized by backbone hydrogen bonds between two strands, and a reverse turn. Accordingly, we consider chiral inversions in both the turn and the β -sheets. We further classify the inversions in the β -sheet portion

of the peptide as cross-strand matching mutations (hydrogen-bonded residue pairs in the native fold) and mismatching inversions, as shown in Table 1. Since only 15 out of GB1 hairpin's 16 residues are chiral amino acids, the maximum number of L- to D-inversions is 15, and the peptide with 15 L- to D-inversions is the inversed-peptide of the naturally occurring GB1 domain.

Following a similar approach to that of our earlier work,¹ we impose the chirality of asymmetric tetrahedral carbon atoms as an initial condition. Peptide initial conditions are generated by the CHARMM program⁵⁵ as extended (unfolded) configurations, using the SwissParam⁵⁶ database for the CHARMM internal coordinates table, and are relaxed in gas-phase simulations for 500 steps. The monomer of each peptide is solvated in a truncated octahedron box containing 2724 water molecules and 3 Na⁺ ions, which are included in order to neutralize the net charge of the peptide.

2.2 | Model

Peptides are modeled using the Amber03* protein force field⁵⁷ and water is modeled using the TIP3P model.⁵⁸ This particular protein/water model combination has been successfully used to infer the folding mechanisms and thermodynamic properties of hairpins, yielding close agreement with experiments.^{46,52}

The force field contains an asymmetric dihedral angle term.⁵⁷ We simply adjust this term for D-amino acids by inverting the sign of the phase shift angle of the term and we generate the D-configuration internal coordinates as an initial condition for mutated residues. We simulate the systems using the GROMACS 2016.3^{59,60} MD engine and the PLUMED 2.3.1 plugin for metadynamics calculations.⁶¹ GROMACS-compatible force field files are available upon request.

2.3 | Enhanced sampling

To ensure the equilibrium sampling of each peptide in aqueous solution, we use a combination of parallel tempering and well-tempered metadynamics, specifically employing parallel sampling in the well-tempered ensemble (PTWTE).^{62–64} We perform PTWTE simulations for at least 500 ns/replica for each peptide. Convergence is monitored by the cumulative average of the β -sheet fraction for the 300 K replica for each peptide (Supporting Information Figures S2 and S3). The initial 250 ns/replica is treated as equilibration and excluded from the analysis for all peptides.

We run the PTWTE simulations using 14 replicas. The potential energy is biased using a value of 400 kJ/mol for the Gaussian width and 1.2 kJ/mol for the initial Gaussian height. Gaussian potentials are added every 2000 steps, with a bias factor of 15. Prior to starting the PTWTE simulations, unbiased simulations are performed at each temperature for 200 ps, in order to equilibrate the potential energy of the replicas. The 14 replica temperatures are distributed geometrically, spanning a range of from 300 to 475 K. For the given system size, this setting yields an average replica exchange acceptance rate of 30%. All runs are performed in the NPT ensemble. Results at 300 K are reported.

Systems are propagated using the leap-frog algorithm with a 2 fs time step. The temperature of each replica is maintained using the Nosé-Hoover thermostat^{65,66} with a 1 ps time constant. The pressure

TABLE 1 Primary structure of the peptides studied in this work [Colour table can be viewed at wileyonlinelibrary.com]

Category		Name	Sequence	No. of D-inversions
Wild-type		GB1	GEWTYDD DA TKTFTVTE	0
Turn positions inverted		GB1 t1	GEWTYD d ATKtFTVTE	1
		GB1 t2	GEWTYD Da TKTFTVTE	1
		GB1 t3	GEWTYD DA tKtFTVTE	1
		GB1 t4	GEWTYD DA TkTFTVTE	1
β -sheet positions inverted	Cross-strand matching pairs	1 pair		
		GB1 b11	GeWTYD DA TKTFTVtE	2
		GB1 b12	GEWtYD DA TKTftVTE	2
		GB1 b13	GEWTYd DA TKtFTVTE	2
		2 pairs-block		
		GB1 b21	GewTYD DA TKTftVtE	4
		GB1 b22	GEWtyD DA TKTftVTE	4
		2 pairs-alternating		
		GB1 b23	GEWtYD DA TKTftVtE	4
		GB1 b24	GEWtyD DA TKTftVTE	4
	3 pairs-block			
	GB1 b31	GewtYD DA TKTftvTE	6	
	3 pairs-block			
	GB1 b32	GEWtyD DA TKTftvTE	6	
	3 pairs-alternating			
	GB1 b33	GEWtYd DA TKtFtVtE	6	
	4 pairs-block			
	GB1 b4	GewtyD DA TKTftvtE	8	
	5 pairs-block			
	GB1 b5	Gewtyd DA TKtftvtE	10	
Mismatching	one strand			
	GB1 m11	GeWTYD DA TKTFTVTE	1	
	GB1 m12	GEWtYD DA TKTFTVTE	1	
	GB1 m13	GEWtYD DA TKTFTVTE	1	
	GB1 m14	GEWtYd DA TKTFTVTE	1	
	GB1 m15	GEWTYd DA TKTFTVTE	1	
	GB1 m16	GEWTYD DA TKTFTVtE	1	
	GB1 m2	GewTYD DA TKTFTVTE	2	
	GB1 m3	GewtYD DA TKTFTVTE	3	
	GB1 m4	GewtyD DA TKTFTVTE	4	
GB1 m5	Gewtyd DA TKTFTVTE	5		
cross-strand				
GB1 m6	GeWTYD DA TKTFTVtE	3		
GB1 m7	GEWtYD DA TKTftVtE	5		
GB1 m8	GewtyD DA TKTftvTE	9		
GB1 m9	Gewtyd DA TKtftvTE	11		
All inverted		GB1 d	Gewtyd dat ktftvTE	15

Residues in turn positions are in bold font. Lower-case characters represent the D-amino acids and upper-case letters denote the canonical L-amino acids. Color coding of the peptide name follows the classification in Figure 1.

of each replica is maintained at 1 bar using an isotropic Parrinello-Rahman barostat with a coupling constant of 2 ps. Electrostatic interactions are calculated using the particle-mesh Ewald method⁶⁷ with a real space cutoff distance of 1 nm. A 1 nm cutoff distance is used for the van der Waals interactions.

