Vicky Hwang

Final Research Report for PRIME

Introduction:

Alzheimer's Disease affects more than 20 million people worldwide and has no known cure or treatment. One protein that may contribute to the pathology of Alzheimer's Disease is the dual-specificity phosphatase (DSP) called slingshot-homolog 2 (SSH-2).[1] It dephosphorylates and activates cofilin, which then causes actin filaments to form rod-shaped bundles. These rods lead to the accumulation of amyloid precursor protein, which results in the plaques that characterize Alzheimer's disease. Inhibiting the activity of SSH-2 may thus prevent cofilin from forming the rods that contribute to the pathology of Alzheimer's disease. However, compounds that inhibit SSH-2 may also affect other members of the DSP family, causing unwanted side effects.

Over the course of our project, we investigated potential inhibitors of SSH-2 to ensure that they do not bind well to other DSPs. Since testing every single inhibitor from a list of about 20,000 compounds against every single DSP protein would be time-consuming, scientists have developed "virtual screening" computer programs to simulate the interactions between proteins and ligands, docking each potential inhibitor onto the active site of the protein. In addition, the use of grid computing shares hardware resources to allow for multiple docking algorithms to be run at the same time, expediting the process.[2]

Our objective was to use DOCK6 to determine the binding interactions between the predicted inhibitors of SSH-2 and the closely related DSP proteins whose structures were predicted and modeled by the MODELLER program.

Methods:

The first step was to prioritize the proteins to be docked based on how similar the DSP was to SSH-2 as well as how accurate and stable the models were. ClustalW 2.1 was used to perform sequence alignment.[3] To optimize the protein models, UCSF Chimera was used to minimize energy, then individual loops near the active site of the protein were refined using the loopmodel of MODELLER. The resultant protein models were analyzed for stability using the online program MolProbity.

The protein was then prepared for docking by generating spheres and a box around the active site, where the protein-ligand interaction would occur. DOCK6 was used to "dock" each of the compounds onto the active sites of the modeled DSP proteins and predict their binding interaction. A grid system containing such values as the score for partial charges and Van der Waals forces was calculated for the active site. The ligands were then matched to each of the spheres on the grid system and grid scores were generated based on how well the ligands fit. In physiological conditions, the receptor often changes shape to better fit the ligand in a process known as "induced fit," but the grid-based scoring does not account for this. Another scoring system called AMBER can simulate a changing shape for the receptor and calculate a binding score, but is more time-intensive. The grid scores and AMBER scores were ranked and averaged to obtain a consensus score, determining how well the different ligands bind to the active site of the modeled proteins from the DSP family.

Results:

In the consensus file, the greater the difference in rank, the less likely the SSH-2 inhibitor would also inhibit that DSP, and the more specific and better that inhibitor would be. SSH-1 and SSH-3 were modeled and had complete scores for both grid-based (dock) and AMBER scoring. DSP 19 and DSP 21 have dock results only. Many of the top 10 inhibitors of SSH-2 also appear

to bind well to DSP 19, even better than they bind to SSH-1 or SSH-3. Of the top 10 inhibitors of SSH-2, the three that seem most promising are 1, 4, and 6, which seem to bind poorly to the DSPs we tested in that they yield a greater difference from the SSH-2 rank, especially for DSP 19.

zinc id	ssh2 rank	ssh1 rank	difference	ssh3 rank	difference	dsp21 rank (dock only)	difference	dsp19 rank (dock only)	difference
ZINC05260817	1	7289	-7288	5959	-5958	17661	-17660	433	-432
ZINC03869281	2	4039	-4037	1340	-1338	7378	-7376	39	-37
ZINC04543673	3	745	-742	510	-507	4534	-4531	121	-118
ZINC02384698	4	524	-520	609	-605	5217	-5213	956	-952
ZINC03869935	5	5349	-5344	3653	-3648	16822	-16817	25	-20
ZINC04521532	6	8330	-8324	6928	-6922	3676	-3670	737	-731
ZINC04543675	7	665	-658	1233	-1226	5716	-5709	134	-127
ZINC02522549	8	1228	-1220	628	-620	409	-401	135	-127
ZINC04652516	9	44	-35	538	-529	1179	-1170	14381	-14372
ZINC02637978	10	303	-293	364	-354	105	-95	3556	-3546

Figure 1: Consensus results from grid-based scoring and/or AMBER for the top 10 inhibitors of SSH-2.

