JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

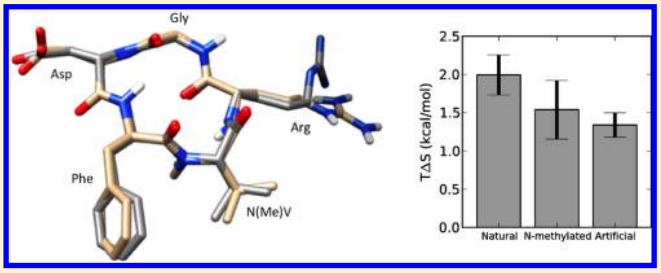
Molecular Simulation of Conformational Pre-Organization in Cyclic RGD Peptides

Amanda E. Wakefield, William M. Wuest, and Vincent A. Voelz*

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

Supporting Information

ABSTRACT: To test the ability of molecular simulations to accurately predict the solution-state conformational properties of peptidomimetics, we examined a test set of 18 cyclic RGD peptides selected from the literature, including the anticancer drug candidate cilengitide, whose favorable binding affinity to integrin has been ascribed to its pre-organization in solution. For each design, we performed all-atom replica-exchange molecular dynamics simulations over several microseconds and compared the results to extensive published NMR data. We find excellent agreement with experimental NOE distance restraints, suggesting that molecular simulation can be a useful tool for the computational design of pre-organized solution-state structure. Moreover, our analysis of conformational populations estimates that, despite the potential for increased flexibility due to backbone amide isomerizaton, N-methylation provides about 0.5 kcal/mol of reduced conformational entropy to cyclic RGD peptides. The combination of pre-organization and binding-site compatibility explains the strong binding affinity of cilengitide to integrin.



INTRODUCTION

Peptidomimetics—molecules that mimic the folded structure of a protein—represent an important class of compounds able to target protein−protein interactions (PPIs).[1,2] Most peptidomimetics designed to disrupt PPIs achieve enhanced binding affinity through pre-organization of the unbound state causing rigidification of the chemical scaffold.[3,4] Stapled peptides, for example, contain a hydrocarbon olefinic cross-link to help constrain helix mimetics to their folded conformation.[5,6] Computational design of pre-organized structure has the potential to accelerate drug design by eliminating costly synthesis and characterization. Yet, detailed prediction of the conformational and chemical properties of peptidomimetics remains an important challenge.[7−14]

In the early 1970s, Ruoslahti and Piersbacher first characterized the Arg-Gly-Asp (RGD) peptide sequence, which adopts a conserved hairpin loop structure of integrin-binding proteins such as fibronectin, fibrinogen, and vitronectin.[15] This seminal work inspired a great number of chemists to develop rigidified scaffolds to lock the conserved tripeptide motif in the desired state, providing a test bed for a number of peptidomimetic designs. Most notably, the Kessler group was able to develop a successful class of molecules by screening conformationally constrained cyclic RGD mimics containing N-methylated residues and D-amino acids.[16] This work culminated in the discovery of cyclic pentapeptide cyclo(RGDf-[*N*-Me]V),[17] the former cancer drug candidate known as cilengitide[18] (Figure 1). Cilengitide inhibits the binding of tumor cells to the extracellular matrix, which prevents vascularization and induces cell apoptosis.

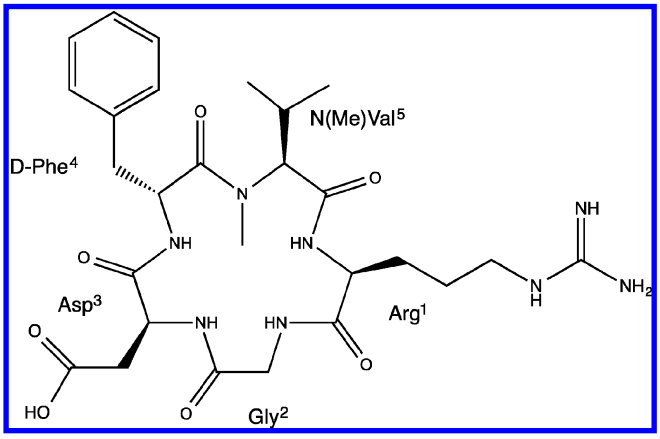


Figure 1. Chemical structure of cyclo(RGDf-[*N*-Me]V), the former cancer drug candidate known as cilengitide.

Notably, this successful design did not come from consideration of the receptor structure (a high-resolution crystal structure of integrin's three head groups was not solved until 2001),[19] but from a deliberate strategy of screening for mimics with decreased flexibility in solution. As a result, there exists a wealth of published data on both the binding affinities of designed mimics and the solution-state structures obtained by NMR. According to NMR studies, cilengitide—the highest-affinity binder—has a remarkably pre-organized solution

Received: December 26, 2014
Published: March 5, 2015

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

Table 1. 18 Cyclic RGD Peptides Simulated in This Study, With Descriptions of NOE Measurements, Where Available[a]

	peptide	type	ref	violations	no. NMR distances	avg. violation (Å)
1	cyclo(RGDkV)	natural	23	16	29	0.129
2	cyclo(RGDpV)	natural	23	7	22	0.133
3	cyclo(RGDwV)	natural	23	10	40	0.150
4	cyclo(RGDfK)	natural	23	3	31	0.087
5	cyclo(RGDKv)	natural	23	4	29	0.069
6	cyclo(RGDWv)	natural	23	17	37	0.312
7	cyclo(RGDFV)	natural	25	16	24	0.135
8	cyclo(VfdGr)	natural	25	15	24	0.209
9	cyclo(vfdGR)	natural	25	14	33	0.147
10	cyclo(vfdGr)	natural	25	16	24	0.259
11	cyclo(RGDfV)	natural	25			
12	cyclo(RGD“tic”V)	artificial	23	9	29	0.082
13	cyclo(RGD″R-ANC″)	artificial	24	13	51	0.395
14	cyclo(N(Me)R-GDfV)	N-methylated	17			
15	cyclo(*R*-Sar-DfV)[b]	N-methylated	17			
16	cyclo(RG-N(Me)D-fV)	N-methylated	17			
17	cyclo(RGD-N(Me)f-V)	N-methylated	17	15	41	0.159
18	cyclo(RGDf-N(Me)V)	N-methylated	17			

[a]Lower case letters denote D-amino acid residues. [b]Sar denotes the sarcosine residue, (*N*-methyl)glycine.

structure.[17] We looked toward this class of compounds as an ideal model system for computational design, as numerous synthetic analogs have been constructed and characterized by extensive 2D NMR studies.[16,17,20−25]

To test the extent to which molecular simulations can accurately model conformational ensembles of designed peptidomimetics, we simulated 18 cyclic RGD peptide designs from the Kessler group (Table 1) using all-atom replica-exchange molecular dynamics (REMD) simulations over several microseconds. We find excellent agreement with experimental NOE distance restraints, suggesting that molecular simulation can be a useful tool for the computational design of pre-organized solution-state structure. A survey of all simulations performed offers insight on how chemical modifications confer structural pre-organization in solution.