2.4 | Analysis

Secondary structure assignment for each configuration is performed using the DSSP algorithm, which is based on hydrogen bonding patterns.⁶⁸ β -sheet assignment is performed in two steps. This calculation iterates over each conformer in the ensemble. For each conformer, backbone hydrogen bonds are first analyzed to assign β -bridges for each residue. If any two nonoverlapping blocks of three residues ($i - 1, i, i + 1$ and $j - 1, j, j + 1$) form hydrogen bonds

between ($i - 1, j$ and $j, i + 1$) or ($j - 1, i$ and $i, j + 1$), a parallel bridge is assigned. If hydrogen bonds are formed between (i, j and j, i) or ($i - 1, j + 1$ and $j - 1, i + 1$), an antiparallel bridge is assigned for those residues. In the second step, contiguity is checked. If there is a set of two or more contiguous bridges of the same type (parallel or antiparallel), those residues are finally assigned to be in a β -sheet instantaneous configuration. For each conformer, the total number of residues assigned to be in a β -sheet configuration is counted and normalized by the total number of residues, yielding the β -sheet fraction. The average β -sheet fraction is then calculated as the ensemble average of the above-defined instantaneous quantity. The native structure of GB1, C-terminal β -hairpin extracted from the Protein Data Bank ID: 1GB1¹⁶ (Supporting Information Figure S1) has 12 out of its 16 residues in β -sheet configuration, which sets the highest expected β -sheet fraction to be 12/16, or 0.75.

Based on the DSSP algorithm, turn structure is assigned based on the hydrogen bonds between residues $i, i + n$ where n can be 3, 4, or 5. In the next stage where contiguity of the turn is checked, if the minimum number of consecutive n -turns criterion is satisfied, turn assignment is switched to a helix assignment which ensures that a turn is not a helical residue. At least two consecutive n -turns are needed for a helix assignment. Residues forming any n -turn and not satisfying contiguity criterion are finally assigned to be in a turn configuration. In this work, we only make β -sheet and turn assignments. Readers are referred to the work by Kabsch and Sander for other assignments.⁶⁸

We now define the order parameters used to compute free energy diagrams. The distance root-mean-square, d_{RMS} , is calculated using $d_{\text{RMS}} = [N_{\text{bb}}^{-1} \sum_{(ij)} (r_{ij} - r_{ij}^0)^2]^{1/2}$. The sum runs over the N_{bb} backbone native contact pairs (intramolecular) (i, j) which are separated by distance r_{ij} in the configuration of interest and by r_{ij}^0 in the native state. Based on the native structure (Supporting Information Figure S1), a contact between a backbone atom of residue a and a backbone atom of residue b is defined to be native if $|a - b| > 2$ and the distance between the two atoms is less than 4.5 Å.⁴⁵

The fraction of native-like hydrogen bonds, Q_{hb} , is calculated using the backbone donor and acceptor pairs of GB1's native fold, determined according to a distance criterion of 2.8 Å. There are eight backbone hydrogen bonds identified according to this distance criterion in the reference native fold of GB1 hairpin. The distance between each of these donor-acceptor pairs is calculated for each configuration and a hydrogen bond is assigned to be present if the donor-acceptor distance is smaller than 3.92 Å. The number of hydrogen bonds is then normalized by the number of hydrogen bonds in the native fold. Clustering is performed based on structural similarity of the backbone heavy, that is, non-hydrogen, atoms following the GROMOS algorithm⁶⁹ using a 0.15 nm root mean square deviation (RMSD) cutoff distance. For each conformer in the ensemble, the number of other conformers for which the RMSD is 0.15 nm or less is calculated. The structure with the highest number of similar conformers forms the

most populated cluster, together with the other conformers within the given cutoff distance. We report the most populated cluster with the percentage of the population (number of conformers in this cluster with respect to the total number of conformers).

3 | RESULTS AND DISCUSSION

In order to assess the structural changes introduced by L- to D-inversions, we quantify the overall peptide β -sheet fraction, based on the DSSP algorithm (see Section 2.4), as shown in Figure 1. In contrast to the helical peptide that we studied recently,¹ the GB1 hairpin does not show a monotonic change in the secondary structure content along the perturbation coordinate of number of L- to D-inversions. While the β -sheet fraction is the same for GB1 (0 L- to D-inversions) and GB1 d (15 L- to D-inversions) within statistical error, as expected, other inversions yield a variety of different β -sheet fractions, some of which are higher than that of GB1 (or GB1 d). An analogous observation was not present for the α -helical peptide that we studied recently, that is, the highest helicity was obtained for poly-L-alanine (α_R) or poly-D-alanine peptide (α_L),¹ and opposite chirality inversions perturbed the helical structure, following a monotonic trend as a function of number inversions.

Three of the four inversions belonging to the turn category cause a decrease in the β -sheet fraction of the peptide, namely, the GB1 t1, t2, and t3 peptides have reduced β -sheet fractions, whereas the GB1 t4 peptide has an increased β -sheet fraction compared to native GB1 (red symbols in Figure 1). The GB1 hairpin forms a type I turn according to the ϕ, ψ Ramachandran angles occupied by the $i + 1$ and $i + 2$ turn residues (4-residue turn, residues indexed from i to $i + 3$).⁷⁰ In addition, the turn structure in the GB1 hairpin requires the $i + 3$ turn residue to sample the torsion angles corresponding to the α_L region¹⁶ (Supporting Information Figure S1). Adoption of an α_L configuration in this position has been found to be strictly correlated with folding in atomistic simulations^{46,71} and has accordingly been

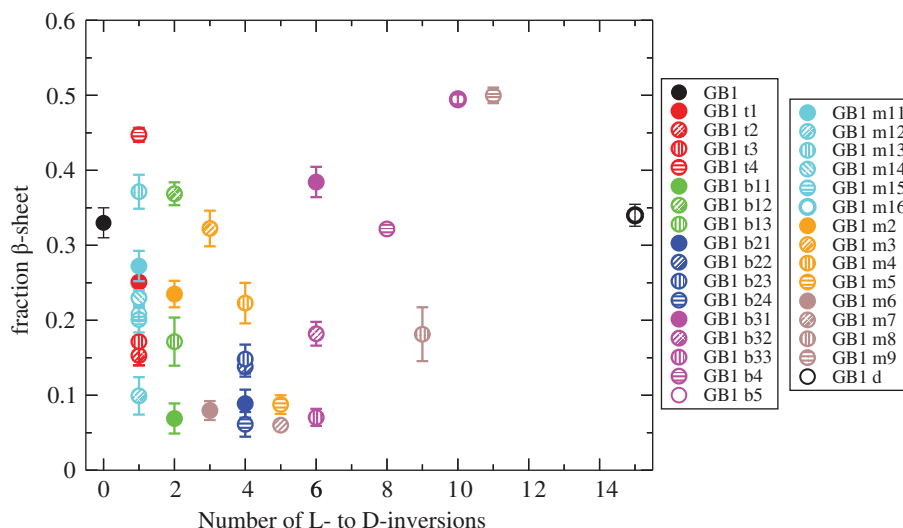


FIGURE 1 Average fraction of β -sheet content as a function of number of L- to D-inversions. Errors are calculated from blocked SEs, using two equal, nonoverlapping blocks of data ($T = 300$ K)

