Probing Peptide Folding with Molecular Dynamics Simulations

Hossein Mohammadiarani and Harish Vashisth

Department of Chemical Engineering, University of New Hampshire, Durham, NH

Abstract

Owing to noteworthy efforts in determining genome sequences for more than three decades, a huge database of gene sequences from different organisms has been developed. The difference between the number of protein structures and the number of protein sequences increases each year. However, only sequence information, without three-dimensional protein structures, is not sufficient for determining biological function of proteins. This point highlights the importance of knowing the three dimensional structure of protein and how a protein folds. The most accurate experimental methods for determination of protein structures are X-ray crystallography and NMR spectroscopy, which are labor-intensive methods. Molecular Dynamics (MD) simulation is a powerful computational approach that has been proposed for protein structure determination, and it does not have typical drawbacks of experimental methods. In this poster, I will present results from MD simulation of small peptide sequences with therapeutic application in diseases such as diabetes and cancer. In particular, small peptide systems of the insulin family were studied for their folding properties and thermodynamics in aqueous and membrane environments. The simulations show that the predicted fold of an insulin-mimetic peptide is highly similar to the known native structure of insulin. MD simulations of the transmembrane domain (TMD) of the insulin receptor in lipid membranes were conducted and the predicted structures are comparable to recent NMR studies. Overall, these results have implications for the design of novel peptide therapeutics targeting membrane receptors.

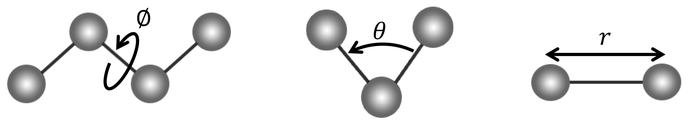
Methods and Software

NAMD (**NA**noscale **M**olecular **D**ynamics) is a parallel molecular dynamics code designed for high performance simulation of large systems of particles based on Newtonian equation of motion:

$$m_i\vec{\ddot{r}}_i = -\frac{\partial}{\partial \vec{r}}U_{total} \quad i = 1, 2, 3, \dots, N$$

$$U_{total} = U_{bond} + U_{angle} + U_{dihedral} + U_{vdw} + U_{coulomb}$$

$$U_{Coulomb} = \frac{q_1 q_2}{4\pi\varepsilon_0 r_{12}} \qquad U_{vdw} = 4\varepsilon_{12}\left[\left(\frac{\sigma_{12}}{r_{12}}\right)^{12} - \left(\frac{\sigma_{12}}{r_{12}}\right)^{6}\right]$$



$$U_{dihedral} = k_d(1 + \cos(n\emptyset - \gamma)) \qquad U_{angle} = k_a(\theta - \theta_0)^2 \qquad U_{bond} = k_b(r - r_0)^2$$

VMD (**V**isual **M**olecular **D**ynamics) is a molecular visualization software for displaying and analyzing MD simulations.

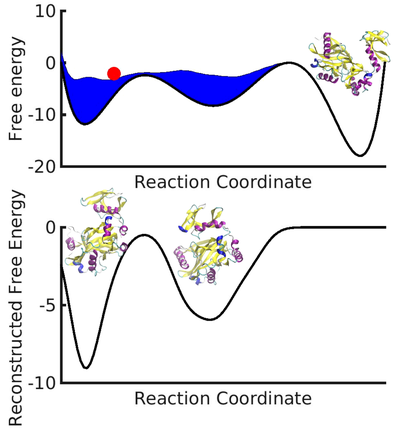
Metadynamics is a technique for enhanced exploration of the free-energy landscape of biomolecules.

Reconstructed free energy is given by:

$$W \sum_{t'=n\tau, t'<t} \exp\left(\frac{(S(x) - s(t'))^2}{2\delta s^2}\right)$$

which is a function of *S(x)* and *t*.

- *n* : *is an integer number*
- *S* : *current reaction coordinate*
- *s* : *all previous reaction coordinate*
- *W* : *the Gaussian height*
- *2δs: the Gaussian width*
- *τ* : *the frequency at which the Gaussians are added*



Insulin B-chain

Sequence alignment of human insulin B-chain with the insulin mimetic peptide. Residues which are conserved in the alignment between mimetic peptide and the B-chain of human insulin are highlighted.

