JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

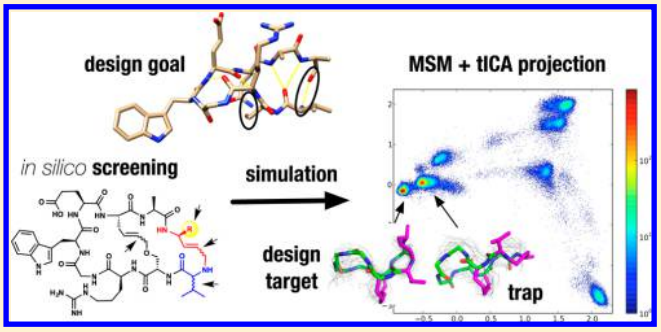
Computational Screening and Selection of Cyclic Peptide Hairpin Mimetics by Molecular Simulation and Kinetic Network Models

Asghar M. Razavi, William M. Wuest,* and Vincent A. Voelz*

Department of Chemistry, Temple University, Philadelphia, Pennsylvania 19122, United States

Supporting Information

ABSTRACT: Designing peptidomimetic compounds to have a preorganized structure in solution is highly nontrivial. To show how simulation-based approaches can help speed this process, we performed an extensive simulation study of designed cyclic peptide mimics of a β-hairpin from bacterial protein LapD involved in a protein−protein interaction (PPI) pertinent to bacterial biofilm formation. We used replica exchange molecular dynamics (REMD) simulation to screen 20 covalently cross-linked designs with varying stereochemistry and selected the most favorable of these for massively parallel simulation on Folding@home in explicit solvent. Markov state models (MSMs) built from the trajectory data reveal how subtle chemical modifications can have a significant effect on conformational populations, leading to the overall stabilization of the target structure. In particular, we identify a key steric interaction between a methyl substituent and a valine side chain that acts to allosterically shift population between native and near-native states, which could be exploited in future designs. Visualization of this mechanism is aided considerably by the tICA method, which identifies degrees of freedom most important in slow conformational transitions. The combination of quantitative detail and human comprehension provided by MSMs suggests such approaches will be increasingly useful for design.



INTRODUCTION

Peptidomimetics, molecules that mimic the folded structure of specific protein epitopes, are an important platform for targeting protein−protein interactions (PPIs). Examples of such molecules to date include stapled peptides and rationally designed scaffolds able to target previously "undruggable" sites.[1] While a number of α-helix mimetics have been developed,[2−4] the number of β-hairpin peptidomimetics remains small, despite their importance as molecular recognition motifs in many PPIs.[5]

A promising class of β-hairpin mimics are cyclic peptides that have a strong propensity to fold into a target structure similar to the bound state. For these molecules, cyclization and cross-linking can provide preorganization that will enhance binding affinity.[5−9] It remains challenging, however, to predict the nontrivial structural and dynamic consequences of these constraints due to the foldameric nature of cyclic peptides.[10] Therefore, it is imperative to develop improved computational approaches for screening and selecting favorable peptidomimetic designs based on their conformational properties.[11]

Herein, we focus specifically on the unbound state of cyclic peptide ligands. We ask: Can molecular simulation methods help guide the design of β-hairpin mimics that are highly preorganized in solution toward a native-like conformation? Successful simulation methods should be able to predict which chemical cross-links will achieve a solution-state conformational ensemble most "aligned" to a desired target structure. Furthermore, such methods should be able to accurately estimate the populations of competing conformational states and leverage this information to guide the design of chemical substitutions that will either energetically reward the target conformation or penalize off-pathway conformational states.

As a proof-of-principle demonstrating progress toward these goals, we present a molecular simulation study of several putative β-hairpin mimetics inspired by the bacterial PPI between LapD and LapG. A key step in *Pseudomonas* bacterial biofilm formation involves a PPI in which a β-hairpin of LapD modulates the activity of protease LapG.[12,13] We explored a series of seven-residue (7mer; Figure S1) and nine-residue (9mer; Figure 1) covalently cross-linked cyclic peptides to mimic the LapD hairpin from *P. fluorescens*, which in practice could be synthesized by olefin metathesis strategies[8,14,15] (Figure 1). Our goal was to design compounds with solution-state structures predicted to be similar to the native hairpin crystal structure of *P. fluorescens* LapD (3PJV). Although a cocrystal structure of LapD bound to LapG has not yet been published, the pattern of conserved residues in both LapD and LapG, as well as the crystal structure of LapG (from related species *L. pneumophila*), suggest that the LapD hairpin binds in a similar conformation.[16]

Received: February 17, 2014
Published: April 23, 2014

dx.doi.org/10.1021/ci500102y | J. Chem. Inf. Model. 2014, 54, 1425−1432

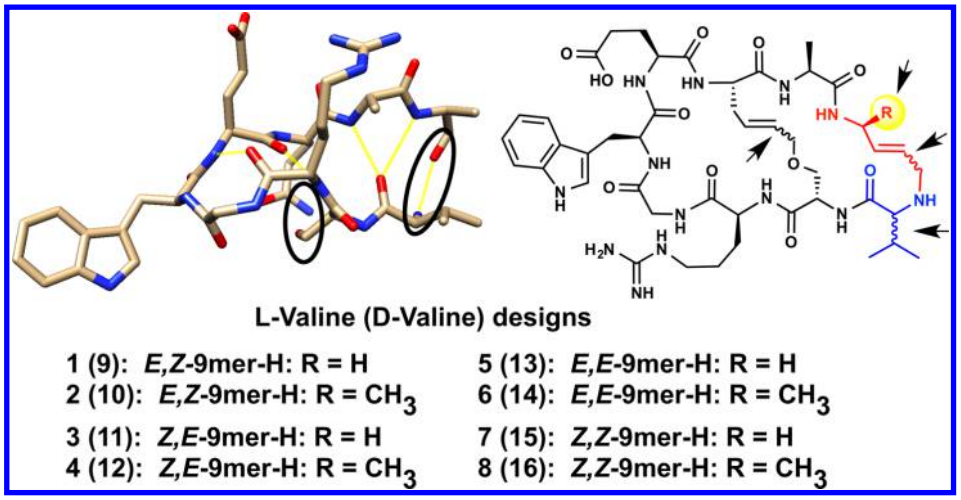


Figure 1. Chemical structure of LapD β-hairpin and cyclic peptide 9mer designs inspired by the *P. fluorescens* hairpin sequence VSRGWEQAA. The *E*/*Z* naming scheme describes the olefin geometry of the central and peripheral linkers, respectively. The valine position is shown in blue. Arrows indicate the relevant moieties varied in these designs (linker stereochemistry, D- vs L-valine, and hydrogen vs methyl at the R group). Structures of 7mer designs **17**−**20** are shown in Figure S1.