suggested to be one of the key events in the GB1 hairpin's folding. Moreover, replacing the L-lysine residue in this position with glycine has been found to increase the stability of the hairpin.^{34,52} This increased stability has been explained as follows. Adoption of an α_L configuration is a rare event for an L-lysine residue, given the steric constraints associated with an L-amino acid adopting torsional angles characteristic of its D-enantiomer. For glycine, on the other hand, the left (negative ϕ angles) and right (positive ϕ angles) sides of Ramachandran map are equally favorable; therefore, adoption of an α_L configuration is energetically favorable for glycine compared to L-lysine. The GB1 t4 peptide corresponds to a D-inversion at the turn's $i + 3$ position, that is, L-lysine is mutated to D-lysine. The GB1 t4 peptide is expected to have a more stable hairpin fold as positive ϕ angles are energetically more favorable to D-amino acids. The higher β -sheet fraction of the GB1 t4 peptide is consistent with this expectation. However, as we will discuss below, the GB1 t4 fold has certain differences compared to the native fold of the GB1 peptide.

As defined in Table 1, the inversions that we perform are not limited to the turn group. Figure 1 follows a color code classification consistent with Table 1, where red (turn), green (β -sheet, cross-strand matching, 1 pair), blue (β -sheet, cross-strand matching, 2 pairs), magenta (β -sheet, cross-strand matching, 3 or more pairs), cyan (β -sheet, one strand, 1 inversion), orange (β -sheet, one strand, 2-5 inversions), and brown symbols (β -sheet, cross-strand mismatching, 3 or more inversions) denote chirally perturbed peptides and black symbols (filled or empty) are homochiral peptides. We do not find any clear patterns of structural change within any of these classes. For example, within the red group (GB1 t1, t2, t3, and t4), GB1 t4 causes an increase in the β -sheet fraction, as discussed above, whereas the others caused a decrease. Among the single β -sheet inversions, represented by the cyan group (GB1 m11, m12, m13, m14, m15, and 16) no common effect emerges, either. One of them, GB1 m13, shows an increased β -sheet fraction compared to GB1, whereas the others cause a decrease in the β -sheet fraction. Similar observations are also true for cross-strand pairwise L- to D-inversions, denoted as GB1 bx peptides in Table 1. One of the single cross-strand pair inverted peptides (the green group in Figure 1), GB1 b12, shows an increased β -sheet fraction, while the others show the opposite behavior. This nonsystematic behavior continues as the number of inverted pairs is increased (GB1 b2x and GB1 b3x peptides). Pairwise alternating (GB1 b23, b24, b33) and mismatching alternating inversions (GB1 m7), however, always lead to a reduced β -sheet fraction. Additionally, the group of peptides where the number of L- to D-inversions is increased progressively by inverting each neighbor residue on one of the strands until the strand is completely composed of D-amino acids (GB1 m11, m2, m3, m4, and m5 peptides), shows no discernable progressive trend. The first two inversions, m11 and m2, lead to a decrease in the β -sheet fraction, whereas the next inversion (m3) causes an increase in the β -sheet fraction. The successive inversions, m4 and m5, again give rise to a decreasing β -sheet fraction trend.

Per-residue secondary structure propensities are particularly useful to assess the local effect of L- to D-inversions. These quantities represent the equilibrium ratio of the number of times in which a

residue is assigned to a β -sheet or turn configuration to the total number of snapshots (observations). Figure 2 shows ensemble-averaged per-residue β -sheet and turn fractions of the studied peptides. As expected, the wild-type GB1 and GB1 d have identical equilibrium per-residue structure propensities. β -sheet fraction reduction toward the termini is expected, as termini are more prone to structural fluctuations.

Comparison of per-residue secondary structure content within the turn group inversions shows how the turn structure is affected by L- to D-inversions. The GB1 t1, t2, and t3 peptides have significantly reduced turn propensities compared to the native GB1. Their β -sheet fractions in the β -sheet region are also decreased. The GB1 t4 peptide, on the other hand, has a turn fraction very close to 1 in the turn region and significantly higher β -sheet fraction as well.

A parallel change in the turn and β -sheet fractions, that is, both increasing or decreasing, is expected as the overall stability of the hairpin is influenced both by the formation of a stable turn structure and by hydrogen bonding and hydrophobic core packing by β -sheet residues. The majority of β -sheet inversions do in fact give rise to parallel changes in turn and β -sheet regions. Some of the inversions in the β -sheet region, however, have a notably different influence on the turn fraction. For example, the GB1 b5, m13, m5, m9, m15, and b33 peptides have turn structures that are shorter by 1 residue on the N-terminus side of the turn (residue 7, aspartic acid: D), for which the turn fraction is nearly zero. Although the turn is shorter by one residue, two of these five peptides (GB1 b5 and m9) are more stable than the native GB1, where stability is assessed by the total β -sheet fraction (Figure 1). In both peptides, residue 7 is switched to β -sheet, in other words, the β -sheet region on the N-terminus side is elongated by 1 residue, indicating that these peptides have different folds than native GB1. The common stereochemical feature among these six peptides is that residue 6 is a D-aspartic acid. Indeed, these six peptides comprise the full list of all peptides where residue 6 is a D-aspartic acid. This is, therefore, a general effect that we identify in this work, namely, that the turn is shortened when a stereochemical inversion is applied at the residue right in front of the turn. This, however, does not imply a correspondingly general effect on stability, as explained above.