METHODS

Molecular Simulation. Model peptides were constructed using the AmberTools *tleap* program. The non-natural amino acids were built using UCSF Chimera. Generalized AMBER force field[26] (GAFF) parameters were assigned for non-natural amino acids, with partial charges computed using the AM1-BCC method.[27] The AMBER ff96 forcefield[28] was used in all simulations, along with the OBC (Onufriev, Bashford, and Case) GBSA solvation model.[29] This combination has been shown to accurately reproduce secondary and tertiary structural properties,[30] and has been applied successfully to ab initio protein folding.[31]

All-atom REMD simulations were performed using GROMACS 4.5.4,[32] on the Owlsnest high-performance computing cluster at Temple University. REMD simulations were performed with 8 temperature replicas exponentially spaced from 300 to 450 K. The number of replicas was chosen to ensure thorough conformational sampling, as predicted by overlap in sampled energy distributions (Figure S1a). The maximum temperature was limited to 450 K to prevent backbone amide *cis*/*trans* isomerization for non-methylated residues. Exchanges between neighboring temperature replicas were attempted every 10 ps, with average acceptance probabilities ranging between 0.65 and 0.79 for all simulations. Each simulation ran for 2.4 μs, with snapshots recorded every 10 ps, for a total of 19.2 μs of trajectory data for each design.

A particular problem in simulating N-methylated residues is the sampling of *cis*/*trans* amide isomers, which are separated by high energy barriers. To accelerate sampling over this degree of freedom, ω-angle torsional barriers were scaled for the highest temperature replicas. Replicas 1 through 5 had torsion scaling parameters of 1.0, while replicas 6, 7, and 8 had values 0.5, 0.25 and 0, respectively. This resulted in excellent sampling of *cis* ⇌ *trans* ω-angle transitions, with decorrelation on the time scale of tens of picoseconds, while non-N-methylated residues equilibrate exclusively in the *trans* state (Figure S1b−d). Given the finite number of dihedral backbone conformations available to a cyclic pentapeptide, and the vast amount of simulation (8 replicas × 2.4 μs) compared to this decorrelation time, it is clear that our simulations exhaustively sample all relevant states.

Conformational Clustering. MSMBuilder2[33] was used for conformational clustering of the lowest-temperature replica data, using a backbone dihedral-based distance metric (specifically, for n dihedral angles φ_j, j = 1, ..., n, the distance is the L2-norm of differences in complex vectors $[\exp(i\varphi_1), \exp(i\varphi_2), ..., \exp(i\varphi_n)]$). The k-centers algorithm was used to derive conformational clusters of similar conformational volume suitable for calculating conformational entropies. The dihedral cluster radius (1.8, corresponding to an average dihedral angle difference of ∼27°) was chosen by examining how the average variance in NOE distance observables for cilengitide changed as a function of the number of clusters (see Results). This cluster radius was then used for the conformational clustering of all other trajectory data.

Estimates of conformational entropy for each design were computed as $S = -\Sigma_i\, p_i \ln p_i$, where p_i are the equilibrium populations estimated for each cluster, computed from the lowest-temperature replica. Although more sophisticated methods for entropy estimation exist,[34−36] the cluster-based approach used here is straightforward and computationally robust. Cluster-based entropy estimates calculated from backbone-rmsd clustering (with a radius of 0.21) gave very similar results (data not shown). We would like to emphasize

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

that our cluster-based entropies are used here to compare the pre-organization of different peptide designs and that more precise methods are required to calculate absolute entropies that can be compared to thermodynamic experiments.

Comparison of Simulations against Experimental NOE Measurements. NOE measurements for 13 of the 18 cyclic RGD peptides were published in a series of papers by Kessler et al.[17,23−25] (Table 1). These data are summarized in the Supporting Information (Tables S2−S14). The data consist of between 22 and 51 NOE interproton distance measurements for each peptide, represented by upper and lower distance bounds derived from the NOE intensities as well as the average interproton distances observed in their original molecular dynamics refinement studies. The original refinement studies were limited to picosecond dynamics simulations with distance restraints.[37]

To compare simulated conformational ensembles to experimental NOE measurements, distributions of interproton distances were computed from lowest-temperature replica data, enabling comparison with the published upper and lower distance bounds. Ensemble-averaged distances were computed as $\langle r^{-6}\rangle^{-1/6}$ to enable comparison with the published NOE distances. The number of distance violations was computed as the number of interproton distances for which the computed $\langle r^{-6}\rangle^{-1/6}$ value fell outside of the upper and lower bounds. Average distances were used for chemically equivalent protons.

Bayesian Inference of Conformational Populations (BICePs). Briefly, BICePs[38] uses a Bayesian approach to infer a distribution of conformational populations, by sampling a posterior probability function:

$$P(X|D) \propto P(D|X)P(X)$$

where X is one of a set of molecular conformations, and D represents the experimental data. $P(X)$ is a prior probability function, calculated from the results of molecular simulation. $P(D|X)$ is a likelihood function representing experimental restraints. In this case, X is a set of 100 conformational clusters obtained using methods described below, and the prior $P(X_i)$ is calculated from the population of each cluster in lowest-tempertature simulation replica.. For experimental restraints, we used Gaussian distributions of the deviation between r^{-6}-averaged interproton distances d_j, and the r^{-6}-average of experimental upper and lower distance bounds d_j^{exp}. All distance distributions were parametrized by a single standard deviation σ_{d} and distance scaling factor $\gamma' = \gamma^{-1/6}$ (see ref 38 for details), resulting in the likelihood function:

$$P(D|X, \sigma_{\mathrm{d}}, \gamma'') = \Pi_j[(2\pi\sigma_{\mathrm{d}}^2)^{-1/2}\exp(-w_j(d_j - \gamma' d_j^{\mathrm{exp}})^2/2\sigma_{\mathrm{d}}^2) / P_{\mathrm{ref}}(d_j|\beta_j)]$$

where groups of n equivalent protons were given equal weights of $w_j = 1/n$. The σ_{d} and γ' are treated as nuisance parameters whose distribution is estimated from posterior sampling along with $P(X|D)$. The reference potentials $P_{\mathrm{ref}}(d_j|\beta_j)$ for each distance are exponential distributions whose first moments β_j are set to the estimated mean distance across all conformational clusters. Ten million steps of Markov chain Monte Carlo (MCMC) in $(X, \sigma_{\mathrm{d}}, \gamma')$ were performed for each BICePs calculation.