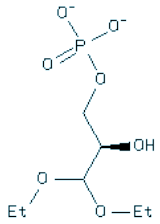

Figure 2: ZINC02384698, [(2R)-3,3-diethoxy-2-hydroxy-propyl]-phosphate, SSH-2 rank 4.

Discussion:

The modeled structures of SSH-1 and SSH-3 did not include parts of the protein that were from far the active site, since lack of data from crystallized proteins prevented an accurate prediction of the entire protein's structure. Therefore, the difference in binding of the ligands may be larger than what the results would indicate. Results from this project may be used as a benchmark to evaluate the accuracy of the protein structures predicted by the MODELLER program. The sequences of SSH-1 and SSH-3 were quite similar to that of SSH-2, with SSH-3

differing slightly more from the other two, which agrees with a previous study that found SSH-1 and SSH-2 to have very similar properties while SSH-3 had lower phosphatase activity.[4] However, the dock results indicated that some of the top ten inhibitors for each protein varied rather considerably, while SSH-3 often did not have a greater difference from SSH-2 rank as compared to SSH-1. Furthermore, DSP 19 has a greater difference in sequence from SSH-2 based on results from ClustalW protein alignment, yet appeared to bind very well to the SSH-2 inhibitors, often better than the more similar SSH-1 or SSH-3. This may indicate that MODELLER did not make very accurate predictions for the protein structures, or that the slight difference in sequences, even in residues far from the active site, made a significant difference in how the ligands bound to the active site.

Challenges encountered:

Many technical issues were encountered during the course of the project. One problem was due to the jobs being run using the extremely resource-intensive service Tomcat, which was only intended for use by one person. Consequently, the entire cluster would occasionally crash when everyone attempted to run a job. Another problem was the limited number of nodes and clusters available, which reduced the number of jobs that could be run. We solved both problems by having only one person run the jobs while the other two prepared the ligands for docking or looked for more clusters to use.

Other technical problems were solved by writing scripts. AMBER is prone to errors when preparing the ligands to dock, and in previous years, the solution to this was to manually find the molecule that caused the error and remove it. The process was tedious and often error-prone, so a script was written and successfully implemented this year to automate the process. Another

useful script developed this year automated the process of compressing and labeling the results generated by both grid-based and AMBER scoring.

Future directions:

Our findings suggest three promising ligands that may be specific to SSH-2 and leave the other DSPs we docked unaffected. These compounds will need to be tested in actual wet lab experiments to determine if they can specifically inhibit SSH-2 in cells. Furthermore, our results are still incomplete; we have data from grid-based energy scoring for DSP-19 and DSP-21, but not AMBER scoring. Other members of the DSP family also still need to be docked. Possible future virtual screening projects may dock ligands built from recombining parts from known inhibitors in order to achieve an optimal inhibitor. Eventually, a drug may be developed that will specifically inhibit SSH-2 and cofilin activation, allowing for the treatment of Alzheimer's disease.

References:

1. Maloney MT, Bamburg JR. “Cofilin-mediated neurodegeneration in Alzheimer's disease and other amyloidopathies”. Mol Neurobiol 2007; 35:21–44.

2. M. J. Levesque, K. Ichikawa, S. Date, and J. H. Haga. "Design of a grid service-based platform for in silico protein-ligand screenings". Computer Methods and Programs in Biomedicine 2009; 93:73-82.

3. Larkin MA, Blackshields G, Brown NP, et al. “ClustalW and ClustalX version 2”. Bioinformatics 2007; 23(21): 2947-2948.

4. Ohta Y, Kousaka K, Nagata-Ohashi K, et al. “Differential activities, subcellular distribution and tissue expression patterns of three members of Slingshot family phosphatases that dephosphorylate cofilin”. Genes Cells 2003; 8:811–824.