MATERIALS AND METHODS

The cross-linked hairpin designs were motivated by envisioning covalent tethers replacing the intramolecular hydrogen bonds seen in the crystal structure of LapD (Figure 1, circled). 7mer and 9mer designs were explored because they were long enough to preserve backbone hydrogen bonds defining the hairpin. Residues S2 and Q7 were chosen for the inner cross-link, preserving the potentially advantageous salt-bridging properties of R3 and E6.

To explore the conformational properties of these putative designs, we employed a two-step computational protocol, utilizing simulations performed with the GROMACS simulation package.[17] First, in an initial screening phase, we performed replica-exchange molecular dynamics (REMD) simulations in implicit solvent to assess the relative stability of each design. Next, we selected the most promising designs for large-scale parallel simulation in explicit solvent on the Folding@home distributed computing platform. The resulting trajectory data, which totaled over a millisecond of simulation time, was used to construct Markov state models (MSMs) of conformational dynamics to identify the relevant conformational states and the transition rates between each state, providing a complete picture of thermodynamics and kinetics of the designs.

Force Fields. Initial screening simulations were performed using the AMBER ff96 potential with generalized Born surface area (GBSA) solvation, while simulations using AMBER ff99sb-ildn + TIP3P explicit solvent were used at a later stage for the most promising designs. The use of AMBER ff96 (which is biased toward beta-sheet secondary structure) is to correct for a well-known artifact of GBSA that biases the protein backbone toward α-helical secondary structure.[18] In 2008, Shell et al. tested the accuracy of several force field + GBSA combinations available in AMBER and found that the optimal combination for predicting known peptide structures was AMBER ff96[19] + the GBSA model of Onufriev and co-workers.[20] This combination has since been validated in large-scale simulation studies of the 39-residue mixed α/β protein NTL9, the first published example of *ab initio* folding simulated for a millisecond folder,[21] and others.[22,23] For the explicit simulations performed on Folding@home, we chose AMBER ff99SB-ildn, which has been shown to be very accurate when combined with TIP3P and TIP4P-Ew solvent models in long-time-scale simulations.[24]

Screening of Cyclic Hairpin Designs. For initial screening of designs (Figure 1), and linear and cyclic (cysteine disulfide-bonded) peptide controls, the AMBER ff96 potential was used with the OBC GBSA implicit solvation model.[25] The GAFF force field[26] was used for non-natural amino acid residues and cross-linkers, with partial charges computed using the AM1-BCC[27] method, as implemented in the AmberTools software package.[28] Replica-exchange molecular dynamics (REMD) was performed with 24 replicas and temperatures exponentially spaced from 300 to 1000 K to ensure broad conformational sampling. The total simulation time of each replica ranged from 1 to 2 μs, with snapshots recorded every 1 ps. Only configurations having all *trans* backbone amides were used in the analysis (as isomerization can occur at high temperatures). Exchanges were attempted every 10 ps between neighboring temperature replicas, with exchange probabilities ranging from 0.82−0.95, enabling extensive traversal of all temperatures. For peptide control simulations, 12 replicas with temperatures 300−450 K were used. Exchanges were attempted every 10 ps, with exchange probabilities ranging from 0.65−0.80. Stochastic (Langevin) dynamics was performed using a time step of 2 fs, with a water-like viscosity of 91/ps, hydrogen bonds restrained using LINCS, and nonbonded cutoffs of 9 Å. Free energy profiles versus rmsd for the initial screening simulations were computed using the MBAR estimator of Shirts and Chodera,[29] from the 12 lowest-temperature replicas, with trajectory data subsampled every 5 ps.

The best candidate designs were selected by computing free energy profiles as a function of rmsd to the native LapD hairpin structure. The rmsd was computed with respect to the backbone and C_β atoms of the six residues in the hairpin turn (SRGWEQ). Using rmsd as our selection metric is a simplification, as we have no information about possible structure−activity relationships, and thus our guiding design principle is simply that the binding mode of the hairpin must be similar in the PPI. We note that in the results presented below, the target rmsd's of our best designs are small enough (0.5−1.0

Å) that rmsd should be a very good proxy for overall structural similarity.

Coarse-Grained Clustering of Screening Simulation Data. To better evaluate the conformational landscapes for designs **9** and **10**, the MSMBuilder2 software package[30] was used to cluster the trajectory data into 10 coarse-grained states. Hybrid *k*-medoid conformational clustering was applied to a combined trajectory data set of the lowest-temperature replicas (T = 300 K) for designs **9** and **10**, using a distance metric of backbone + C_β rmsd. For each design, conformational free energies for each of the 10 states were crudely estimated by assigning the remaining temperature replicas to cluster generators with weight $\exp(-\langle U\rangle/k_BT)$, where $\langle U\rangle$ is the average potential energy of each temperature replica.

Parallel Simulation on Folding@home. Massively parallel molecular dynamics simulations of designs **1**, **2**, **9** and **10** were performed on the Folding@home distributed computing platform[31] at both 300 and 350 K, for over a millisecond of total simulation time. The AMBER ff99sb-ildn + GAFF potential was used in combination with the TIP3P explicit solvent model. NVT molecular dynamics were performed at 300 and 350 K using a Berendsen thermostat, for a cubic periodic box of volume $(31.72\ Å)^3$ filled with solute, two Na^+ ions, two Cl^- ions, and 1000 water molecules. Hydrogen bonds were restrained using LINCS, and PME electrostatics was used with nonbonded cutoffs of 9 Å. For each system, distributed clients generated upward of 500 parallel trajectories, each with a minimum length of 10 ns. Because of the stochastic nature of distributed computing, a roughly exponential distribution of trajectory lengths was obtained, having a mean trajectory length of ∼500 ns, with the longest trajectory reaching 3−4 μs. Snapshots were recorded every 100 ps.