RESULTS

All 18 cyclic RGD peptide designs (Table 1) were simulated using all-atom REMD, as described in Methods. The peptides can be classified into three categories: peptides composed of natural amino acids, peptides containing a single N-methylated amino acid, and peptides containing a non-natural (synthetic) residue. Chemical structures of synthetic residues TIC and (R)-ANC are shown in Figure 2.

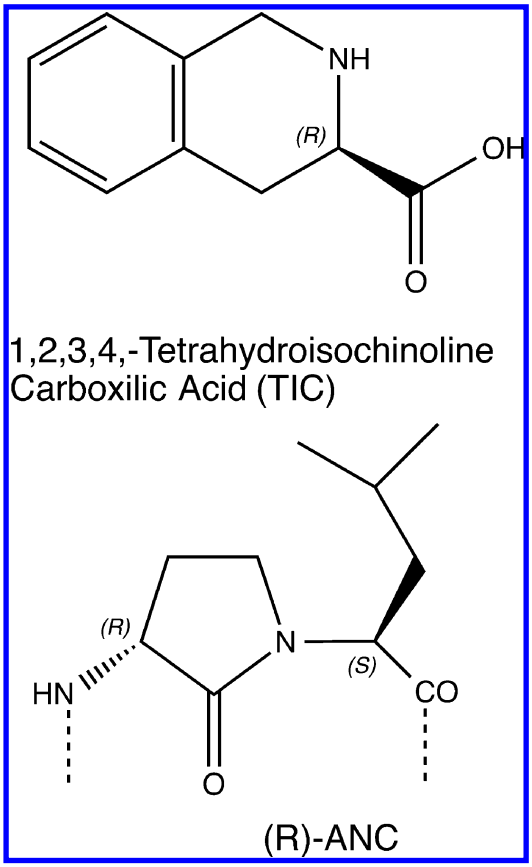


Figure 2. Synthetic amino acid 1,2,3,4-tetrahydroisochinoline carboxylic acid (TIC), structure shown above, was used in the cyclic peptide RGD"TIC"V (Table S7). The synthetic amino acid ANC, with *R*-stereochemistry, structure shown above, was used in the cyclic peptide RGD "*R*-ANC" (Table S3).

Simulated Conformational Ensembles Show Excellent Agreement with Experimental NMR Measurements. On average, simulations of the 13 cyclic RGD peptides with published NOE data show interproton distance distributions that agree very well with experiment. Figure 3 shows examples of such distributions for Arg HN-Arg HA, Asp HN-Asp HA, Phe HN-Gly HA, and Val HA-Val HB observed in simulations of cilengitide. The $\langle r^{-6}\rangle^{-1/6}$ average distance for most pairs fall between the upper and lower NOE bounds, and for distances falling outside these bounds, violations are found to be small. For cilengitide, simulated and experimental interproton distances agree with a correlation coefficient of $R^2 = 0.829$, and the average violation of distance bounds was only 0.16 Å (Figure 4). Similar results were found for the other 12 simulated molecules, with average correlation coefficients of $R^2 = 0.700$, and an average violation of 0.176 Å (Figure S2). This extent of violation is modest compared to the violations typically accepted in NMR structural refinements (<0.5 Å), underscoring how closely our simulations were able to reproduce the published NOE data.

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

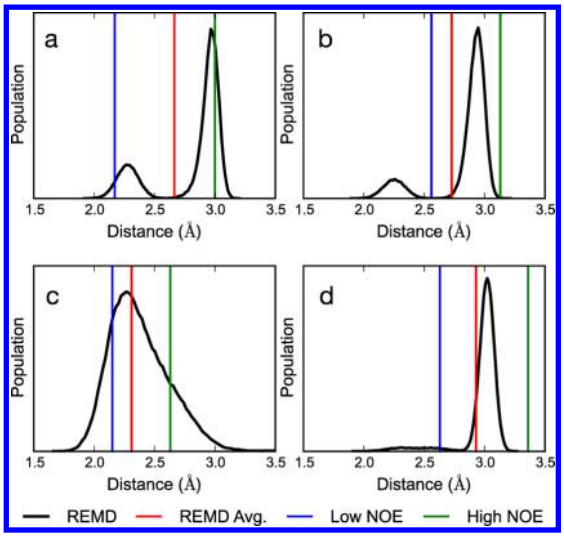


Figure 3. NOE distances of interproton pairs determined from 2.4 μs simulations of cilengitide show excellent agreement with previously published NOE ranges. Shown as examples are four of the 41 interproton distances for cilengitide: Arg HN-Arg HA (a), Asp HN-Asp HA (b), Phe HN-Asp HA(c), and Val HA-Val HB (d). The NOE distances are shown in black with the $\langle r^{-6}\rangle^{-1/6}$ average shown as a red line. The experimental ranges are $\langle r^{-6}\rangle^{-1/6}$ values. Lower and upper bounds of the range are shown as blue and green lines, respectively.

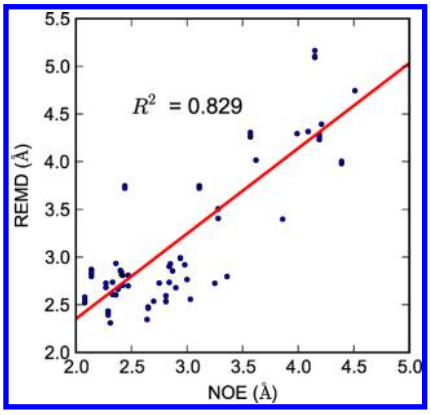


Figure 4. The $\langle r^{-6}\rangle^{-1/6}$ average NOE distances for interproton pairs determined from the simulation of cilengitide show excellent agreement with previously published NOE data. The experimental average distances were calculated by taking the $\langle r^{-6}\rangle^{-1/6}$ average of the lower and upper bounds.