Native Insulin B-chain:	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Mimetic peptide (S371):	G	S	L	D	E	S	F	Y	D	W	F	E	R	Q	L	G	K	K

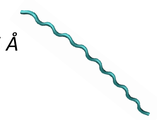
Simulations Parameters

Equilibration simulation:
- ✓Peptide immersed in water box (50 *Å* x 75 *Å* x 72 *Å*)
- ✓Neutralized in saline water (0.05 mol/L NaCl)
- ✓Equilibration time : 200 ps, time step : 2.0 fs
- ✓NPT ensemble with periodic boundary conditions
- ✓ Number of atoms: 25119 atoms (including hydrogen)
- ✓CHARMM 22 force field with CMAP correction

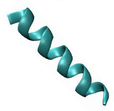
Metadynamics simulation:
- ✓Simulation time : 100 ns
- ✓Time step : 1.0 fs
- ✓Gaussian height : 0.1
- ✓Gaussian width : 0.2 Å
- ✓Gaussian frequency : 1 ps
- ✓Reaction coordinate grid spacing : 0.2 Å

Reaction coordinates ranges :

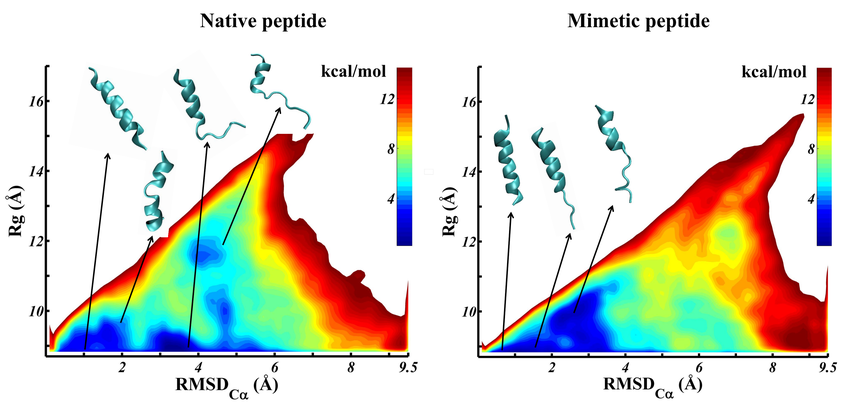
RMSD = 9.5 Å
Rg = 18.7 Å



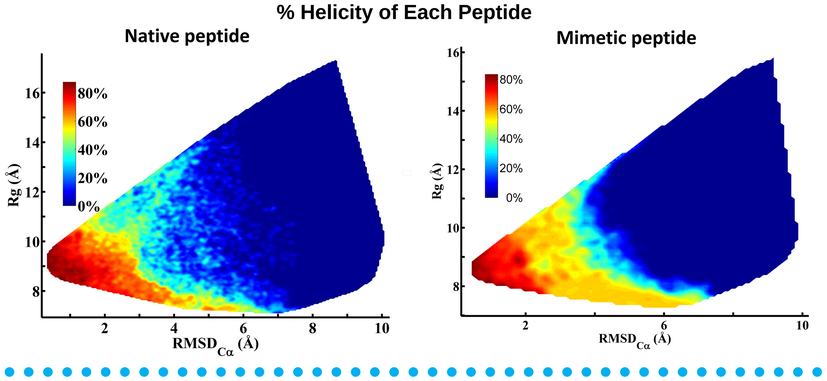
RMSD = 0 Å
Rg = 8.7 Å



Folding Thermodynamics (Potentials of Mean Force)



% Helicity of Each Peptide



Sequence Alignment of Receptor Transmembrane Domain

Human insulin receptor (IR):
PTYFYVTDYLDVPSNIAKIIIGPLIFVFLFSVVIGSIYLFLRKRQPDGPLG
Human type-1 insulin-like growth factor receptor (IGF1R):
DPVFFYVQAKTGYENFIHLIIALPVAVLLIVGGLVIMLYVFHRKRNNSRLG

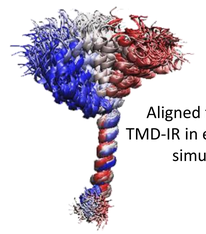
Hydrophobic
Polar or Charged
Proline

Proline residues induce a kink in the peptide structure.