Although the changes in β -sheet and turn fractions are unambiguous measures of structural changes in the peptide, they exclusively address the secondary structure content. Peptides can adopt a variety of different configurations with similar β -sheet/turn content. In the rest of this section, we address this structural degeneracy. Figure 3 shows a representative structure of the most populated cluster of the peptides (see Section 2.4). The most populated clusters of the native GB1 and GB1 d peptides share a similar population (occurrence) percentage, their structures being mirror images of each other. A variety of shapes and population percentages result upon backbone stereochemical mutations. One of the most notable is GB1 t4 (Figure 3, top row, fifth from the left). This peptide has the turn mutation at position 10 (K). As mentioned above, this position imposes steric constraints that penalize the L-enantiomer. Accordingly, this mutation results in increased stability. The representative structure of its most populated cluster shows correctly registered β -strands, as in native GB1, however, one of the strands is flipped

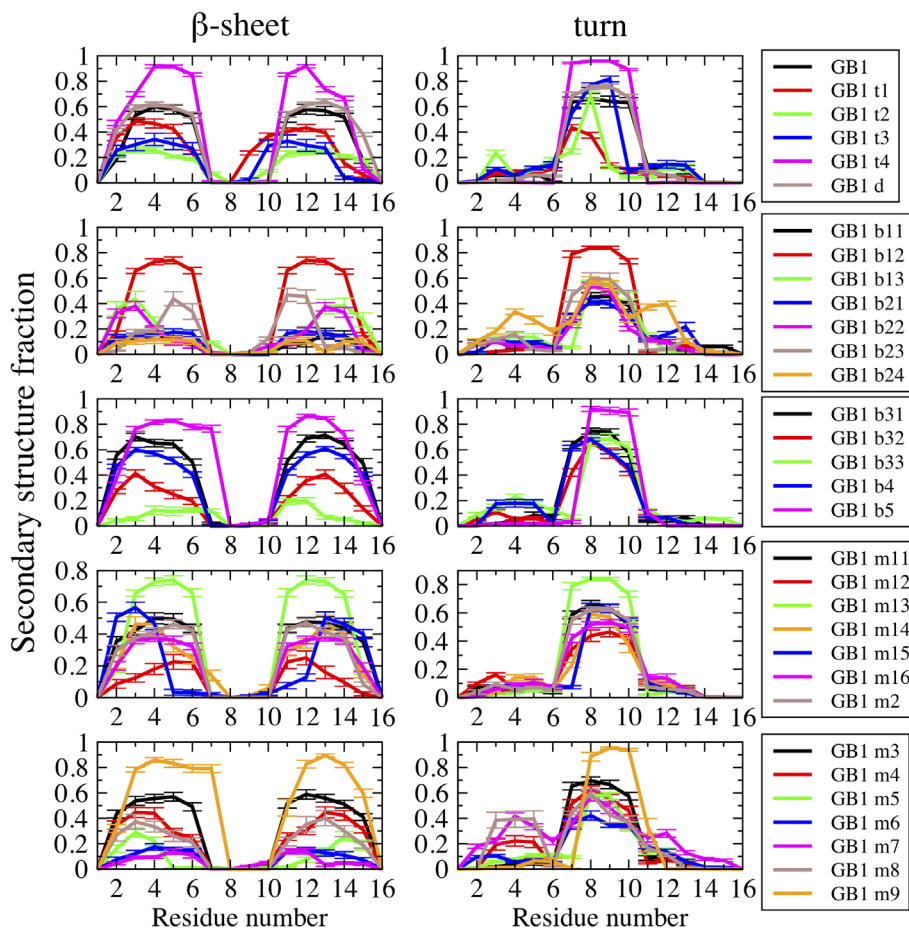


FIGURE 2 Fraction of β -sheet (left) and turn (right) per residue. Errors are calculated from blocked SEs using two equal, nonoverlapping blocks of data ($T = 300$ K) [Color figure can be viewed at wileyonlinelibrary.com]

compared to native GB1, as is evident from the relative orientation of the aromatic side chains. In native GB1 (as well as in GB1 d), the aromatic side chains (three total, one on one strand, two on the other strand) are on the “inside,” bounded by the β -strands (Figure 3), as opposed to the most populated structure of GB1 t4 in which one of the three aromatic side chains faces the “outside,” as its strand is flipped. GB1 t4 is not the only peptide featuring such a flip: similar structures are also highly populated for hairpins such as GB1 m9 and b5. We analyze the backbone torsion angles (ϕ and ψ) of the peptides with β -sheet fraction similar to or higher than that of GB1 and inverse GB1 (Supporting Information Figures S4–S9), which includes GB1 t4, m9, and b5, in order to search for an explanation of this flip. As a reference, we perform the same analysis for native GB1 and GB1 d (Supporting Information Figures S10 and S11, respectively). The Ramachandran map of the GB1 d peptide is rotated by 180° with respect to that of native GB1 peptide, as expected. However, the GB1 t4 peptide has its L-TYR-5 residue adopting an α_R configuration ($-100^\circ \leq -30^\circ$, $-70^\circ \leq \psi \leq 0^\circ$) instead of an extended configuration (for L-amino acids: $-180^\circ \leq \phi$, $\psi \leq -120^\circ$, for D-amino acids: $120^\circ \leq \phi$, $\psi \leq 180^\circ$). This α_R configuration introduces a twist to the backbone, causing the flip of the strand, as shown in Figure 3 for GB1 t4. A similar twist is also present in the GB1 m9 and b5 peptides in the D-ASP-6 residue. In those

cases, the D-ASP-6 residue adopts an α_L configuration ($30^\circ \leq \phi \leq 100^\circ$, $0^\circ \leq \psi \leq 70^\circ$) instead of an extended configuration.

There are other notable structures emerging from stereochemical inversions. With approximately 80% fraction, GB1 b12 populates a highly stable, correctly registered hairpin which is significantly bent compared to native GB1. The shape variability of peptides arising from the joint presence of both L- and D-amino acids has been discussed in the literature before, with the possibility of boat, canoe, or bracelet-like structures addressed.¹⁵ Among the single-pair cross-strand inversion group, only one peptide (GB1 b12) yields a significantly bent hairpin with high stability. Other members of single-pair inversion group destabilize the hairpin. Moreover, other cross-strand pair inversions of two consecutive or two alternating pairs (GB1 b2x peptides) also destabilize the peptide. We do not observe other significantly bent, boat, canoe, or bracelet-like¹⁵ stable hairpin folds by any other stereochemical inversion patterns investigated in this work. While the design of specific peptide folds is beyond the scope of this work, the broad range of L- to D-inversions that we consider here indicates that stability is a crucial concern when applying L- to D-inversions in a hairpin. Even if one can design different patterns of L/D amino acids to achieve a specific shape, the folding equilibrium may shift to destabilize the folded state, in which case a reasonable stability of the desired shape may not be attained.