Determining an Optimal Resolution for Conformational Clustering. To examine the extent of pre-organization in solution, we performed conformational clustering as described in Methods. Our goal was to assign the trajectory data into a set of discrete conformational states, fine-grained enough that structural observables (NOE distances) are well-defined within each cluster, but not so fine-grained that cluster populations suffer from finite sampling error. To choose an appropriate resolution, we performed multiple conformational clustering calculations on the cilengitide trajectory data, specifying fixed numbers of clusters ranging from 2 to 150. In each instance, we calculated the average variance in the interproton distances within each cluster (Figure 5). These studies show that the average variance decreases rapidly as the number of clusters increases, but begins to stabilize beyond 80 clusters. Based on these results, we chose a dihedral cluster radius of 1.8 (yielding ~100 clusters for cilengitide) for all subsequent clustering calculations.

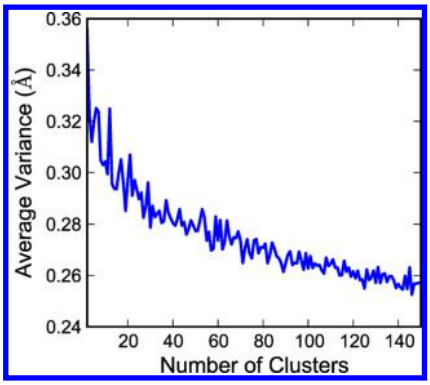


Figure 5. Multiple conformational clustering calculations were performed on the cilengitide trajectory data, specifying fixed numbers of clusters ranging from 2 to 150. In each instance, we calculated the average variance in the interproton distances within each cluster. These studies show that the average variance decreases rapidly as the number of clusters increases, but begins to stabilize beyond 80 clusters. Subsequently we chose a dihedral cluster radius of 1.8 (yielding ~100 clusters for cilengitide) for all subsequent clustering calculations.

BICePs Calculations Show That Simulations Produce Objectively Better Models of Conformational Populations than Could Be Inferred from Experimental Constraints Alone. Since these RGD peptides are cyclic, could it be that the good agreement we find is a trivial consequence of the proximity of protons in cyclic pentapeptides? To test this idea, we performed an analysis of the cilengitide simulations using our recently developed Bayesian inference of conformational populations (BICePs) method.[38] A key feature of BICePs is the proper use of reference potentials representing the "null distribution" of interproton distances, so that the information in experimental restraints can be fairly compared with information from molecular simulations. For example, a short-distance NOE measurement is informative only if our null hypothesis is that such distances are typically large.

Briefly, BICePs uses MCMC to sample from a Bayesian posterior model consisting of a likelihood function (specified by the experimental constraints and the reference potentials, see Methods) and a prior (obtained from computational results). In this way, posterior distributions of conformational populations and model parameters are estimated. For the full technical details of BICePs, we refer the reader to Methods and to ref 38.

The BICePs results for cilengitide show that conformational populations estimated from simulations are highly compatible with the experimental restraints (Figure 6). Our simulation data predict a small number of conformational states to be highly populated, much more so than can be inferred from the NOE measurements alone. The most over-represented of these conformational states (cluster 53) matches the pose seen in the crystal structure of cilengitide bound to integrin (see below). Rank-ordering of 100 conformational clusters by their computed correlation coefficients (R^2) between simulated and experimental interproton distances show that the clusters predicted to have the highest populations are also the ones that agree best with experimental measurements. Posterior estimates of distributions σ_d are centered about 0.75 Å, indicating sub-Å deviations between simulated and experimental distances.

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

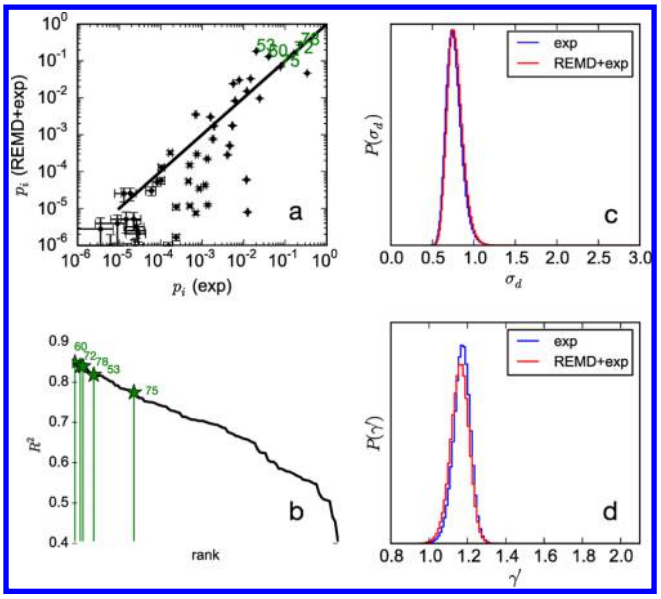


Figure 6. BICePs results for cilengitide show that predicted conformational populations are in excellent agreement with experimental measurements. (a) Predicted conformational populations for a model which only considers the experimental restraints (exp) versus a model that additionally includes information from simulation (REMD + exp). In green are labeled the indices of conformational states (100 total) predicted by simulations to have populations >5%. The most over-represented of these (cluster 53) matches the pose seen in the crystal structure of cilengitide bound to integrin (see main text). (b) Rank-ordering of the 100 conformational clusters by their computed correlation coefficients (R^2) between simulated and experimental interproton distances). Posterior estimates are shown for (c) distributions of σ_d and (d) distributions of scaling parameter γ'.

Posterior distributions of the scaling parameter γ' are centered near 1 (~1.17). Moreover, we use BICePs to compute Bayes factors $P(M_1)/P(M_2)$, where M_1 is a model using information from both simulation and experiment, and M_2 is the "null hypothesis" model with no information from simulation (a uniform prior for all conformations). We obtain an estimated Bayes Factor of 1.95, indicating that not only does our simulation agree well with the experimental data, it constitutes an objectively better model.