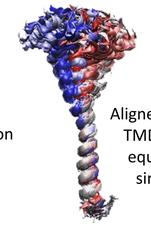
MD Equilibration of Initial Structure

Equilibration simulation parameters:
- ✓IR immersed in membrane (80 Å x 80 Å)
- ✓Neutralized in saline water (0.05 mol/L KCl)
- ✓NPT ensemble with periodic boundary conditions
- ✓ Number of atoms: ~75000 atoms
- ✓CHARMM 22 force filed with CMAP correction

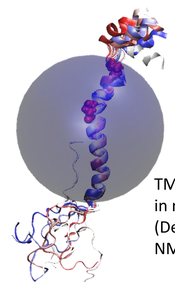
Equilibration steps:
1- 0.5 ns, Water and protein are frozen
2- 5 ns, Protein is frozen
3- 50 ns, all molecules are relaxed



Aligned frames of TMD-IR in equilibration simulation



Aligned frames of TMD-IGF1R in equilibration simulation

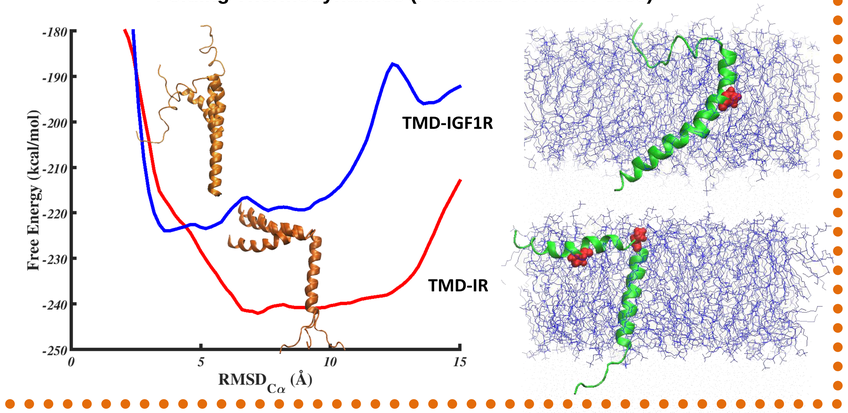


TMD-IR structure in micelle (2MFR) (Determined by NMR)

Conformational Sampling of Membrane Peptide with Metadynamics

Metadynamics Parameters:
- ✓Time step : 1.0 fs
- ✓Reaction coordinate: 0 < RMSD < 15 Å
- ✓Reaction coordinate grid spacing : 0.2 Å
- ✓Simulation time ~ 160 ns
- ✓Gaussian height : 0.1
- ✓Gaussian width : 0.2 Å
- ✓Gaussian frequency : 1 ps

Folding Thermodynamics (Potential of Mean Force)



References and Acknowledgements

1. L. Schaffer, et al., 2003, PNAS, 100:4435–4439.
2. J. Henin, et al., 2010, J. Chem. Theory Comp., 6:35–47.
3. J. C. Phillips, et al., 2005, J. Comp. Chem., 26:1781-1802.
4. M. K. Baker, et al, 2014, Biochim. Biophys. Acta, 1838:1396–1405.
5. C.W. Ward, et al., 2009, BioEssays, 31:422–434.
6. Q. Li, et al, 2014, Biochim. Biophys. Acta, 1838:1313–1321.
7. W. Humphrey, et al., 1996, J. Mol. Graphics, 14:33-38.

Computations were performed on Trillian, a Cray XE6m-200 supercomputer at UNH supported by the NSF-MRI program under grant PHY-1229408.
Summer TA Fellowship (Mohammadiarani; UNH Graduate School)
Summer Faculty Fellowship (Vashisth; UNH Graduate School)