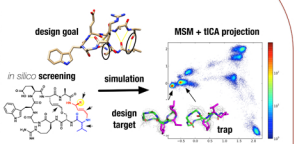
Computational screening and selection of cyclic peptide hairpin mimetics by molecular simulation and kinetic network models

Asghar M. Razavi , William M. Wuest and Vincent A. Voelz
Department of Chemistry, Temple University, Philadelphia, PA 19122, USA

Abstract

Designing peptidomimetic compounds to have pre-organized structure in solution is highly non-trivial. Simulation-based approaches can help speed this process. Here, we present an extensive simulation study of designed cyclic peptide mimics of a ß-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 twenty covalently cross-linked designs with varying stereochemistry, and selected the most favorable of these for massively parallel simulation on Folding@home in explicit solvent. Using Markov State Models (MSMs) we identified a key steric interaction between a methyl substituent and a valine sidechain that acts to allosterically shift population between *native* and *near-native* states, which could be exploited in experimental design rounds. Visualization of this mechanism is aided considerably by the tICA method, which identifies degrees of freedom most important in slow conformation transitions. The combination of quantitative detail and human comprehension provided by MSMs suggests such approaches will be increasingly useful for design.



Introduction

A **biofilm** is any group of microorganisms in which cells stick to each other on a surface.

Biofilm hazards: Dental plaques, Infectious kidney stones, Water pollution, etc.

Mechanism of biofilm growth: first Microorganisms stick to a surface, then they reproduce and cover more surface area until they reach a point that the surface is no longer large enough to hold all the microorganism. At this point the colony bursts and microorganisms propagate and cover new surface area.

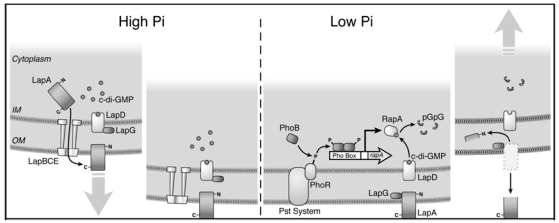
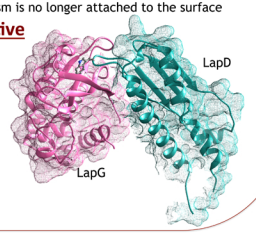


Figure 1. (From ref. 1) Details of surface adhesion of microorganisms: A protein called LapA binds to surfaces at high concentration of environmental phosphate. When phosphate is low, LapD-LapG complex breaks and now LapG is free to bind to LapA. Once LapG binds to LapA, it cleaves N-terminus of LapA and now microorganism is no longer attached to the surface

Objective

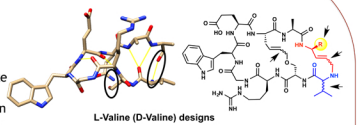
Long Term Goal:
Designing a hairpin mimetic to bind LapG, thus, preventing LapD-LapG complex formation.

Short Term Computational Goal:
Can we design a small hairpin molecule that is well structured in solution?



Results

Computational Design of Cyclic Hairpins. Hairpins with 7 residues (7-mer) or 9 residues (9-mer) are designed to mimic the binding hairpin of LapD in solution. To further stabilize the designed hairpins we introduced central and peripheral linkers between ß-sheets. By changing steriochemistry of Valine and introducing bulky groups in peripheral linker we also investigated effects of sterioselectivity on hairpin stability. The 9-mer designed (Fig. 1) show more native like behavior in solution than 7-mer designs and they selected for further analysis.



1 (9): ***E,Z*-9mer-H: R = H**
2 (10): ***E,Z*-9mer-H: R = CH_3**
3 (11): ***Z,E*-9mer-H: R = H**
4 (12): ***Z,E*-9mer-H: R = CH_3**
5 (13): ***E,E*-9mer-H: R = H**
6 (14): ***E,E*-9mer-H: R = CH_3**
7 (15): ***Z,Z*-9mer-H: R = H**
8 (16): ***Z,Z*-9mer-H: R = CH_3**

Figure 2. 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).

Fast initial screening using replica exchange molecular dynamics (REMD) simulations. REMD results reveal that designs **1** and **2** from L-Valine and designs **9** and **10** from D-Valine have highest native-like populations in the lowest RMSD region (Fig.2). Designs 1, 2, **9**, and 10 are then prepared for massive parallel explicit water simulations on Folding@Home.