Markov State Model Construction. Markov state models (MSMs) were constructed using the MSMBuilder2 software package.[30] Conformational clustering was performed to define microstates, using time-structure-based independent component analysis (tICA).[32] The tICA method calculates (here, from the time course of 861 backbone + C_β interatomic distances) a time-lagged correlation matrix, from which can be computed a subspace that best captures the slowest conformational motions. It has been shown that conformational clustering in this reduced-dimension subspace results in much improved metastable state definitions and implied time scales that are very robust with respect to the lag time used.[32,33] Because of the similarity in the different simulation systems, we were able to combine the simulation data for designs **1** and **2** and combine the simulation data for designs **9** and **10**. To validate this approach, we performed several studies comparing the results of clustering and constructing MSMs using separate versus combined and found similar results (see Figures S5 and S6).

The time-lagged covariance matrix was constructed using a 2 ns time lag. This, along with a corresponding covariance matrix, was used to compute the tICA eigenvalues and eigenvectors.[32] The trajectory data were projected to a subspace defined by the two largest tICA eigenvectors, and conformational clustering was performed using a hybrid *k*-medoid algorithm to obtain a MSM with 2000 microstates. The average distance of any conformation to its cluster generator was 0.0251 units (in the tICA subspace). We found that 2000 microstates was a number large enough to provide a great deal of microscopic resolution, yet not too large as to suffer from finite sampling error in state transitions. The maximum likelihood estimator (MLE) of Beauchamp et al. was used to estimate a microstate transition matrix from the observed transition counts obeying detailed balance.[30] A lag time of 2 ns was chosen for building MSMs for each data set. The BACE algorithm[34] was used to lump 2000 microstates into 20 macrostates. When projected to the two largest tICA components, the macrostate definitions thus obtained align very well with the metastable basins seen in the raw trajectory data (see Figures S8 and S9).

RESULTS

Initial Screening of Designed Cyclic Peptides Show Improved Conformational Preferences Compared to Peptide Controls. As controls, we additionally simulated the linear peptide sequence Ac-SRGWEQA-Nme, the cyclic variant CSRGWEQC with disulfide-bonded cysteines, and corresponding linear (Ac-MDGWMQA-Nme) and cyclic (CMDGWMQC) sequences from *L. pneumophila*. Simulations of the control sequences reveal that the MDGWMQA sequence from *L. pneumophila* has a stronger propensity to form a native-like hairpin structure in solution, compared to the SRGWEQA sequence from *P. fluorescens*, which has a strong helical propensity. In both cases, the simulations predict that cyclization due to cysteine disulfide bonding serves to destabilize the native hairpin structure with respect to other compact conformational states (Figures S2a,b).

Simulations of our cyclic hairpin designs show improved conformational preferences. While our cyclized 7mer designs **17**−**20** are comparable to the control sequences (Figures S2c,d), the 9mer designs **1**−**16** display a near-native conformational basin near 1.5 Å rmsd and another at <1 Å, which we deem to be the "native-like" conformational basin (Figure S3). It is not a coincidence that such states exist in all the designs, as the cross-linkers force the hairpin into only a small number of competent hydrogen bonded hairpin configurations near the native state.

For both L-Val and D-Val versions, designs with a central cross-linker having the (*E*)-form olefin geometry are more native-like (∼0.7 Å rmsd) than those of the (*Z*)-form (∼0.9 Å rmsd), while the (Z)-form is generally preferred for the peripheral linker. We ascribe this difference to the geometry of the two linker forms. The C−C═C−C distance in the (*E*) form is ∼3.9 Å, but in the (*Z*) form it is ∼3.1 Å.

Overall, the *E*-*Z*-9mer designs emerge from our screening simulations as the most favorable, with design **1** (L-Val *E*-*Z*-9mer-H) ranking the best, having a native-like basin comparable in free energy to the near-native basin. Further analysis of the *E*-*Z*-9mer free energy profiles reveals the interesting observation that design **2** (L-Val *E*-*Z*-9mer-Me), while differing from **1** only by a methyl group, has a destabilized native-like basin (by ∼0.5 kcal/mol). The reverse situation is found for designs **9** (D-Val *E*-*Z*-9mer-H) and **10** (D-Val *E*-*Z*-9mer-Me); in this case, methylation is predicted to *stabilize* the native-like basin by ∼1 kcal/mol (Figure 2).

Coarse-grained Conformational Clustering Suggests Sequence-Dependent Population Shifts in Metastable States. To gain more insight into the role of methylation in controlling the conformational preferences of *E*-*Z*-9mer designs, we performed a coarse-grained conformational clustering of the combined REMD data (see Methods in SI) into 10 states based on backbone dihedral angles. An analysis of state population differences suggests that steric mechanisms

dx.doi.org/10.1021/ci500102y | J. Chem. Inf. Model. 2014, 54, 1425−1432

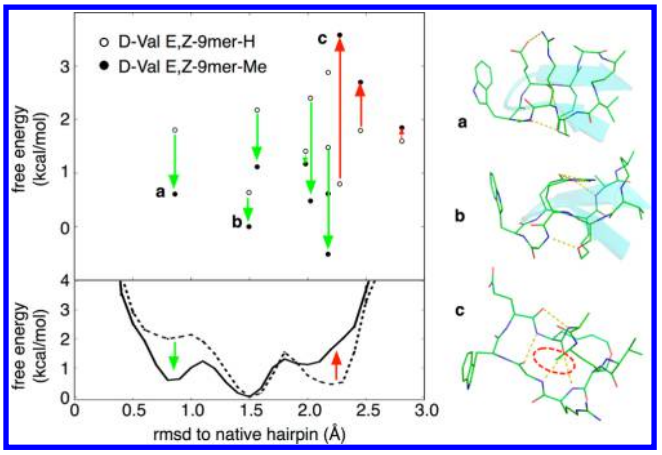


Figure 2. Coarse-grained conformational clustering reveals population shifts upon methylation. (Bottom left): Free energy profiles for designs **9** (D-Val *E-Z*-9mer-H, dashed line) and **10** (D-Val *E-Z*-9mer-Me, solid line). Arrows show the population shifts that methylation induces. (Top left): Free energies of conformational macrostates (computed from populations p_i as $-kT \log p_i$, T = 300 K) versus rmsd for design **9** (filled circles) and design **10** (open circles). Arrows show the population shifts that methylation induces. (Right): Structural representations of macrostates a, b, and c that are labeled on the top left panel. The red dashed circle denotes a steric clash with the methyl group.

might explain the shift in state populations upon methylation (Figure 2).

