Mapping and Visualizing the Free Energy Landscape of p53 using Vectors

Nabeel Kemal[1], Kelly M Thayer[1,2]

1. Department of Mathematics and Computer Science 2. Department of Chemistry

Introduction

In allosteric regulation, the energy landscape of a protein, a statistical representation of a proteins potential energy, can be altered to achieve various results. Understanding allostery provides exciting insight into the possibilities of more targeted and effective drugs made using allosteric design. To better understand the shifts in the free energy landscape that occur as a result of allosteric effectors, we can use vectors, programmed with magnitude and angle data from simulation outputs, and couple them with residues. This allows for an instantaneous visual representation of how a protein reacts to various allosteric effectors in a simulation. This visualization allows for continuous refinements in order to create allosteric effectors that most favorably shift the free energy landscape, with the eventual goal of engineering allosteric effectors to reactivate native functionality in proteins. Each vector will be centered on alpha carbon of a residue, using a polar coordinate system the vector will represent the net force felt by its respective residue. This net force on a residue is a product of its neighboring residues inflicting external forces upon it.

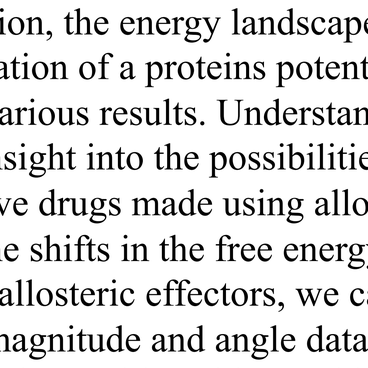

Figure 1: Example of a Thayer Vector, uses a local polar coordinate system centered on each residue's alpha carbon. Pictured with three neighboring residues *i,j,k*

The feasibility of this approach, in which we capture the free energy landscape using vectors will be demonstrated using the p53 tumor suppression protein. Mutations in p53 are present in nearly 50% of ovarian, esophageal, colorectal, head and neck, larynx, and lung cancers in humans[1]. Most of which are both lethal and undruggable, however, if this method proves to succeed in providing the missing link between identifying allosteric control points and which protein substate will be selected, it will be an integral step toward a new class of allosteric drugs with targeted control of the biological processes in any protein.

Acknowledgments

This research was funded by Grants NIH R15 GM128102-01, CNS-0619508, and CNS-095085 Thank you to my colleagues Prof. David Beveridge, Prof. Michael Weir as well as lab members Rujun Yan, Ben Cowan, Randy Tyng, Chunyue Ma, Abdelrahman Elsayed, Anastasia Saar, and Carol Delgarno for their helpful insight and thoughtful discussion. Many thanks as well to Hank Meij for maintaining Wesleyan's HPCC and finally thanks to the Wesleyan Summer Research Program.

References

1. Olivier, Magali et al. "TP53 mutations in human cancers: origins, consequences, and clinical use." *Cold Spring Harbor perspectives in biology* vol. 2,1 (2010)

Methods

Two programs were coded in order to create and implement vectors to capture the free energy landscape of p53. The first of which parses the output data of AMBER Molecular Dynamics (MD) simulations using cpptraj and extracts magnitude and angle data. The script is stored and operated on the HPCC (High Performance Computing Cluster) and is coded in bash.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	#Frame	p538_pep[bond]	p538_pep[angle]	p538_pep[dih]	p538_pep[vdw14]	p538_pep[elec14]	p538_pep[vdw]	p538_pep[elec]	p538_pep[total]					
2	1	6.7479	15.6217	24.7386	5.2769	15.6948	-3.0423	-90.8041	-25.7665					
3	2	40897.9951	2252.9596	87.2133	40701.1663	-50.9442	72270.8277	-33.8690	156125.3489					
4	3	59942.2332	1939.2856	81.1410	364429.5075	-50.6506	24013.6564	-57.0811	450298.0920					
5	4	47545.0194	2283.3472	70.4327	1771.1226	-89.5301	6154556.5080	-37.1547	6206099.7451					
6	5	38153.4562	2763.9278	101.9633	2860542.4699	-139.7331	32868694.6193	-137.4137	35769979.2897					
7	6	39236.0583	2447.8939	86.2725	6997.9165	-30.5999	2019109216.1424	-139.5449	2019157814.1388					
8	7	62318.0720	3032.0620	85.0139	4143227.5499	-58.0144	9537.8397	-104.6547	4218037.8685					
9	8	2656349.6292	3084.4650	69.3460	19135.8920	-1.1083	381301.7986	-72.4825	3059867.5400					
10	9	50273.2794	2918.3157	92.4352	12162.7284	-75.1862	17604326.3954	85.4245	17669783.3924					
11	10	50430.5399	3198.9328	91.3264	95687.5797	-29.3312	438251.2547	-71.4858	587558.8164					
12	11	52804.9640	2629.7348	76.2685	7504.2612	-43.1307	10696.1861	-167.4540	73500.8299					

Figure 2: Example output data from AMBER MD simulation using cpptraj

The second program is a python script that takes in the magnitude and angle data from the AMBER MD simulation as an input and outputs a vector for every frame, which represents the instantaneous magnitude and angle of the force felt by the residue.