Simulations of Cilengitide Reveal a Highly Pre-Organized Solution State. Conformational clustering of our cilengitide simulations reveals a remarkable similarity between the structure of the most populated cluster (5.3%) and the crystal structure of cilengitide bound to integrin (PDB: 1L5G), with backbone atoms agreeing to 0.25 Å rmsd (Figure 7). Other populated clusters are conformationally similar. In agreement with the seminal NMR studies by Kessler et al.,[17] simulations of cilengitide show the RGD sequence to largely populate a γ-turn at the Gly position (required for selective binding to integrin $\alpha_V\beta_3$) and inverse-γ turns at the Arg and Asp positions. In this respect, cilengitide is not unlike many of the other N-methylated designs. To further examine the structural features that distinguish cilengitide from the other N-methylated peptides, we compared distributions of φ and ψ backbone dihedral angles for the Arg, Asp, and Gly residues in the turn region (Figure 8). This analysis shows that while other peptides are equally (or even more) pre-organized in solution, only cilengitide has RGD backbone dihedral angles that perfectly align with those of bound structure. In part, this is

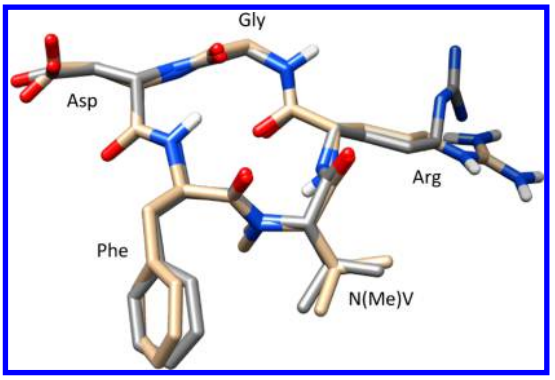


Figure 7. Structural comparison of the simulated solution structure of cilengitide and the experimental crystal structure of cilengitide bound to integrin. The generator of the most populated cluster of cilengitide (brown) observed in simulations represents 5.3% of the total population; other populated clusters are conformationally similar. The predicted structure agrees remarkably with the crystal structure of cilengitide (gray) bound to integrin (PDB 1L5G) within 0.25 Å rmsd for all backbone atoms.

to be expected, due to the fact that the only available cocrystal structure contains cilengitide, not other N-methylated peptides. Still, the high binding affinity of cilengitide versus other N-methylated designs suggests that what sets cilengitide apart is its pre-organization toward a specific and highly complementary bound-state conformation.

The *cis*-amide populations of N-methylated residues, calculated from the lowest-temperature replica, were found to be small (Table 2). This agrees with the low (or absent) percentage of *cis*-amide population found in experimental studies, particularly for residues with branched side chains.[39]

Relative Conformational Entropy Estimates for Classes of Designed Peptidomimetics. Conformational clustering was performed on the lowest-temperature trajectories of all 18 cyclic RGD peptides simulated, yielding estimates of conformational entropies for each design (Table 3). Although the cluster-based entropy estimates are too crude to make detailed comparisons between competing designs, we can use them to examine the relative conformational entropies of the three categories of cyclic RGD peptides simulated: peptides composed of natural amino acids, peptides containing a single N-methylated amino acid, and peptides containing a non-natural (synthetic) residue.

Conformational entropy estimates averaged over these three groups yield an overview of the factors that influence conformational heterogeneity in solution (Figure 9). According to our estimates, N-methylation reduces the conformational entropy ($T\Delta S$) of natural peptides by approximately 0.5 kcal/mol. Substitution of an artificial amino acid such as TIC or (R)-ANC provides the greatest increase in the rigidity for cyclic peptides. These results are consistent with experimental studies showing that, despite the potential for increased flexibility via backbone amide *cis*/*trans*-isomerization, N-methylation of cyclic peptides often induces steric constraints that lead to "frozen" conformations with enhanced pre-organization.[40,41]

■ DISCUSSION

Our unrestrained REMD simulations of cyclic RGD peptides show excellent quantitative agreement with experimental NOE distances and reproduce the conformational properties originally reported by Kessler et al.,[17,23−25] including a highly

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

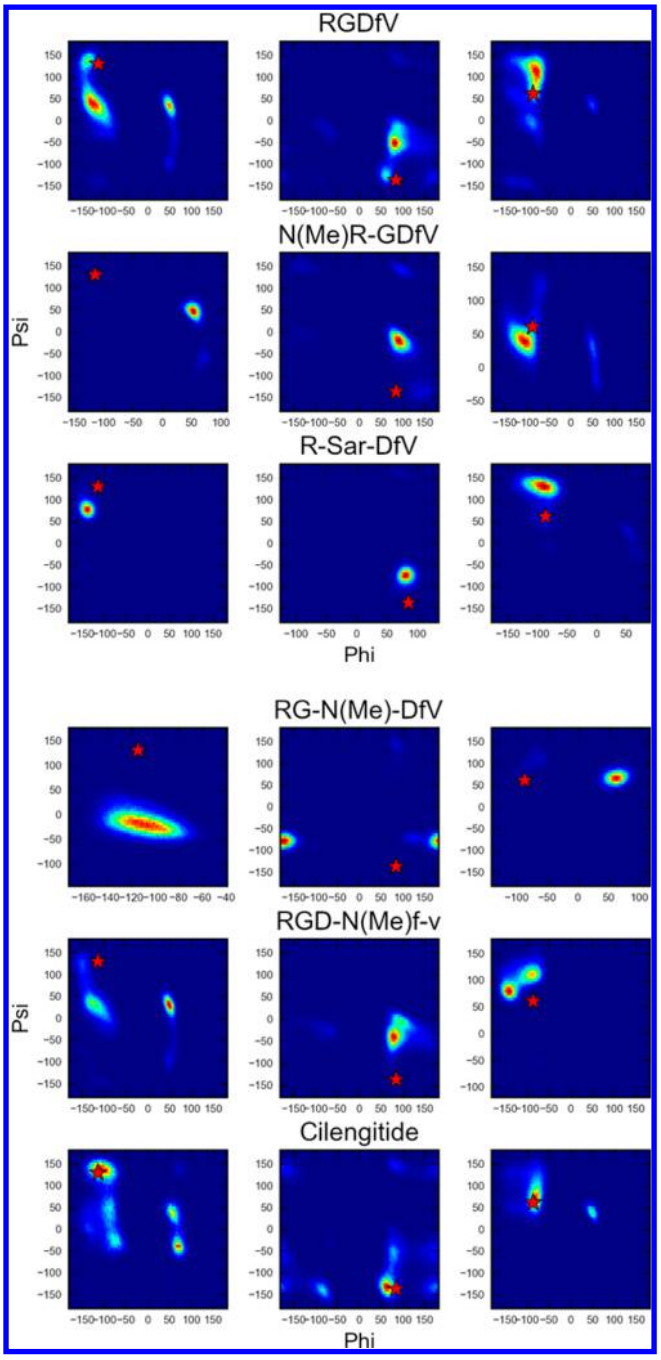


Figure 8. Backbone dihedral angle distributions for Arg[1], Gly[2], and Asp[3] residues (left, center, and right columns, respectively) for N-methylated variants of cyclo(RGDfV). Red stars represent the φ and ψ backbone angles of the bound cilengitide crystal structure. Whereas many of the designs show reduced conformational flexibility in solution due to N-methylation, only the solution simulations of cilengitide (bottom row) show a precise preference for the bound state.