Markov State Models of Cyclic Hairpin Dynamics Reveal a Steric Mechanism for the Sequence-Dependent Stability of Methylated Designs. We then simulated designs **1**, **2**, **9**, and **10** in explicit water (TIP3P) at 300 K on the massively parallel Folding@home distributed computing platform, using the 10 cluster generators as starting conformations. The MSMBuilder2 software package was used to construct MSMs from the resulting trajectory data. Considerable assistance in computing the relevant state definitions and visualization was provided by a new tICA approach,[32,33] in which a time-lagged correlation matrix is diagonalized to generate a projection of atomic distance coordinates onto a low-dimensional subspace corresponding to the slowest kinetic degrees of freedom. A range of tICA subspaces was explored to find the optimal subspace in which to perform clustering to discover metastable states. The two largest tICA components (derived from all pairwise backbone and C_β atomic distances) were used to cluster all simulation data into 2000 microstates, which we found to be sufficient to identify the relevant degrees of freedom. To better understand conformational similarities or differences, MSMs were built by combining simulation data for designs **1** and **2** and designs **9** and **10**. We found that the L-Val designs (**1** and **2**) and D-Val designs (**9** and **10**) needed to be clustered separately, as the conformational states of valine play a predominant role in the slowest conformational transitions (Figure S4), thus creating poor overlap in the tICA subspace. For methylated and nonmethylated designs, however, the state definitions overlap very well in the tICA subspace. MSMs constructed from separate versus combined trajectory data predict similar spectra of relaxation times (Figure S5), and the implied time scales of microstate vs macrostate MSMs are similar[30,35] (Figure S6).

Choosing a lag time of τ = 2 ns (long enough for the implied time scales to plateau), we find that the slowest relaxation time scale for all designs are ~3 μs, on the order of previous measurements of folding times for β-hairpins.[36] For the L-Val designs, the first and second tICA components (i.e., the reaction coordinates along which the slowest structural rearrangements occur) capture the inversion of the central (*E*)-form linker across the plane of the hairpin, as well as motions of the peripheral (*Z*)-form linker (Figure S7). A 20-macrostate decomposition performed using the BACE

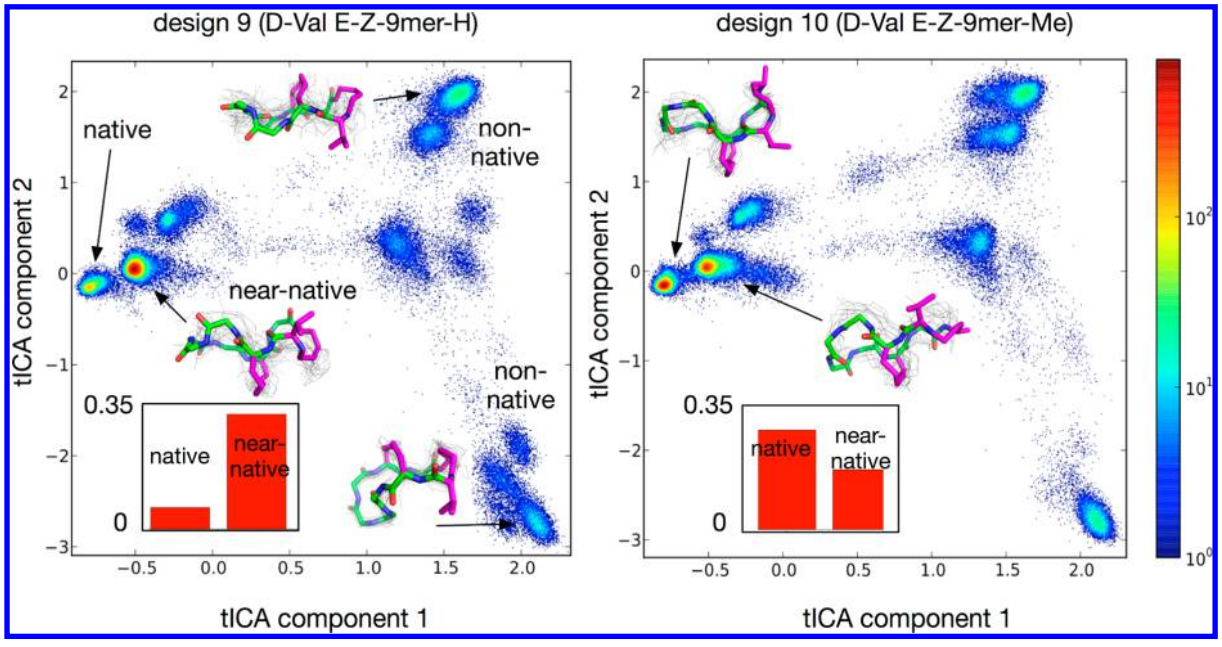


Figure 3. Projections of explicit-solvent simulation data for designs **9** (D-Val *E-Z*-9mer-H) and **10** (D-Val *E-Z*-9mer-Me) onto the two largest tICA components. These reflect the slowest conformational transitions, which involve motions of the peripheral linker, and the inversion of the central linker across the plane of the hairpin. Macrostate conformations are shown for selected basins, with central and peripheral linkers shown in magenta. Addition of a methyl group shifts macrostate populations toward the native state (bar graphs).