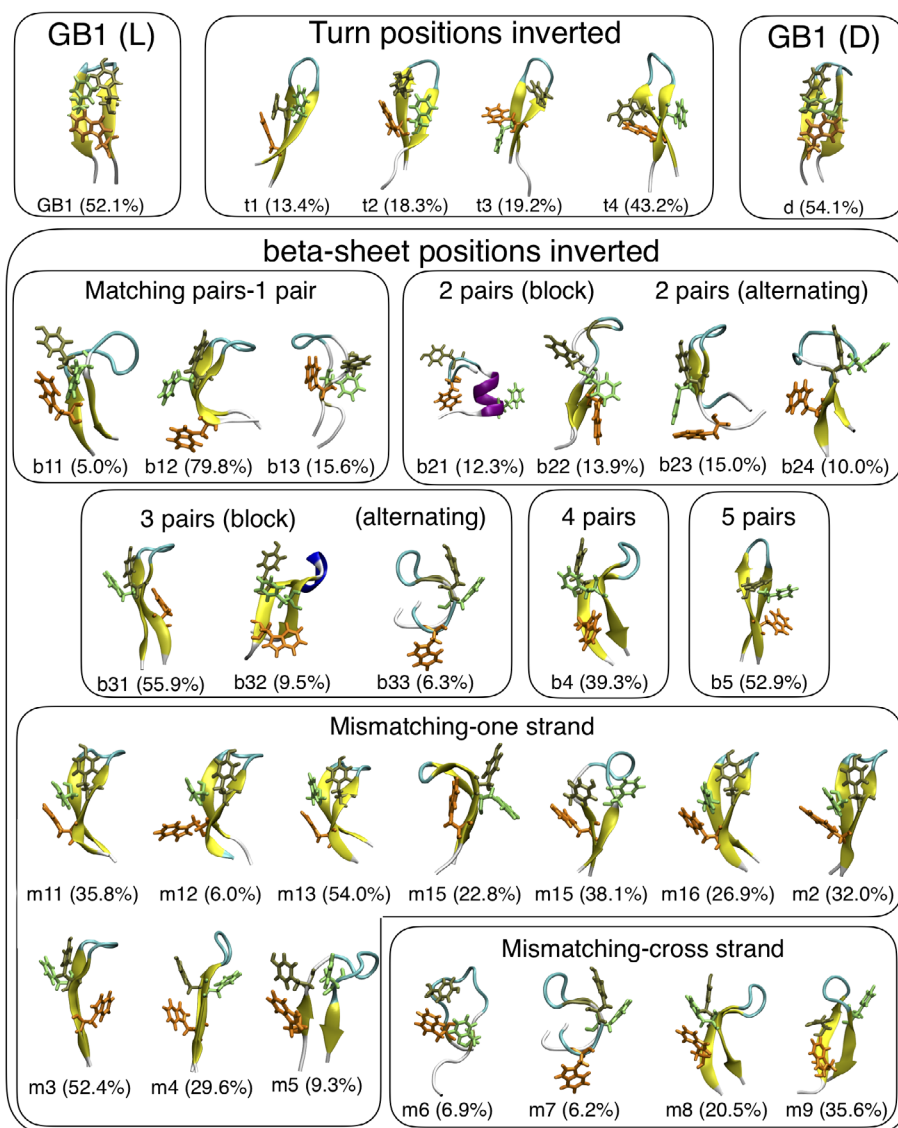


FIGURE 3 Representative structures of the most populated clusters of the peptides. Reported below each structure is the percent occurrence of the most populated cluster. Color coding of the structural elements (backbone) shown in cartoon representation is as follows: yellow: β -sheet, cyan: turn, blue: 3(10)-helix, purple: α -helix, and white: coil. Aromatic side-chains are also shown in licorice representation following the color code of gray: tyrosine, green: phenylalanine, and orange: tryptophan [Color figure can be viewed at wileyonlinelibrary.com]

Structural similarity-based metrics, such as backbone d_{RMS} or fraction of native-like contacts, are commonly used order parameters for thermodynamic analysis of β -hairpins.^{45,72} Here, we provide a free energy analysis using two such order parameters, backbone d_{RMS} and fraction of native-like hydrogen bonds, Q_{hb} (see Section 2.4 for detailed description) as shown in Figure 4. GB1 hairpin configurations with backbone d_{RMS} less than 1.5 Å with respect to the reference structure and Q_{hb} greater than 0.8 are considered folded.⁴⁶ Folded structures similar to native GB1, therefore, populate the top left region of the diagrams in Figure 4. As both backbone d_{RMS} and Q_{hb} rely on internal distances, neither can differentiate the mirror image structures. The same folded basins are, therefore, identified in native GB1 and in the GB1 d peptide. Among the other peptides, GB1 b12, b23, b31, b4, m11, m15, m2, m4, and m8 have similar folded basins too. Some of the other peptides, however, have

another folded basin, which is not one of the low free energy basins seen in native GB1: instead, this basin has a low backbone d_{RMS} (less than 5 Å) and a Q_{hb} of around 0.5. GB1 t4, b5, and m9 peptides, all of which have higher β -sheet fraction than native GB1, have this basin in their free energy surfaces (Figure 4). Its low d_{RMS} indicates that the structures that populate this basin have high backbone similarity to the native fold of GB1 in terms of intramolecular distances, but its lower Q_{hb} indicates that the structures do not form the same backbone hydrogen bonds as in the native fold of GB1, suggesting that the more stable peptides do not necessarily have the same fold. In particular, the folded population of GB1 t4 lacks the termini hydrogen bonds, whereas the GB1 b5 and m9 folded populations lack the hydrogen bonds that the ASP-6 forms with the turn residues in the native structure. We note that the GB1 b5 and m9 peptides have D-ASP at that position. As we show in Figure 3, the turn