pre-organized solution state for cilengitide.[18] In those studies, molecular dynamics simulation was used primarily as a tool for structure refinement, using trajectory lengths of ∼150 ps to anneal conformational ensembles with NOE distance restraints.[37] With 20 years of hindsight (and faster computers), we can now perform unrestrained molecular simulations over tens of microseconds to obtain converged equilibrium conformational properties of cyclic RGD peptides. The

Table 2. Estimated *cis*-Amide Populations from the Lowest-Temperature Replica (300 K) of 2.4 μs REMD Simulations

peptide sequence	*cis*-amide population (%)
cyclo(RGDfV)	0.007
cyclo(N(Me)R-GDfV)	0.21
cyclo(*R*-Sar-DfV)	0.16
cyclo(RG-N(Me)D-fV)	8.4
cyclo(RGD-N(Me)f-V)	0.4
cyclo(RGDf-N(Me)V) (cilengitide)	0.12

Table 3. Backbone Conformational Entropies of the 18 Cyclic RGD Peptides Simulated in This Study, Estimated at 300 K Using a Cluster-Based Approach[a]

		hybrid backbone cluster	
	peptide sequence	no. of clusters	conformational entropy $T\Delta S$ (kcal/mol)
1	cyclo(RGDkV)	64	3.502
2	cyclo(RGDpV)	47	2.985
3	cyclo(RGDwV)	73	3.591
4	cyclo(RGDfK)	69	3.374
5	cyclo(RGDKv)	71	3.040
6	cyclo(RGDWv)	47	2.262
7	cyclo(RGDFV)	108	3.685
8	cyclo(VfdGr)	82	3.437
9	cyclo(vfdGR)	122	3.856
10	cyclo(vfdGr)	99	3.560
11	cyclo(RGDfV)	84	3.515
12	cyclo(RGD″tic″V)	32	2.436
13	cyclo(RGD″*R*-ANC″)	31	2.061
14	cyclo(N(Me)*R*-GDfV)	39	2.388
15	cyclo(*R*-Sar-DfV)	37	1.904
16	cyclo(RG-N(Me)D-fV)	40	2.162
17	cyclo(RGD-N(Me)f-V)	67	2.978
18	cyclo(RGDf-N(Me)V) (cilengitide)	101	3.485

[a]The values are useful only for relative comparison.

accuracy of our predictions suggests we can use simulation alongside experiment for the purpose of molecular design. Indeed, REMD simulations coupled to docking studies have recently been employed to successfully design and optimize a set of cyclic peptides for an anticancer target.[42] Similar simulations have helped to screen β-hairpin mimetics for solution-state pre-organization and structural similarity to a known bound state.[9]

Our simulation work also provides partial insight into the role of pre-organization in determining binding affinity. Among the N-methylated variants, we find no straightforward correlation between low conformational entropy in solution and binding affinity (IC_{50}s) to $\alpha_v\beta_3$ integrin (Table S1). Although it is possible that errors in our cluster-based entropy estimator obfuscates any such correlation, a more likely reason is simply the fact that ligand binding affinity depends sensitively on many factors beyond conformational entropy loss, such as optimal enthalpic interactions in the binding site, shape complementarity and desolvation. Rationalizing the competing roles of conformational entropy and enthalpy for ligand binding[43] a priori, even using sophisticated modeling approaches,[44] can be highly non-nontrivial. In the future, we hope to perform large-scale free energy perturbation studies[45,46] of relative binding affinity for a series of N-methylated RGDfV pentapeptides, in order to better understand the key

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

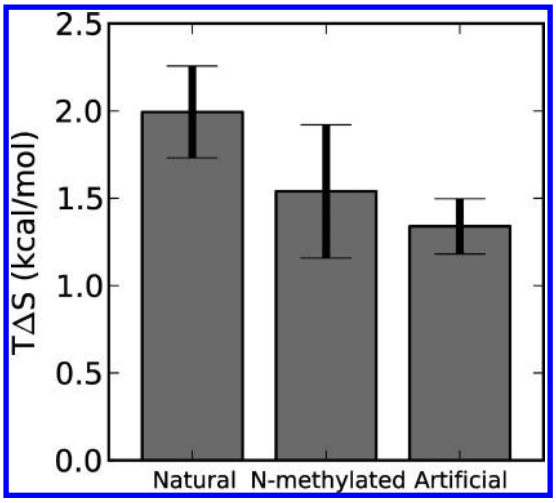


Figure 9. Solution-state simulations of cyclic RGD peptides predict that N-methylation reduces the conformational entropy ($T\Delta S$) of natural peptides by approximately 0.5 kcal/mol, with even more rigidity provided by artificial amino acids TIC or (*R*)-ANC and provides the greatest increase in the rigidity for cyclic peptides. Entropies were estimated using a cluster-based calculation at 300 K, averaged over the three categories of natural (11 designs), artificial (2 designs), and N-methylated (5 designs) cyclic peptides. Error bars reflect standard deviations within each category.

factors that influence the binding affinity of cyclic RGD peptides to integrin.