dx.doi.org/10.1021/ci500102y | J. Chem. Inf. Model. 2014, 54, 1425−1432

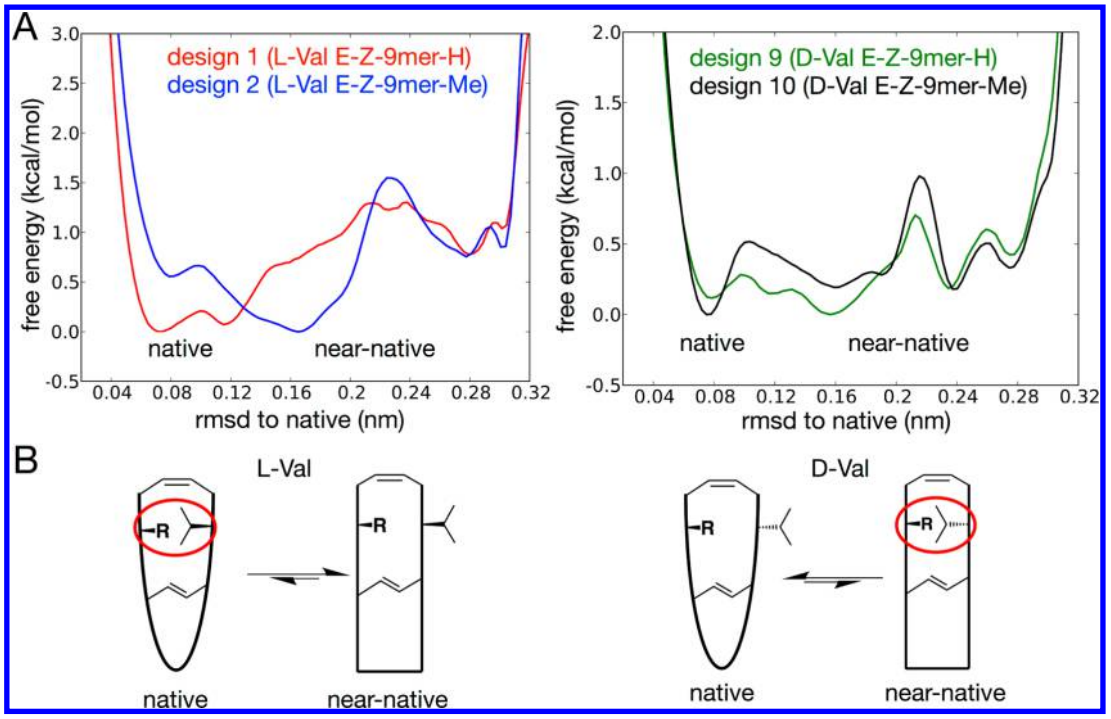


Figure 4. (A) Free energy profiles from explicit-solvent simulations performed on Folding@home, consistent with implicit-solvent REMD. (B) A schematic mechanism for steric-induced stabilization of the hairpin native fold, R = H or CH_3. The wedge/dash indicates L- or D-Val, with steric clash indicated by an inward-facing side chain. Circled in red is the steric clash occurring when R = CH_3, resulting in a population shift (arrows).

algorithm[34] shows that inverted non-native forms indeed have large populations (Figures S8−12). The discovery of these conformations was unexpected, which further underscores the need for accurate simulation-based tools for rational design, as consequences of conformational constraints may be highly nontrivial.

A projection of the simulation data onto the two largest tICA components is especially useful in helping to visualize competing metastable states. In this view, the native-like state and near-native states first seen in the screening simulations are clearly differentiated into distinct metastable basins. For the D-Val designs **9** and **10**, the methylated version displays a population shift toward the more native-like basin, confirming the results of our implicit-solvent screening simulations (Figure 3, Figure S9). The reverse effect is seen for the L-Val designs (Figures S7 and S8).

A detailed inspection of the MSM macrostate decomposition reveals a specific allosteric mechanism for the methylation-dependent population shifts, namely that the presence of a methyl group creates a steric clash with the valine side chain. For the L-Val designs, the native backbone configuration has the valine side chain above the hairpin plane; the additional methyl group destabilizes this configuration (Figure 4, Figure S10). The opposite is true for the D-Val designs; in that case, the additional methyl destabilizes non-native backbone configurations having the valine side chain above the plane (Figure 4, Figure S11).

DISCUSSION AND CONCLUSION

This proof-of-principle study highlights important implications for the computational design of peptidomimetics. To date, the most successful computational approaches for peptide and peptidomimetic design have incorporated knowledge of bound-state structures and physics-based information into Monte Carlo searches across conformational and sequence spaces. Prominent examples include many functional beta-peptides from the DeGrado group designed using the NAPOLI algorithm,[37−39] MDM2-binding spiroligomers designed using the CANDO algorithm,[3] as well as several de novo PPIs designed using the Rosetta algorithm,[40−42] which has recently been adapted to work with non-natural amino acids and other peptidomimetic polymers such as peptoids.[43,44] Other approaches have relied on more expensive physics-based methods such as free energy perturbation to refine initial leads.[11,45] It is interesting to note that for many of these successes (as well as successful noncomputational foldamer designs[46,47]), a great deal of preorganization already exists in the starting material, reducing the role of conformational heterogeneity and/or misfolding.

One lesson from our work here simulating cross-linked cyclic peptides is that the problem of rationally designing conformational constraints can be highly nontrivial. This underscores the need for transferrable, atomically detailed information about all relevant metastable states, including misfolded states.[9] Molecular simulation-based approaches like the kind presented here are well suited to provide this information. An advantage of the kind of large-scale distributed computing performed here is that we can obtain very thorough conformational sampling in explicit solvent, to achieve a high level of molecular detail. To our knowledge, this work represents the most in-depth simulation study of proposed cyclic peptide hairpins to date (over a millisecond of total trajectory data). A disadvantage of such massive conformational sampling efforts is the potentially prohibitive computational expense. Furthermore, such large-scale simulation efforts may not be suitable for studying larger proteins and/or constructs. That said, the good agreement of our explicit-solvent results with the earlier implicit-solvent screening simulations suggests that less expensive simulation

methods may be accurate enough for future studies, depending on the type of system and the time scales of conformational motions. New Markov state model (MSM) methods may also help to make our approach more amenable to studying larger systems, as these methods can elucidate long-time scale conformational dynamics using ensembles of short-time scale trajectories.