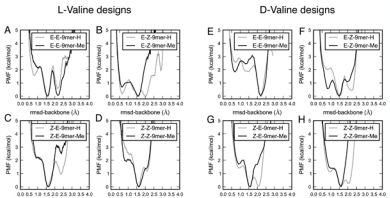
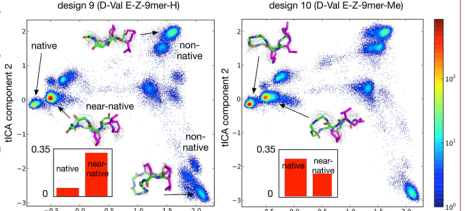


Figure 3. Potential of mean force vs. backbone RMSD from native structure for D-valine and L-valine designs.

Folding@Home simulation results confirm REMD results. Qualitative agreement between REMD and F@H results is obtained by comparing free energy of cyclic hairpin designs. This confirm reliability of implicit REMD simulations and the fact that designs **1**, **2**, **9**, and **10** are the most potent (see conclusion figure).

Markov State Models of Cyclic Hairpin Designs Reveal a Steric Mechanism for Stability of Methylated D-Val. Free energy landscape of designs **1**, **2**, **9**, and **10** show two distinct basins below 2 Å RMSD; one at 1.6 Å (*near-native*) and one at 0.7 Å (*native*). Design **10** (methylated D-Val.) shows higher stability in *native* basin compared to design **9** (unmethylated D-Val). To see conformational differences of *native* and *near-native* basins we projected the simulation data on tICA landscape (a newly developed dimensional reduction technique than identifies the slowest reaction coordinates in a system).



MSM built on tICA landscape is capable of distinguishing *native* and *near-native* states: the *native* state has Val. and methyl facing opposite from each other, while in the *near native* conformations, the Val. and methyl are facing close to each other, which causes steric clash between them and increases free energy of the *near-native* conformations. MSMs built on conventional free energy landscapes like RMSD-RG cannot identify *native* and *near-native* basins since there is no clear boundary between these states (Fig. 5). tICA components (eigenvectors of a time-lagged covariance matrix normalized with the covariance matrix) reveal the valine is playing a major role in controlling the kinetics of the system (Fig. 4) as well as governing the thermodynamics through allosteric interactions with the turn region (Fig. 7).

Results

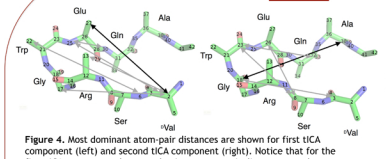


Figure 4. Most dominant atom-pair distances are shown for first tICA component (left) and second tICA component (right). Notice that for the first tICA component the most dominant atom-pair distances are the ones that connect D-val to turn region.

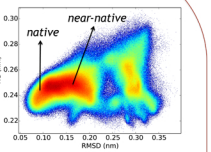


Figure 5. Root Mean Square Deviation (RMSD) versus Radius of Gyration (RG) for design **10** (methylated D-Val.). Notice that there is no clear boundary between *near-native* and *native* states. As a result MSM built from this landscape cannot distinguish these two structures.

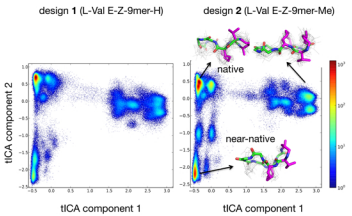


Figure 6. tICA landscapes for L-Val designs. Here introducing methyl group at the peripheral linker shifts population toward *near-native* state which has higher RMSD than *native* state.

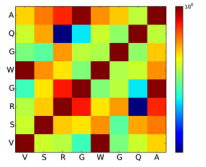
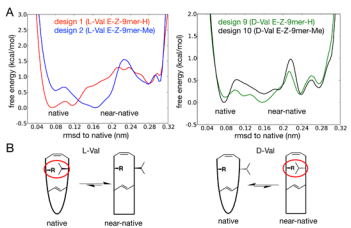


Figure 7. correlation matrix for psi dihedral angles of design **9**. Notice the high correlation between valine and tryptophan. This reveals existence of a direct allosteric interaction between peripheral and turn region in design **9**, which otherwise recognizing it is highly challenging to human comprehension.

Conclusion

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. 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.



One lesson from our work here is that the problem of rationally designing conformational constraints can be highly nontrivial. We showed here that how subtle changes in one end of a cyclic hairpin can allosterically effect conformation on the turn region leading to significant stabilization or destabilization of a target structure.

References

1. Peter D. Newell, Chelsea D. Boyd, Holger Sondermann, George A. O'Toole, A c-di-GMP Effector System Controls Cell Adhesion by Inside-Out Signaling and Surface Protein Cleavage. PLoS Biology, www.plosbiology.org , 2011, Volume 9, Issue 2, e1000587
2. 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
3. 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
4. Asghar M. Razavi, William M. Wuest, and Vincent A. Voelz, Computational Screening and Selection of Cyclic Peptide Hairpin Mimetics by Molecular Simulation and Kinetic Network Models. J. Chem. Inf. Model. 2014, 54, 1425-1432