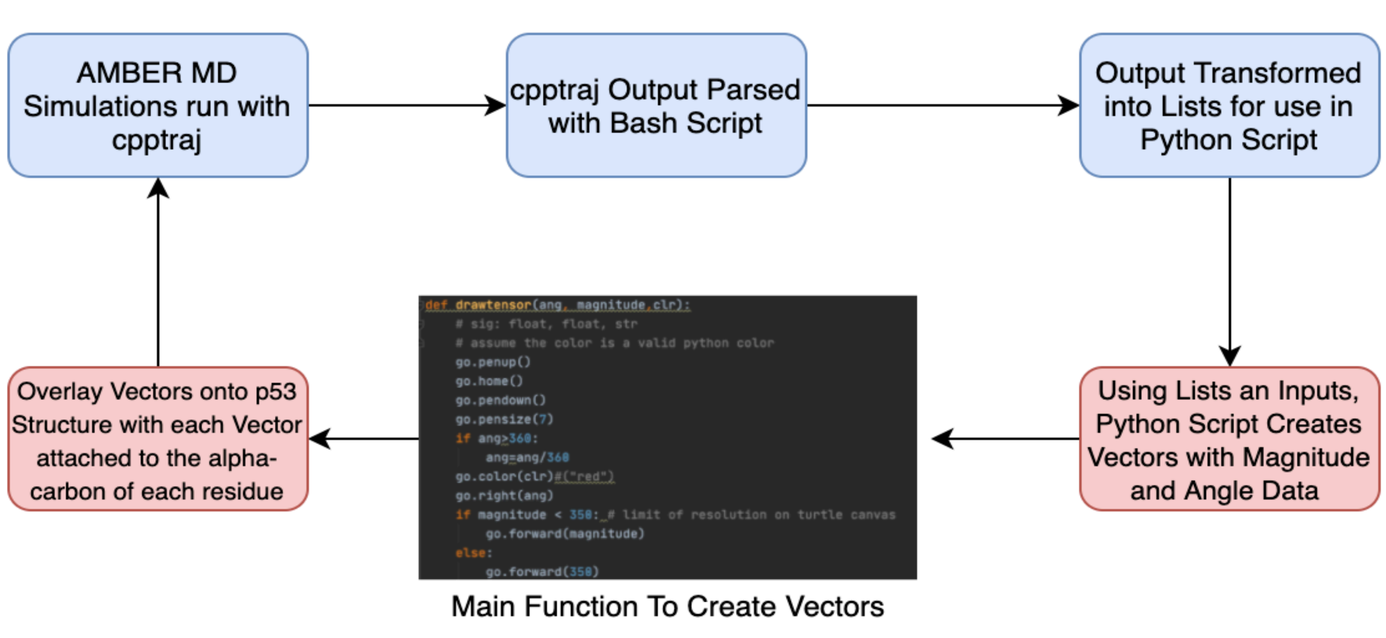

Figure 4: Flowchart detailing free energy landscape capture process as well as use of Python Script to create vector representations

Results

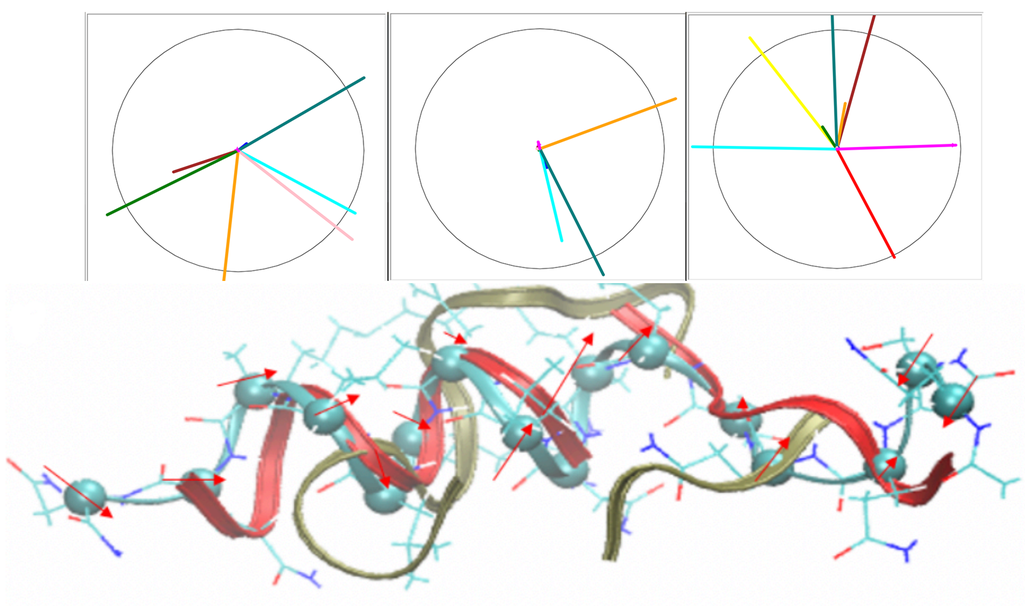

Figure 5, 6: (5) Output of Python script, an 11-frame vector representation of the free energy landscape, made using Turtle in Python (6) Output vectors mapped onto SAH-p53-8 starting structure, pointing towards the centroid 0 structure (red)

Future Direction

- Use Machine Learning to create a program that uses the vector data to predict peptide conformational substate selection
- Create automated visual overlay of vectors onto targeted molecules, with each vector centered at the alpha carbon of each residue
- Extend method beyond p53 and attempt on additional proteins