We have shown that MSM models, along with the tICA method, provide a powerful new way to make sense of complicated conformational landscapes that may otherwise challenge our human comprehension. In particular, it is the power and versatility of the tICA approach that makes this possible. Whereas projections onto reaction coordinates such as radius of gyration and rmsd are highly diffuse, tICA projections are well-defined and reveal clearly the relevant metastable states and their role in conformational dynamics. Armed with the tICA approach, an MSM description of a peptidomimetic scaffold thus can be extremely useful: one could use the knowledge of competing metastable states to propose further chemical modifications anticipated to further shift state populations. Alternatively, MSM adaptive sampling methods[48,49] could be used in an automated fashion to efficiently discern the effects of many proposed chemical substitutions.

Finally, this study suggests that preorganization can be enhanced by subtle chemical modifications designed to induce allosteric population shifts,[50] much in the same way that post-translational modifications or binding interactions can transmit far-reaching conformational changes. In the vast literature of previous work on cyclic peptides and hairpin mimics,[5,10,51,52] we note that several observations are consistent with the predicted allosteric effects of methylation. Examples include so-called "magic methylation" effects—in which a simple methylation imparts profound improvements on the affinity and biological activity of small-molecule ligands—which are thought to be a result of steric interactions that enhance particular conformational biases.[53] *N*-methylation of cyclic peptides has been shown to induce preorganized structure in solution,[54,55] as well as influence conformational-dependent permeability[56] and enhance binding affinity.[57] Reminiscent of the findings presented here, a previous study found that disulfide-bonded cyclic hairpin conformations can be "fine-tuned" by subtle differences in side chains and that significant stabilization can be achieved by the substitution of a valine residue adjacent to the cysteine cross-link.[58]

We stress that the computationally observed allosteric effect of methylation is strictly a prediction, and one that needs to be fully validated by experimentation. We are continuing to study further cyclic designs mimicking the LapD hairpin in conjunction with experimental studies to confirm these predictions, to be described in future publications. We are also currently working to demonstrate how simulation-based approaches can reveal allosteric effects in previously characterized experimental systems. In the future, we expect that the MSM approach described here will be a natural way to discover such interactions, which could be employed synergistically to design significantly improved conformational properties of peptidomimetics and other ligands.

■ ASSOCIATED CONTENT

Supporting Information

Figures S1−S11. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: wwuest@temple.edu.
*E-mail: voelz@temple.edu.

Author Contributions

The manuscript was written through contributions of all authors.

Funding

This research was supported in part by the National Science Foundation through major research instrumentation grant number CNS-09-58854 and funds from Temple University.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Christian Schwantes for help with the tICA algorithm, Holger Sondermann and George O'Toole for helpful conversations, and the many participants of Folding@home.

■ ABBREVIATIONS

PPI, protein−protein interaction; REMD, replica exchange molecular dynamics; MSM, Markov state model; tICA, time-structure-based independent component analysis

■ REFERENCES

(1) London, N.; Raveh, B.; Schueler-Furman, O. Druggable protein-protein interactions - from hot spots to hot segments. *Curr. Opin. Chem. Biol.* **2013**, *17*, 952−959.

(2) Naduthambi, D.; Bhor, S.; Elbaum, M. B.; Zondlo, N. J. Synthesis of a Tetrasubstituted Tetrahydronaphthalene Scaffold for α-Helix Mimicry via a MgBr 2-Catalyzed Friedel−Crafts Epoxide Cycloalkylation. *Org. Lett.* **2013**, *15*, 4892−4895.

(3) Brown, Z. Z.; Akula, K.; Arzumanyan, A.; Alleva, J.; Jackson, M.; Bichenkov, E.; Sheffield, J. B.; Feitelson, M. A.; Schafmeister, C. E. A Spiroligomer α-Helix Mimic That Binds HDM2, Penetrates Human Cells and Stabilizes HDM2 in Cell Culture. *PLoS One* **2012**, *7*, e45948.

(4) Becerril, J.; Rodriguez, J. M.; Wyrembak, P. N.; Hamilton, A. D. Inhibition of Protein−Protein Interactions by Peptide Mimics. *Protein Surf. Recognit.* **2010**, 1−27.

(5) Robinson, J. A. β-Hairpin Peptidomimetics: Design, Structures and Biological Activities. *Acc. Chem. Res.* **2008**, *41*, 1278−1288.

(6) Mas-Moruno, C.; Rechenmacher, F.; Kessler, H. Cilengitide: the first anti-angiogenic small molecule drug candidate. Design, synthesis and clinical evaluation. *Anti-Cancer Agents Med. Chem.* **2010**, *10*, 753−768.

(7) Syud, F. A.; Stanger, H. E.; Gellman, S. H. Interstrand Side Chain−Side Chain Interactions in a Designed β-Hairpin: Significance of Both Lateral and Diagonal Pairings. *J. Am. Chem. Soc.* **2001**, *123*, 8667−8677.

(8) Martín-Gago, P.; Ramón, R.; Aragón, E.; Fernández-Carneado, J.; Martin-Malpartida, P.; Verdaguer, X.; López-Ruiz, P.; Colás, B.; Cortes, M. A.; Ponsati, B.; Macias, M. J.; Riera, A. A tetradecapeptide somatostatin dicarba-analog: Synthesis, structural impact and biological activity. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 103−107.

(9) Lama, D.; Quah, S. T.; Verma, C. S.; Lakshminarayanan, R.; Beuerman, R. W.; Lane, D. P.; Brown, C. J. Rational Optimization of Conformational Effects Induced By Hydrocarbon Staples in Peptides and their Binding Interfaces. *Sci. Rep.* **2013**, *3*, 3451.

(10) Gibbs, A. C.; Kondejewski, L. H.; Gronwald, W.; Nip, A. M.; Hodges, R. S.; Sykes, B. D.; Wishart, D. S. Unusual beta-sheet periodicity in small cyclic peptides. *Nat. Struct. Biol.* **1998**, *5*, 284−288.

