Project Title:

Computational structure-based design of protein inhibitors

Name:○Kam Zhang, Arnout Voet , Ashutosh Kumar , Kamlesh Sahu , Muhammad Muddassar , Tom Venken, Xiaoyin Lee

Laboratory:

Zhang Initiative Research Unit, RIKEN Advanced Science Institute, RIKEN Wako Institute

1. Background and purpose of the project, relationship of the project with other projects

Fragment based drug design has developed significantly over the past decade and is now recognized as successful method of hit identification and lead generation. The method was proposed in 1996 and has gained more and more interest in following years to finally become a tangible alternative to high throughput screening. Fragment based drug design focuses on low molecular weight compounds that target sub-pockets within the overall active site. It samples chemical space more efficiently because the number of fragments needed to cover all reasonable chemotypes is many orders of magnitude smaller than that for more drug-like compounds, where multiple chemotypes are combined. Fragment hits are expected to be more suitable starting point for hit to lead optimization due to their reduced complexity, which leaves more freedom for multidimensional property optimization of the fragment hits, usually by addition of new chemical functions or alternatively by linking of two fragment hits binding in adjacent pockets.

In the past, this strategy has led to the discovery of new scaffolds that were later combined or grown into high affinity inhibitors. However, the small size of fragments and low binding affinity makes it particularly difficult to detect them in standard biochemical assays. Instead biophysical methods such as NMR, protein crystallography and surface Plasmon resonance are used to test up to hundreds or several thousands of compounds. These methods additionally benefit from the fact that they yield structural information about fragment binding poses but despite their utility, there are significant time, labor, and materials costs associated with biophysical screening. Also, considering that estimated size of chemically relevant fragment space is in the order of 10^8 and the number of commercially available fragments is over 300000, the throughput and the associated cost limit the application of biophysical screening.

Structure based virtual screening of drug like libraries has already been proved to be an efficient technology in hit discovery. Furthermore structure based virtual screening was integrated with experimental screening providing drug like hits for large variety of targets. Based on positive experiences one can conclude structure based virtual screening could also support fragment based drug design by molecular docking and prioritizing fragments using various ranking and scoring methodologies such as molecular dynamics simulations. Despite broad application of structure based virtual approaches, *in silico* fragment screening still plays a minor role. Although there have been some case studies reported that demonstrate virtual screening to be a powerful tool to discover potent fragments, this approach is still often considered to be too unreliable using popular computer tools. Indeed, facing the shortcomings of computational protocols with respect to ranking and affinity prediction of small fragments, more elaborate protocols have to be evolved departing from routine virtual screening strategies.

The objective of this project is to develop a virtual screening pipeline integrating structure based virtual screening approaches and utilize them

in the rational identification of low molecular weight compounds with weak binding affinity towards a variety of therapeutically relevant protein targets. A key theme is the development and utilization of computational approaches that incorporate protein flexibility considerations into structure based discovery by using flexible docking and molecular dynamics simulation to identify desired fragments can be later combined or grown into high-affinity inhibitors.

In this project, we proposed to develop a virtual screening protocol and to use it for the discovery of inhibitors of a protein lysine methyltransferase G9a and Sumoylation E1 and E2 enzymes as well as SUMO:SBM interactions. However, the approach is inherently generic and can be applied to many fragment based drug discovery programs for disease like cancer, inflammation, neurodegenerative diseases.

Histone methylation has important roles in regulating transcription, genome integrity and epigenetic inheritance. G9a is a protein lysine methyltransferase that catalyzes methylation of lysine 9 of histone 3 (H3K9) and lysine 373 of tumor suppressor p53. G9a inihibition is an attractive therapeutic strategy for cancer as G9a is over expressed in human cancers and its knockdown inhibits cancer cell growth.

Sumoylation is a post-translational modification that plays an important role in a wide range of cellular processes including DNA replication and repair, chromosome packing and dynamics, genome integrity, nuclear transport, signal transduction and cell proliferation. Among the proteins involved in sumoylation pathway, Ubc9 is the sole E2-conjugating enzyme required for sumoylation and plays a central role by interacting with almost all the partners required for sumoylation. Ubc9 has been implicated in a variety of human malignancies such as ovarian carcinoma, melanoma and lung adenocarcinoma, suggesting that Ubc9 inhibition could be a potential therapeutic approach to control tumorigenesis.