Our simulations of cyclic RGD peptides suggest that an approach to designing high-affinity peptidomimetics is to discover molecules that are not only pre-organized, but pre-organized toward a specific high-affinity bound-state conformation. The simulations we present here clearly show this to be the case for cilengitide. Other peptidomimetics, such as stapled peptides, have been developed using this idea as the guiding design principle. Our calculated estimates of relative conformational entropies for different classes of cyclic peptides give further support to this idea. While our simulations show that artificial amino acids can be used to greatly rigidify solution-state conformational ensembles of cyclic RGD peptides, it is at the risk of stabilizing non-native conformations that bind poorly. N-methylation, on the other hand, provides a significant reduction in conformational entropy, while preserving the structural properties needed for strong binding. Consistent with this idea are recent results showing that di-N-methylation of cyclo(RGDfV) decreases the entropy of the molecule but at the cost of the binding affinity.[47]

CONCLUSION

In this work, we have performed comprehensive REMD simulations of 18 cyclic RGD peptide designs, and find very good quantitative agreement with experimental NOE measurements. Consistent with previous NMR studies, we find a remarkable extent of pre-organization for cilengitide, cyclo-(RGDf-[*N*-Me]V), whose predominant solution-state conformation is nearly identical to its bound-state conformation. Cluster-based estimates of average conformational entropies for different classes of cyclic peptides suggest that N-methylation can provide ∼0.5 kcal/mol of entropic stabilization and that artificial amino acids can provide even more, but at the risk of altering desirable bound-state conformational properties. The quantitative agreement of simulation and experiment, as well as the insights obtained by systematically comparing different classes of cyclic peptide designs, suggests that molecular simulation can be a much more useful tool for the computational design of pre-organized solution-state structure for peptidomimetics.

ASSOCIATED CONTENT

Supporting Information

Figures S1 and S2; Tables S1−S14. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*voelz@temple.edu

Funding

This research was supported in part by the National Science Foundation through MCB-1412508, major research instrumentation grant no. CNS-09-58854, and funds from Temple University.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank members of the Voelz lab for helpful scripts and coding advice, the Kessler lab for helpful clarification of previously published NMR data, and Dr. David Dalton for invaluable guidance.

ABBREVIATIONS

RGD (Arg-Gly-Asp); REMD, replica exchange molecular dynamics; NOE, nuclear Overhauser effect; NMR, nuclear magnetic resonance; GBSA, generalized Born/surface area

REFERENCES

(1) Lao, B. B.; Drew, K.; Guarracino, D. A.; Brewer, T. F.; Heindel, D. W.; Bonneau, R.; Arora, P. S. Rational Design of Topographical Helix Mimics as Potent Inhibitors of Protein-Protein Interactions. *J. Am. Chem. Soc.* **2014**, *136*, 7877−7888.

(2) Mason, J. M. Design and Development of Peptides and Peptide Mimetics as Antagonists for Therapeutic Intervention. *Future Med. Chem.* **2010**, *2*, 1813−1822.

(3) Baek, S.; Kutchukian, P. S.; Verdine, G. L.; Huber, R.; Holak, T. A.; Lee, K. W.; Popowicz, G. M. Structure of the Stapled p53 Peptide Bound to Mdm2. *J. Am. Chem. Soc.* **2012**, *134*, 103−106.

(4) Robinson, J. A. Design of Protein-Protein Interaction Inhibitors Based on Protein Epitope Mimetics. *ChemBioChem* **2009**, *10*, 971−973.

(5) 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.

(6) Kim, Y.-W.; Grossmann, T. N.; Verdine, G. L. Synthesis of All-Hydrocarbon Stapled [alpha]-Helical Peptides by Ring-Closing Olefin Metathesis. *Nat. Protoc.* **2011**, *6*, 761−771.

(7) 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.

(8) Yu, H.; Lin, Y.-S. Toward Structure Prediction of Cyclic Peptides. *Phys. Chem. Chem. Phys.* **2015**, *17*, 4210−4219.

(9) Razavi, A. M.; Wuest, W. M.; Voelz, V. A. Computational Screening and Selection of Cyclic Peptide Hairpin Mimetics by Molecular Simulation and Kinetic Network Models. *J. Chem. Inf. Model.* **2014**, *54*, 1425−1432.