(11) Fuller, J. C.; Jackson, R. M.; Shirts, M. R. Configurational Preferences of Arylamide α-Helix Mimetics via Alchemical Free

Energy Calculations of Relative Binding Affinities. *J. Phys. Chem. B* **2012**, *116*, 10856−10869.
(12) Newell, P. D.; Boyd, C. D.; Sondermann, H.; O'Toole, G. A. A c-di-GMP Effector System Controls Cell Adhesion by Inside-Out Signaling and Surface Protein Cleavage. *PLoS Biology* **2011**, *9*, e1000587.
(13) Navarro, M. V. A. S.; Newell, P. D.; Krasteva, P. V.; Chatterjee, D.; Madden, D. R.; O'Toole, G. A.; Sondermann, H. Structural Basis for c-di-GMP-Mediated Inside-Out Signaling Controlling Periplasmic Proteolysis. *PLoS Biol.* **2011**, *9*, e1000588.
(14) Henchey, L.; Jochim, A.; Arora, P. A Hydrogen Bond Surrogate Approach for Stabilization of Short Peptide Sequences in Alpha-Helical Conformation. *Acc. Chem. Res.* **2008**, *41*, 1289−1300.
(15) Scholl, M.; Ding, S.; Lee, C. W.; Grubbs, R. H. Synthesis and Activity of a New Generation of Ruthenium-Based Olefin Metathesis Catalysts Coordinated with 1,3-Dimesityl-4,5-dihydroimidazol-2-ylidene Ligands. *Org. Lett.* **1999**, *1*, 953−956.
(16) Chatterjee, D.; Boyd, C. D.; O'Toole, G. A.; Sondermann, H. Structural Characterization of a Conserved, Calcium-Dependent Periplasmic Protease from Legionella pneumophila. *J. Bacteriol.* **2012**, *194*, 4415−4425.
(17) 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.
(18) Roe, D. R.; Okur, A.; Wickstrom, L.; Hornak, V.; Simmerling, C. *J. Phys. Chem. B* **2007**, *111*, 1846−1857.
(19) Kollman, P. A. Advances and Continuing Challenges in Achieving Realistic and Predictive Simulations of the Properties of Organic and Biological Molecules. *Acc. Chem. Res.* **1996**, *29*, 461−469.
(20) Onufriev, A.; Bashford, D.; Case, D. A. Exploring protein native states and large-scale conformational changes with a modified generalized born model. *Proteins: Struct., Funct., Bioinf.* **2004**, *55*, 383−394.
(21) Voelz, V. A.; Bowman, G. R.; Beauchamp, K.; Pande, V. S. Molecular Simulation of ab Initio Protein Folding for a Millisecond Folder NTL9(1−39). *J. Am. Chem. Soc.* **2010**, *132*, 1526−1528.
(22) Voelz, V. A.; Jäger, M.; Yao, S.; Chen, Y.; Zhu, L.; Waldauer, S. A.; Bowman, G. R. Slow Unfolded-State Structuring in Acyl-CoA Binding Protein Folding Revealed by Simulation and Experiment. *J. Am. Chem. Soc.* **2012**, *134*, 12565−12577.
(23) Voelz, V. A.; Singh, V. R.; Wedemeyer, W. J.; Lapidus, L. J.; Pande, V. S. Unfolded-State Dynamics and Structure of Protein L Characterized by Simulation and Experiment. *J. Am. Chem. Soc.* **2010**, *132*, 4702−4709.
(24) Lindorff-Larsen, K.; Piana, S.; Palmo, K.; Maragakis, P.; Klepeis, J. L.; Dror, R. O.; Shaw, D. E. Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins: Struct., Funct., Bioinf.* **2010**, *78*, 1950−1958.
(25) Buch, I.; Giorgino, T.; De Fabritiis, G. Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10184−10189.
(26) Wang, J.; Wolf, R. M.; Caldwell, J. W.; Kollman, P. A.; Case, D. A. Development and testing of a general amber force field. *J. Comput. Chem.* **2004**, *25*, 1157−1174.
(27) Jakalian, A.; Jack, D. B.; Bayly, C. I. Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation. *J. Comput. Chem.* **2002**, *23*, 1623−1641.
(28) Case, D. A.; Darden, T. A.; Cheatham, I.; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 12*; University of California: San Francisco, CA, 2012.
(29) Shirts, M. R.; Chodera, J. D. Statistically optimal analysis of samples from multiple equilibrium states. *J. Chem. Phys.* **2008**, *129*, 124105.
(30) Beauchamp, K. A.; Bowman, G. R.; Lane, T. J.; Maibaum, L.; Haque, I. S.; Pande, V. S. MSMBuilder2: Modeling Conformational Dynamics on the Picosecond to Millisecond Scale. *J. Chem. Theory Comput.* **2011**, *7*, 3412−3419.
(31) Shirts, M.; Pande, V. S. Screen savers of the world, unite! *Science* **2000**, *290*, 1903−1904.
(32) Schwantes, C. R.; Pande, V. S. Improvements in Markov State Model Construction Reveal Many Non-Native Interactions in the Folding of NTL9. *J. Chem. Theory Comput.* **2013**, *9*, 2000−2009.
(33) Perez-Hernandez, G.; Paul, F.; Giorgino, T.; De Fabritiis, G.; Noé, F., Identification of slow molecular order parameters for Markov model construction. arXiv preprint arXiv: 1302.6614 2013.
(34) Bowman, G. R. Improved coarse-graining of Markov state models via explicit consideration of statistical uncertainty. *J. Chem. Phys.* **2012**, *137*, 134111.
(35) Prinz, J.-H.; Wu, H.; Sarich, M.; Keller, B.; Senne, M.; Held, M.; Chodera, J. D.; Schütte, C.; Noé, F. Markov models of molecular kinetics: generation and validation. *J. Chem. Phys.* **2011**, *134*, 174105.
(36) Munoz, V.; Thompson, P.; Hofrichter, J. Folding dynamics and mechanism of b-hairpin formation. *Nature* **1997**, *390*, 196−199.
(37) Shandler, S. J.; Korendovych, I. V.; Moore, D. T.; Smith-Dupont, K. B.; Streu, C. N.; Litvinov, R. I.; Billings, P. C.; Gai, F.; Bennett, J. S.; DeGrado, W. F. Computational Design of a β-Peptide That Targets Transmembrane Helices. *J. Am. Chem. Soc.* **2011**, *133*, 12378−12381.
(38) Shandler, S. J.; Shapovalov, M. V.; Dunbrack, R. L., Jr.; DeGrado, W. F. Development of a Rotamer Library for Use in β-Peptide Foldamer Computational Design. *J. Am. Chem. Soc.* **2010**, *132*, 7312−7320.
(39) Korendovych, I. V.; Kim, Y. H.; Ryan, A. H.; Lear, J. D.; DeGrado, W. F.; Shandler, S. J. Computational Design of a Self-Assembling β-Peptide Oligomer. *Org. Lett.* **2010**, *12*, 5142−5145.
(40) Sammond, D. W.; Bosch, D. E.; Butterfoss, G. L.; Purbeck, C.; Machius, M.; Siderovski, D. P.; Kuhlman, B. Computational Design of the Sequence and Structure of a Protein-Binding Peptide. *J. Am. Chem. Soc.* **2011**, *133*, 4190−4192.
(41) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y.-E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popović, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P. ROSETTA3: an object-oriented software suite for the simulation and design of macromolecules. *Meth. Enzymol.* **2011**, *487*, 545−574.
(42) Fleishman, S. J.; Whitehead, T. A.; Ekiert, D. C.; Dreyfus, C.; Corn, J. E.; Strauch, E. M.; Wilson, I. A.; Baker, D. Computational Design of Proteins Targeting the Conserved Stem Region of Influenza Hemagglutinin. *Science* **2011**, *332*, 816−821.
(43) Drew, K.; Renfrew, P. D.; Craven, T. W.; Butterfoss, G. L.; Chou, F.-C.; Lyskov, S.; Bullock, B. N.; Watkins, A.; Labonte, J. W.; Pacella, M.; Kilambi, K. P.; Leaver-Fay, A.; Kuhlman, B.; Gray, J. J.; Bradley, P.; Kirshenbaum, K.; Arora, P. S.; Das, R.; Bonneau, R. Adding Diverse Noncanonical Backbones to Rosetta: Enabling Peptidomimetic Design. *PLoS One* **2013**, *8*, e67051.
(44) Renfrew, P. D.; Choi, E. J.; Bonneau, R.; Kuhlman, B. Incorporation of Noncanonical Amino Acids into Rosetta and Use in Computational Protein-Peptide Interface Design. *PLoS One* **2012**, *7*, e32637.
(45) Michel, J.; Harker, E. A.; Tirado-Rives, J.; Jorgensen, W. L.; Schepartz, A. In Silico Improvement of β3-Peptide Inhibitors of p53·hDM2 and p53·hDMX. *J. Am. Chem. Soc.* **2009**, *131*, 6356−6357.
(46) Schafmeister, C. E.; Po, J.; Verdine, G. L. An All-Hydrocarbon Cross-Linking System for Enhancing the Helicity and Metabolic Stability of Peptides. *J. Am. Chem. Soc.* **2000**, *122*, 5891−5892.
(47) Robinson, J. A. Design of Protein-Protein Interaction Inhibitors Based on Protein Epitope Mimetics. *ChemBioChem.* **2009**, *10*, 971−973.