2. Specific usage status of the system and calculation method

The basic goal of our virtual screening protocol is to computationally screen millions of commercially available small molecules against a specific target protein to prioritize small number of compounds for biological testing and for their further development to high affinity inhibitors. However, *in silico* structure based virtual screening requires intensive computing especially in case of large database with millions of chemical compounds, which is in the range of few Tflops per day on one target protein. Our virtual screening protocol involves the screening of small molecule library using pharmacophore based modeling, flexible molecular docking and molecular dynamics simulation.

For pharmacophore modeling, we have used LigandScout, MOE software suite. We have also used molecular shape and electrostatic potential based criteria to search for initial hit using the ROCS and EON program. For flexible docking, we have used RosettaLigand, Glide and GOLD program. The hits which ranked higher were further prioritized by molecular dynamics simulation based binding free energy calculations. AMBER program and MM-GB(PB)SA approach was employed for this purpose. Our procedure incorporate full protein and ligand flexibility which greatly increases the computation time due to the vast number of conformations need to be explored; however, this allows more accurate treatment of protein ligand interactions.

3. Result

We have used virtual screening techniques such as molecular docking, molecular dynamics simulation and binding free energy calculations to identify potential small molecule inhibitors of SUMO E1. About 24 hits were acquired and tested using *in situ* sumoylation assay. Out of 24 hits, four hits

showed more than 85 % inhibition of sumoylation with the most active compound showed an IC_{50} of 5.4 μM. To improve the potency of these compounds, we have carried out similarity search to identify commercially available derivatives. About 50 derivatives were identified to be tested further. Compounds with improved potency were found in each of four chemical classes. These compounds will be excellent starting points for further chemical optimization.

We used a hybrid structure-based virtual screening protocol that incorporates both ligand-based and structure-based techniques to identify inhibitors of Ubc9. Nineteen compounds were acquired from different chemical vendors and were tested first using in situ sumoylation assay and then by using E1 and E2 SUMO intermediate formation assay. About five compounds showed inhibitory activity against SUMO E2 out of which one compound was selected for further optimization. The similarity search to retrieve commercially available derivatives resulted in 40 compounds that have been acquired and tested for Ubc9 inhibition. Four derivative compounds were found to possess improved activity over the parent compound.

Pharmacophore models were developed using the published SAR of G9a inhibitors. About 1 million compounds from databases were docked onto G9a with and without co-crystallized water molecules. Molecular dynamics and binding energy calculations were performed using crystal structure of G9a (3K5K) in complex with docked inhibitors to establish correlation between published IC_{50} values and calculated binding free energies. The correlation has improved by selecting post filtered poses (using pharmacophore) for MD and by using co-crystallized water as co-factor. Several weak inhibitors have been confirmed by experimental assay. Further optimization of those inhibitors are underway.

In order to identify small molecules that might have the possibility to break SUMO:SBM interaction, a funnel based virtual screening approach was followed. First, the structure was analyzed and because of the abundant presence of K and R amino acids, all molecules should be negatively charged or have more hydrogen bond acceptor than donor atoms. Next, a pharmacophore query was created based on the consensus interactions of the SBM with the SUMO1 protein. The pharmacophore was used as a post processing filter for all docked poses. Those that do not fulfill the key interactions within the receptor were discarded. Lastly, only the molecules that were able to bind to the different receptor conformations in exactly the same binding mode were retained. Out of these remaining compounds, the most promising compounds were selected based on a consensus of the PLP and the DrugScore scoring functions of the docked results as well as divergence and lead-likeness. These compounds have been selected for purchase in order to be tested for inhibition of different SUMO-SBM interactions. The ALPHA-screen assay was used to measure the ability of those compounds to inhibitor SUMO1:SBM interaction and 11 out of the 64 compounds purchased were found to be active (Fig. 1). Further optimization is underway. The binding mode of these active compounds to SUMO1 has been confirmed by NMR STD and HSQC experiments.

4. Conclusion

Computational methods, which include pharmacophore, shape and electrostatics, docking and molecular dynamics, were used to screen a large commercial compound database for inhibitors of the sumoylation enzymes E1 & E2 and the protein lysine methyl transferase G9a. We have targeted both enzyme catalytic sites and protein-protein interfaces for inhibitor discovery. For SUMO E1, four chemical series have been identified and the most active compound showed an IC_{50} of 5.4 μM. Chemical optimization of these inhibitors is underway. For SUMO E2, we have identified five chemical series that showed inhibitory activities and the most potent compounds have IC_{50} of 46 μM. We have identified a

site in SUMO:SBM that can be targeted by small molecule inhibitors. Using a pharmacophore modeling approach, 40 compounds were selected for purchasing and assaying. After several rounds of optimization by analogs, several small molecule inhibitors of SUMO:SBM interaction with IC_{50} in the single digit μM have been identified. We have used a consensus based approach to identify G9a inhibitors. Two compounds with IC_{50} ~100 μM have been identified. These provide good starting points for further optimization.