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813

(10) Fuller, J. C.; Jackson, R. M.; Edwards, T. A.; Wilson, A. J.; Shirts, M. R. Modeling of Arylamide Helix Mimetics in the p53 Peptide Binding Site of hDM2 Suggests Parallel and Anti-Parallel Conformations Are Both Stable. *PLoS One* **2012**, *7*, e43253.
(11) Galan, J. F.; Brown, J.; Wildin, J. L.; Liu, Z.; Liu, D.; Moyna, G.; Pophristic, V. Intramolecular Hydrogen Bonding in ortho-Substituted Arylamide Oligomers: A Computational and Experimental Study of ortho-Fluoro- and ortho-Chloro- N-methylbenzamides. *J. Phys. Chem. B* **2009**, *113*, 12809−12815.
(12) Galan, J. F.; Tang, C. N.; Chakrabarty, S.; Liu, Z.; Moyna, G.; Pophristic, V. Conformational Preferences of Furan- and Thiophene-Based Arylamides: a Combined Computational and Experimental Study. *Phys. Chem. Chem. Phys.* **2013**, *15*, 11883.
(13) Butterfoss, G. L.; Yoo, B.; Jaworski, J. N.; Chorny, I.; Dill, K. A.; Zuckermann, R. N.; Bonneau, R.; Kirshenbaum, K.; Voelz, V. A. De Novo Structure Prediction and Experimental Characterization of Folded Peptoid Oligomers. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 14320−14325.
(14) Voelz, V. A.; Dill, K. A.; Chorny, I. Peptoid Conformational Free Energy Landscapes from Implicit-Solvent Molecular Simulations in AMBER. *Biopolymers* **2010**, *96*, 639−650.
(15) Ruoslahti, E.; Pierschbacher, M. D. New Perspectives in Cell Adhesion: RGD and Integrins. *Science* **1987**, *238*, 491−497.
(16) Chatterjee, J.; Gilon, C.; Hoffman, A.; Kessler, H. N-Methylation of Peptides: A New Perspective in Medicinal Chemistry. *Acc. Chem. Res.* **2008**, *41*, 1331−1342.
(17) Dechantsreiter, M. A.; Planker, E.; Matha, B.; Lohof, E.; Holzemann, G.; Jonczyk, A.; Goodman, S. L.; Kessler, H. N-Methylated Cyclic RGD Peptides as Highly Active and Selective alpha(V)beta(3) Integrin Antagonists. *J. Med. Chem.* **1999**, *42*, 3033−3040.
(18) 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.
(19) Xiong, J. P. Crystal Structure of the Extracellular Segment of Integrin alpha Vbeta 3 in Complex with an Arg-Gly-Asp Ligand. *Science* **2002**, *296*, 151−155.
(20) Aumailley, M.; Gurrath, M.; Muller, G.; Calvete, J.; Timpl, R.; Kessler, H. Arg-Gly-Asp Constrained Within Cyclic Pentapeptides. Strong And Selective Inhibitors of Cell Adhesion to Vitronectin and Laminin Fragment P1. *FEBS Lett.* **1991**, *291*, 50−4.
(21) Bach, A.; Gross, J.; Jackson, S.; Markwalder, J.; Parthasarathy, A.; Wells, G.; Mousa, S.; DeGrado, W. NMR-Studies of Potent RGD-Containing Cyclic Peptide Inhibitors of GPIIb/IIIa. In *Abstracts of Papers: 205th ACS National Meeting*; American Chemical Society: Washington, DC, 1993; Vol. *205*; pp 112-MEDI.
(22) Gurrath, M.; Muller, G.; Kessler, H.; Aumailley, M.; Timpl, R. Conformation/Activity Studies of Rationally Designed Potent Anti-Adhesive RGD Peptides. *Eur. J. Biochem.* **1992**, *210*, 911−21.
(23) Haubner, R.; Gratias, R.; Diefenbach, B.; Goodman, S. L.; Jonczyk, A.; Kessler, H. Structural and Functional Aspects of Rgd-Containing Cyclic Pentapeptides as Highly Potent and Selective Integrin alpha(V)beta(3) Antagonists. *J. Am. Chem. Soc.* **1996**, *118*, 7461−7472.
(24) Haubner, R.; Schmitt, W.; Holzemann, G.; Goodman, S. L.; Jonczyk, A.; Kessler, H. Cyclic RGD Peptides Containing Beta-Turn Mimetics. *J. Am. Chem. Soc.* **1996**, *118*, 7881−7891.
(25) Wermuth, J.; Goodman, S. L.; Jonczyk, A.; Kessler, H. Stereoisomerism and Biological Activity of the Selective and Superactive alpha(V)beta(3) Integrin Inhibitor Cyclo(-Rgdfv-) and its Retro-Inverso Peptide. *J. Am. Chem. Soc.* **1997**, *119*, 1328−1335.
(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) 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.
(29) Onufriev, A.; Case, D. A.; Bashford, D. Effective Born Radii in the Generalized Born Approximation: The Importance of Being Perfect. *J. Comput. Chem.* **2002**, *23*, 1297−1304.
(30) Shell, M. S.; Ritterson, R.; Dill, K. A. A Test on Peptide Stability of AMBER Force Fields with Implicit Solvation. *J. Phys. Chem. B* **2008**, *112*, 6878−6886.
(31) Voelz, V. A.; Bowman, G. R.; Beauchamp, K. A.; 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.
(32) 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.
(33) 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.
(34) Killian, B. J.; Yundenfreund Kravitz, J.; Gilson, M. K. Extraction of Configurational Entropy from Molecular Simulations Via an Expansion Approximation. *J. Chem. Phys.* **2007**, *127*, 024107.
(35) Hensen, U.; Lange, O. F.; Grubmüller, H. Estimating Absolute Configurational Entropies of Macromolecules: The Minimally Coupled Subspace Approach. *PLoS One* **2010**, *5*, e9179.
(36) Polyansky, A. A.; Kuzmanic, A.; Hlevnjak, M.; Zagrovic, B. On the Contribution of Linear Correlations to Quasi-harmonic Conformational Entropy in Proteins. *J. Chem. Theory Comput.* **2012**, 120705154754007.
(37) Mierke, D. F.; Kurz, M.; Kessler, H. Peptide Flexibility and Calculations of an Ensemble of Molecules. *J. Am. Chem. Soc.* **1994**, *116*, 1042−1049.
(38) Voelz, V. A.; Zhou, G. Bayesian Inference of Conformational State Populations from Computational Models and Sparse Experimental Observables. *J. Comput. Chem.* **2014**, *35*, 2215−2224.
(39) Laufer, B.; Frank, A. O.; Chatterjee, J.; Neubauer, T.; Mas-Moruno, C.; Kummerlowe, G.; Kessler, H. The Impact of Amino Acid Side Chain Mutations in Conformational Design of Peptides and Proteins. *Chem.—Eur. J.* **2010**, *16*, 5385−5390.
(40) Chatterjee, J.; Mierke, D.; Kessler, H. N-Methylated Cyclic Pentaalanine Peptides as Template Structures. *J. Am. Chem. Soc.* **2006**, *128*, 15164−15172.
(41) 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.
(42) Maro, S. D.; Zizza, P.; Salvati, E.; Luca, V. D.; Capasso, C.; Fotticchia, I.; Pagano, B.; Marinelli, L.; Gilson, E.; Novellino, E.; Cosconati, S.; Biroccio, A. Shading the TRF2 Recruiting Function: A New Horizon in Drug Development. *J. Am. Chem. Soc.* **2014**, *136*, 16708−16711.
(43) Martin, S. F. Preorganization in Biological Systems: Are Conformational Constraints Worth the Energy? *Pure Appl. Chem.* **2007**, *79*, 193−200.
(44) Shi, Y.; Zhu, C. Z.; Martin, S. F.; Ren, P. Probing the Effect of Conformational Constraint on Phosphorylated Ligand Binding to an SH2 Domain Using Polarizable Force Field Simulations. *J. Phys. Chem. B* **2012**, *116*, 1716−1727.
(45) Shirts, M. R.; Mobley, D. L. An Introduction to Best Practices in Free Energy Calculations. *Methods Mol. Biol.* **2013**, *924*, 271−311.
(46) Chodera, J. D.; Mobley, D. L.; Shirts, M. R.; Dixon, R. W.; Branson, K.; Pande, V. S. Alchemical Free Energy Methods for Drug Discovery: Progress and Challenges. *Curr. Opin. Struct. Biol.* **2011**, *21*, 150−160.
(47) Mas-Moruno, C.; Beck, J. G.; Doedens, L.; Frank, A. O.; Marinelli, L.; Cosconati, S.; Novellino, E.; Kessler, H. Increasing alphaVbeta3 Selectivity of the Anti-Angiogenic Drug Cilengitide by N-Methylation. *Angew. Chem., Int. Ed.* **2011**, *50*, 9496−500.

DOI: 10.1021/ci500768u
J. Chem. Inf. Model. 2015, 55, 806−813