(48) Singhal, N.; Pande, V. S. Error analysis and efficient sampling in Markovian state models for molecular dynamics. *J. Chem. Phys.* **2005**, *123*, 204909.
(49) Bowman, G. R.; Ensign, D. L.; Pande, V. S. Enhanced Modeling via Network Theory: Adaptive Sampling of Markov State Models. *J. Chem. Theory Comput.* **2010**, *6*, 787−794.
(50) Long, D.; Brüschweiler, R. Atomistic Kinetic Model for Population Shift and Allostery in Biomolecules. *J. Am. Chem. Soc.* **2011**, *133*, 18999−19005.
(51) Syud, F. A.; Espinosa, J. F.; Gellman, S. H. NMR-Based Quantification of β-Sheet Populations in Aqueous Solution through Use of Reference Peptides for the Folded and Unfolded States. *J. Am. Chem. Soc.* **1999**, *121*, 11577−11578.
(52) Celentano, V.; Diana, D.; De Rosa, L.; Romanelli, A.; Fattorusso, R.; D'Andrea, L. D. β-Hairpin stabilization through an interstrand triazole bridge. *Chem. Commun.* **2012**, *48*, 762.
(53) Schönherr, H.; Cernak, T. Profound methyl effects in drug discovery and a call for new C-H methylation reactions. *Angew. Chem., Int. Ed.* **2013**, *52*, 12256−12267.
(54) Dechantsreiter, M. A.; Planker, E.; Mathä, B.; Lohof, E.; Hölzemann, G.; Jonczyk, A.; Goodman, S. L.; Kessler, H. N-Methylated Cyclic RGD Peptides as Highly Active and Selective α Vβ 3Integrin Antagonists. *J. Med. Chem.* **1999**, *42*, 3033−3040.
(55) Demmer, O.; Frank, A. O.; Hagn, F.; Schottelius, M.; Marinelli, L.; Cosconati, S.; Brack-Werner, R.; Kremb, S.; Wester, H.-J.; Kessler, H. A Conformationally Frozen Peptoid Boosts CXCR4 Affinity and Anti-HIV Activity. *Angew. Chem., Int. Ed.* **2012**, *51*, 8110−8113.
(56) White, T. R.; Renzelman, C. M.; Rand, A. C.; Rezai, T.; McEwen, C. M.; Gelev, V. M.; Turner, R. A.; Linington, R. G.; Leung, S. S. F.; Kalgutkar, A. s.; Bauman, J. N.; Zhang, Y.; Liras, S.; Price, D. a.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, R. S. On-resin N-methylation of cyclic peptides for discovery of orally bioavailable scaffolds. *Nat. Chem. Biol.* **2011**, *7*, 810−817.
(57) Chatterjee, J.; Rechenmacher, F.; Kessler, H. N-Methylation of Peptides and Proteins: An Important Element for Modulating Biological Functions. *Angew. Chem., Int. Ed.* **2012**, *52*, 254−269.
(58) Russell, S. J.; Blandl, T.; Skelton, N. J.; Cochran, A. G. Stability of Cyclic β-Hairpins: Asymmetric Contributions from Side Chains of a Hydrogen-Bonded Cross-Strand Residue Pair. *J. Am. Chem. Soc.* **2003**, *125*, 388−395.