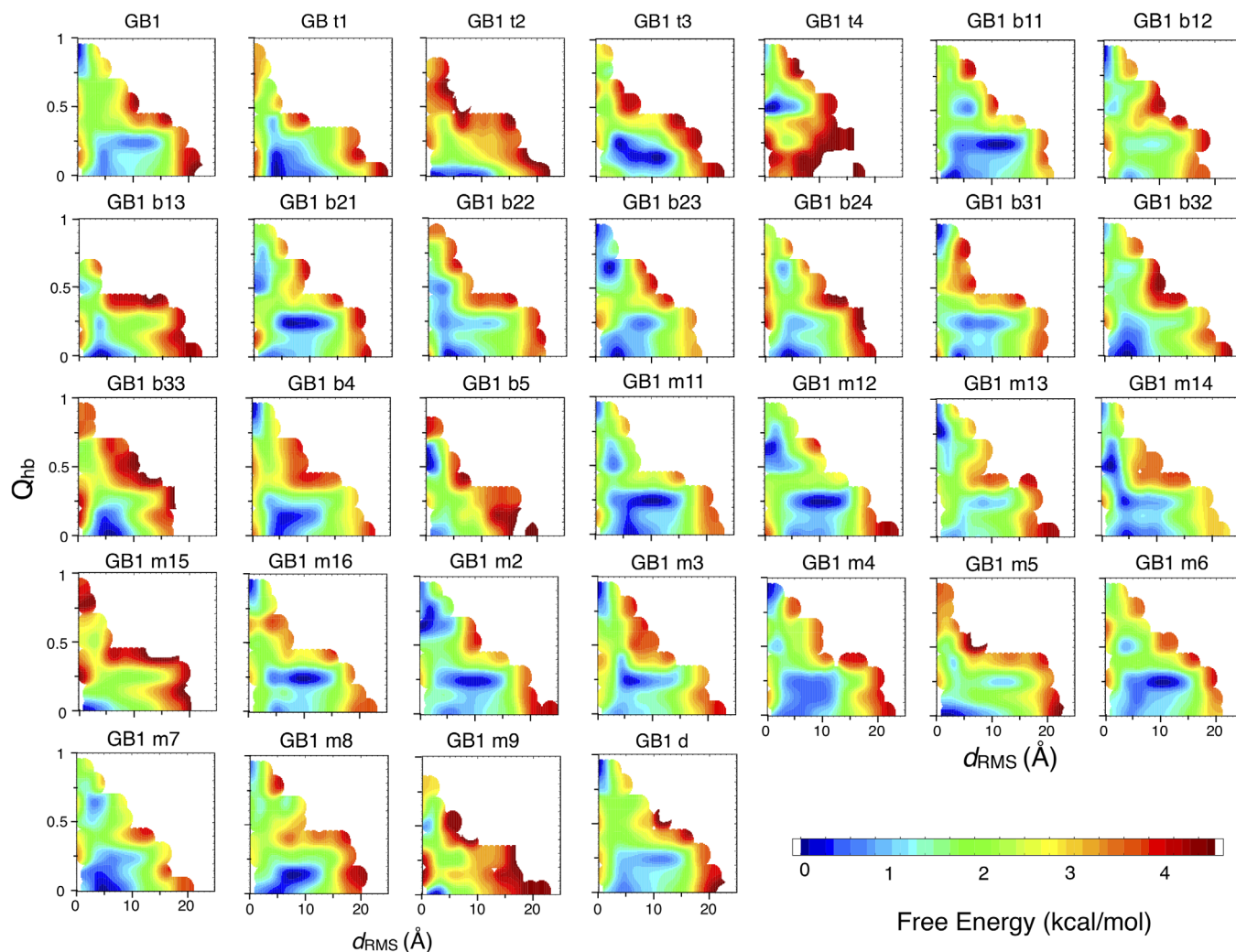


FIGURE 4 Free energy surface of GB1 peptides at 300 K as a function of backbone d_{RMS} and fraction of native hydrogen bonds (Q_{hb}) [Color figure can be viewed at wileyonlinelibrary.com]

structure is shorter by 1 residue, compensated by the elongation of the β -sheet on the N-terminus strand.

4 | CONCLUSIONS

β -hairpins, which are composed of two basic structural elements, β -pleated sheets in an antiparallel arrangement connected with a reverse turn, are ubiquitous protein structural motifs.^{70,73,74} We simulate a number of GB1 hairpin peptides with a variety of stereochemical patterns, classifying the L- to D-inversions into two major groups, according to the position of the inversion(s), as turn or β -sheet inversions (Table 1), and we systematically investigate the effect of stereochemical inversions on folding propensity and stability. While some of the L- to D-inversions cause the β -hairpin structure to deteriorate, whether they belong to the turn or the β -sheet group, the majority preserve the secondary structure content of the peptide. In contrast to the case of a helical peptide,¹ there is no monotonic change in the structure content as a function of number of L- to D-inversions. The effects of L- to D-inversions are instead position-specific for this hairpin peptide. Additionally, as opposed to the recently studied case of a

helical peptide,¹ some inversions yield an increase in the secondary structure content of the peptide. One particularly important example of this increase is the GB1 t4 peptide, which contains a single turn mutation at position 10. Some of the several different turn types of the β -hairpins, such as turn types I', II', and II, require positive ϕ angles at certain turn positions in order to achieve the reversal of the peptide chain direction. Therefore, D-amino acids can be incorporated at these positions, so as to promote turn formation, especially in the design of short peptides and foldamers.^{34,75–77} Although GB1's type I turn does not require a positive ϕ angle in the central turn residues, it does require adoption of an α_{L} configuration at the $i + 3$ position of its 4-residue turn. The GB1 t4 peptide is L- to D-inverted at that location, which promotes an α_{L} configuration. Incorporation of D-amino acids into such positions has been shown to yield an increase in the stability of the peptide as quantified by circular dichroism experiments.^{53,54} Analogously, here we quantify the β -sheet fraction and find that GB1 t4 is also more stable according to this metric. However, the complete three-dimensional structure of the folded conformer shows differences compared to the native structure of the GB1 hairpin as seen qualitatively and quantitatively in Figures 3 and 4, respectively. Therefore, (a) it may not be sufficient to measure only

the secondary structure content in order to claim that the folded state is stabilized. There may be a more stable structure identified, however, that structure may not necessarily be the same expected structure as that of the homochiral native peptide. (b) It is not trivial to introduce a D-inversion even at a location where D-amino acids are more favorable energetically, and concurrently increase stability, as such a mutation may significantly alter the tertiary structure, for example, by flipping a strand, as we observed here. Similar outcomes are also true for the inversions in β -sheet regions, that is, some of the resulting peptides have stable structures, but those structures are different than the native fold of GB1.

ACKNOWLEDGMENTS

GHZ and PGD acknowledge the support from Unilever R&D. The simulations presented in this work are performed on computational resources managed and supported by Princeton Research Computing, a consortium of groups including the Princeton Institute for Computational Science and Engineering (PICSciE) and the Office of Information Technology's High Performance Computing Center and Visualization Laboratory at Princeton University.

ORCID

Gül H. Zerze  <https://orcid.org/0000-0002-3074-3521>

REFERENCES

- Zerze GH, Khan MN, Stillinger FH, DeBenedetti PG. Computational investigation of the effect of backbone chiral inversions on polypeptide structure. *J Phys Chem B*. 2018;122:6354-6363.
- Gevers W, Kleinkauf H, Lipmann F. Peptidyl transfers in gramicidin S biosynthesis from enzyme-bound thioester intermediates. *Proc Natl Acad Sci U S A*. 1969;63:1335-1342.
- Helfman PM, Bada JL. Aspartic acid racemization in tooth enamel from living humans. *Proc Natl Acad Sci U S A*. 1975;72:2891-2894.
- Bada JL. In vivo racemization in mammalian proteins. *Methods Enzymol*. 1984;106:98-115.
- Dunlop DS, Neidle A, McHale D, Dunlop DM, Lajtha A. The presence of free D-aspartic acid in rodents and man. *Biochem Biophys Res Commun*. 1986;141:27-32.
- Kreil G. D-amino acids in animal peptides. *Annu Rev Biochem*. 1997;66:337-345.
- Hashimoto A, Kumashiro S, Nishikawa T, et al. Embryonic development and postnatal changes in free D-aspartate and D-serine in the human prefrontal cortex. *J Neurochem*. 1993;61:348-351.
- Fujii N, Saito T. Homochirality and life. *Chem Rec*. 2004;4:267-278.
- Lam H, Oh D-C, Cava F, et al. D-amino acids govern stationary phase cell wall remodeling in bacteria. *Science*. 2009;325:1552-1555.
- Cava F, Lam H, De Pedro MA, Waldor MK. Emerging knowledge of regulatory roles of D-amino acids in bacteria. *Cell Mol Life Sci*. 2011;68:817-831.
- Fairman R, Anthony-Cahill SJ, DeGrado WF. The helix-forming propensity of D-alanine in a right-handed α -helix. *J Am Chem Soc*. 1992;114:5458-5459.
- Hermans J, Anderson AG, Yun R. Differential helix propensity of small apolar side chains studied by molecular dynamics simulations. *Biochemistry*. 1992;31:5646-5653.
- Krause E, Beyermann M, Dathe M, Rothmund S, Bienert M. Location of an amphipathic α -helix in peptides using reversed-phase HPLC retention behavior of D-amino acid analogs. *Anal Chem*. 1995;67:252-258.
- Krause E, Bienert M, Schmieder P, Wenschuh H. The helix-destabilizing propensity scale of D-amino acids: the influence of side chain steric effects. *J Am Chem Soc*. 2000;122:4865-4870.
- Durani S. Protein design with L- and D- α -amino acid structures as the alphabet. *Acc Chem Res*. 2008;41:1301-1308.
- Gronenborn AM, Filpula DR, Essig NZ, et al. A novel, highly stable fold of the immunoglobulin binding domain of streptococcal protein G. *Science*. 1991;253:657-661.
- Blanco FJ, Rivas G, Serrano L. A short linear peptide that folds into a native stable β -hairpin in aqueous solution. *Nat Struct Mol Biol*. 1994;1:584-590.
- Munoz V, Thompson PA, Hofrichter J, Eaton WA. Folding dynamics and mechanism of β -hairpin formation. *Nature*. 1997;390:196-199.
- Munoz V, Henry ER, Hofrichter J, Eaton WA. A statistical mechanical model for β -hairpin kinetics. *Proc Natl Acad Sci U S A*. 1998;95:5872-5879.
- Dinner AR, Lazaridis T, Karplus M. Understanding β -hairpin formation. *Proc Natl Acad Sci U S A*. 1999;96:9068-9073.
- Pande VS, Rokhsar DS. Molecular dynamics simulations of unfolding and refolding of a β -hairpin fragment of protein G. *Proc Natl Acad Sci U S A*. 1999;96:9062-9067.
- Roccatano D, Amadei A, Nola AD, Berendsen HJ. A molecular dynamics study of the 41-56 β -hairpin from B1 domain of protein G. *Protein Sci*. 1999;8:2130-2143.
- Kobayashi N, Honda S, Yoshii H, Munekata E. Role of side-chains in the cooperative β -hairpin folding of the short C-terminal fragment derived from streptococcal protein G. *Biochemistry*. 2000;39:6564-6571.
- Ma B, Nussinov R. Molecular dynamics simulations of a β -hairpin fragment of protein G: balance between side-chain and backbone forces. *J Mol Biol*. 2000;296:1091-1104.
- Honda S, Kobayashi N, Munekata E. Thermodynamics of a β -hairpin structure: evidence for cooperative formation of folding nucleus. *J Mol Biol*. 2000;295:269-278.
- Klimov D, Thirumalai D. Mechanisms and kinetics of β -hairpin formation. *Proc Natl Acad Sci U S A*. 2000;97:2544-2549.
- Cochran AG, Skelton NJ, Starovasnik MA. Tryptophan zippers: stable, monomeric β -hairpins. *Proc Natl Acad Sci U S A*. 2001;98:5578-5583.
- García AE, Sanbonmatsu KY. Exploring the energy landscape of a β hairpin in explicit solvent. *Proteins*. 2001;42:345-354.
- Zhou R, Berne BJ, Germain R. The free energy landscape for β hairpin folding in explicit water. *Proc Natl Acad Sci U S A*. 2001;98:14931-14936.
- Zagrovic B, Sorin EJ, Pande V. β -hairpin folding simulations in atomistic detail using an implicit solvent model. *J Mol Biol*. 2001;313:151-169.
- Tsai J, Levitt M. Evidence of turn and salt bridge contributions to β -hairpin stability: MD simulations of C-terminal fragment from the B1 domain of protein G. *Biophys Chem*. 2002;101:187-201.
- Bolhuis PG. Transition-path sampling of β -hairpin folding. *Proc Natl Acad Sci U S A*. 2003;100:12129-12134.
- Du D, Zhu Y, Huang C-Y, Gai F. Understanding the key factors that control the rate of β -hairpin folding. *Proc Natl Acad Sci U S A*. 2004;101:15915-15920.
- Fesinmeyer RM, Hudson FM, Andersen NH. Enhanced hairpin stability through loop design: the case of the protein G B1 domain hairpin. *J Am Chem Soc*. 2004;126:7238-7243.
- Olsen KA, Fesinmeyer RM, Stewart JM, Andersen NH. Hairpin folding rates reflect mutations within and remote from the turn region. *Proc Natl Acad Sci U S A*. 2005;102:15483-15487.
- Nguyen PH, Stock G, Mittag E, Hu C-K, Li MS. Free energy landscape and folding mechanism of a β -hairpin in explicit water: a replica exchange molecular dynamics study. *Proteins*. 2005;61:795-808.
- Bolhuis PG. Kinetic pathways of β -hairpin (Un)folding in explicit solvent. *Biophys J*. 2005;88:50-61.
- Muñoz V, Ghirlardo R, Blanco FJ, Jas GS, Hofrichter J, Eaton WA. Folding and aggregation kinetics of a β -hairpin. *Biochemistry*. 2006;45:7023-7035.
- Du D, Tucker MJ, Gai F. Understanding the mechanism of β -hairpin folding via ϕ -value analysis. *Biochemistry*. 2006;45:2668-2678.

40. Bussi G, Gervasio FL, Laio A, Parrinello M. Free-energy landscape for β hairpin folding from combined parallel tempering and metadynamics. *J Am Chem Soc.* 2006;128:13435-13441.
41. Yoda T, Sugita Y, Okamoto Y. Cooperative folding mechanism of a β -hairpin peptide studied by a multicanonical replica-exchange molecular dynamics simulation. *Proteins.* 2007;66:846-859.
42. Juraszek J, Bolhuis PG. Effects of a mutation on the folding mechanism of a β -hairpin. *J Phys Chem B.* 2009;113:16184-16196.
43. Kim J, Keiderling TA. All-atom molecular dynamics simulations of β -hairpins stabilized by a tight turn: pronounced heterogeneous folding pathways. *J Phys Chem B.* 2010;114:8494-8504.
44. Shao Q, Yang L, Gao YQ. Structure change of β -hairpin induced by turn optimization: an enhanced sampling molecular dynamics simulation study. *J Chem Phys.* 2011;135:235104.
45. Best RB, Mittal J. Free-energy landscape of the GB1 hairpin in all-atom explicit solvent simulations with different force fields: similarities and differences. *Proteins.* 2011;79:1318-1328.
46. Best RB, Mittal J. Microscopic events in β -hairpin folding from alternative unfolded ensembles. *Proc Natl Acad Sci U S A.* 2011;108:11087-11092.
47. De Sancho D, Mittal J, Best RB. Folding kinetics and unfolded state dynamics of the GB1 hairpin from molecular simulation. *J Chem Theory Comput.* 2013;9:1743-1753.
48. Markiewicz BN, Yang L, Culik RM, Gao YQ, Gai F. How quickly can a β -hairpin fold from its transition state? *J Phys Chem B.* 2014;118:3317-3325.
49. Kimura S, Kanaya S, Nakamura H. Thermostabilization of *Escherichia coli* ribonuclease HI by replacing left-handed helical Lys95 with Gly or Asn. *J Biol Chem.* 1992;267:22014-22017.
50. Stites WE, Meeker AK, Shortle D. Evidence for strained interactions between side-chains and the polypeptide backbone. *J Mol Biol.* 1994; 235:27-32.
51. Takano K, Yamagata Y, Yutani K. Role of non-glycine residues in left-handed helical conformation for the conformational stability of human lysozyme. *Proteins.* 2001;44:233-243.
52. Zerze GH, Uz B, Mittal J. Folding thermodynamics of β -hairpins studied by replica-exchange molecular dynamics simulations. *Proteins.* 2015;83:1307-1315.
53. Anil B, Song B, Tang Y, Raleigh DP. Exploiting the right side of the Ramachandran plot: substitution of glycines by D-alanine can significantly increase protein stability. *J Am Chem Soc.* 2004;126:13194-13195.
54. Hua Q-X, Nakagawa S, Hu S-Q, Jia W, Wang S, Weiss MA. Toward the active conformation of insulin stereospecific modulation of a structural switch in the B chain. *J Biol Chem.* 2006;281:24900-24909.
55. Brooks BR, Brooks CL III, Mackerell AD Jr, et al. CHARMM: the biomolecular simulation program. *J Comput Chem.* 2009;30:1545-1614.
56. Gfeller D, Michielin O, Zoete V. SwissSidechain: a molecular and structural database of non-natural sidechains. *Nucleic Acids Res.* 2012;41: D327-D332.
57. Best R, Hummer G. Optimized molecular dynamics force fields applied to the helix-coil transition of polypeptides. *J Phys Chem B.* 2009;113: 9004-9015.
58. Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. *J Chem Phys.* 1983;79:926-935.
59. Berendsen HJ, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. *Comput Phys Commun.* 1995;91:43-56.
60. Hess B, Kutzner C, Van Der Spoel D, Lindahl E. GROMACS 4: algorithms for highly efficient, load-balanced, and scalable molecular simulation. *J Chem Theory Comput.* 2008;4:435-447.
61. Bonomi M, Branduardi D, Bussi G, et al. PLUMED: a portable plugin for free-energy calculations with molecular dynamics. *Comput Phys Commun.* 2009;180:1961-1972.
62. Sugita Y, Okamoto Y. Replica-exchange molecular dynamics method for protein folding. *Chem Phys Lett.* 1999;314:141-151.
63. Bonomi M, Parrinello M. Enhanced sampling in the well-tempered ensemble. *Phys Rev Lett.* 2010;104:190601.
64. Deighan M, Bonomi M, Pfendtner J. Efficient simulation of explicitly solvated proteins in the well-tempered ensemble. *J Chem Theory Comput.* 2012;8:2189-2192.
65. Nosé S. A molecular dynamics method for simulations in the canonical ensemble. *Mol Phys.* 1984;52:255-268.
66. Hoover WG. Canonical dynamics: equilibrium phase-space distributions. *Phys Rev A.* 1985;31:1695-1697.
67. Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG. A smooth particle mesh Ewald method. *J Chem Phys.* 1995;103:8577-8593.
68. Kabsch W, Sander C. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Bio-polymers.* 1983;22:2577-2637.
69. Daura X, Gademann K, Jaun B, Seebach D, van Gunsteren WF, Mark AE. Peptide folding: when simulation meets experiment. *Angew Chem Int Ed.* 1999;38:236-240.
70. Sibanda B, Thornton J. β -hairpin families in globular proteins. *Nature.* 1985;316:170-174.
71. Best RB, Mittal J. Balance between α and β structures in ab initio protein folding. *J Phys Chem B.* 2010;114:8790-8798.
72. Hatch HW, Stillinger FH, DeBenedetti PG. Computational study of the stability of the miniprotein trp-cage, the GB1 β -hairpin, and the AK16 peptide, under negative pressure. *J Phys Chem B.* 2014;118:7761-7769.
73. Gunasekaran K, Ramakrishnan C, Balaram P. Beta-hairpins in proteins revisited: lessons for de novo design. *Protein Eng.* 1997;10:1131-1141.
74. Kuhn M, Meiler J, Baker D. Strand-loop-strand motifs: prediction of hairpins and diverging turns in proteins. *Proteins.* 2004;54:282-288.
75. Rose GD, Gierasch LM, Smith JA. Turns in peptides and proteins. *Adv Protein Chem.* 1985;37:1-109.
76. Imperiali B, Moats R, Fisher S, Prins T. A conformational study of peptides with the general structure Ac-L-Xaa-Pro-D-Xaa-L-Xaa-NH₂: spectroscopic evidence for a peptide with significant β -turn character in water and in dimethyl sulfoxide. *J Am Chem Soc.* 1992;114: 3182-3188.
77. Espinosa JF, Syud FA, Gellman SH. Analysis of the factors that stabilize a designed two-stranded antiparallel β -sheet. *Protein Sci.* 2002;11: 1492-1505.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Zerze GH, Stillinger FH, DeBenedetti PG. Effect of heterochiral inversions on the structure of a β -hairpin peptide. *Proteins.* 2019;87:569-578. <https://doi.org/10.1002/prot.25680>