5. Schedule and prospect for the future

We plan to optimize the initial hits for sumoylation enzyme E1 and E2 that have been identified by virtual screen and confirmed by biochemical assays. Various computation tools will be used to optimize these initial hits into more potent inhibitors. The computation time in RICC is critical for us to achieve these goals.

We also plan to optimize those small molecule protein-protein interaction inhibitors (SMPPII) of the sumoylation pathway that we have discovered thus far. SMPPIIs have been considered as high-hanging fruits in drug discovery. They offer a clear advantage over traditional enzymatic site inhibitors. We have established a pharmacophore based protocol for the discovery of SMPIIs. We will use the RICC time to optimize those SMPPIIs for sumoylation. These inhibitors can potentially be developed into drugs for treating various diseases such like cancer.

6. If no job was executed, specify the reason.

N/A.

RICC Usage Report for Fiscal Year 2011

Fiscal Year 2012 List of Publications Resulting from the Use of RICC

[Publication]

1. Voet, A., Banwell, E. F., Sahu, K. K., Heddle, J. G., Zhang, K. Y. J. (2013) Protein interface pharmacophore mapping tools for small molecule protein:protein interaction inhibitor discovery. *Curr. Med. Chem.*, in the press.
2. Tsuchiya, A., Asanuma, M., Hirai, G., Oonuma, K., Muddassar, M., Nishizawa, E., Koyama, Y., Otani, Y., Zhang, K. Y. J., Sodeoka, M. (2013) CDC25A-inhibitory RE Derivatives Bind to Pocket Adjacent to the Catalytic Site. *Mol. BioSyst.*, DOI:10.1039/C3MB00003F
3. Kumar, A., Zhang, K. Y. J. (2013) Computational investigation of SENP:SUMO protein-protein interaction for structure based drug design. *Molecular Informatics*, DOI:10.1002/minf.201200124.
4. Voet, A., Helsen, C., Zhang, K. Y. J., Claessens, F. (2013) The discovery of novel hAR receptor antagonist chemotypes using a combined pharmacophore screening procedure. *ChemMedChem.*, DOI:10.1002/cmdc.201200549.
5. Jose, R.A., Voet, A., Broos, K., Jakobi, A.J., Bruylants, G., Egle, B., Zhang, K. Y. J., De Maeyer, M., Deckmyn, H., De Borggraeve, W. M. (2012) An integrated fragment based screening approach for the discovery of small molecule modulators of the VWF-GPIbα interaction. *Chem. Commun.*, **48**, 11349-11351.
6. Kumar, A., Voet, A., Zhang, K. Y. J. (2012) Fragment based drug design: from experimental to computational approaches. *Curr. Med. Chem.*, **19**, 5128-5147.
7. Voet, A., Zhang, K. Y. J. (2012) Pharmacophore modeling as a virtual screening tool for the discovery of small molecule protein-protein interaction inhibitors. *Current Pharmaceutical Design*, **18**, 4586-4598.
8. Mishra, V., Kumar, A., Ali, V., Nozaki, T., Zhang, K.Y.J., Bhakuni, V. (2012) Novel protein-protein interactions between Entamoeba histolyticad-phosphoglycerate dehydrogenase and phosphoserine aminotransferase. *Biochime*, **94**, 1676-1686. doi:10.1016/j.biochi.2012.02.028

[Proceedings, etc.]

None.

[Oral presentation at an international symposium]

1. Bio-IT World Asia Conference, June 6 - 8, 2012, Singapore. Invited Speaker, “Scaffold-based drug design: an efficient tool for the discovery of new molecular entities”.
2. 4th International Conference on Drug Discovery and Therapy, Feb. 12-15, 2012, Dubai, UAE. Invited Speaker, “Scaffold-based drug design: an efficient tool for the discovery of new molecular entities”.

3. EFMC-ISMC 2012, 22nd International Symposium on Medicinal Chemistry, Sept. 2 - 6, 2012, Berlin, Germany, Poster presentation, Ashutosh Kumar, Akihiro Ito, Minoru Yoshida and Kam Y. J. Zhang, *"In silico* Screening of Inhibitors for Sumoylation Enzyme Ubc9".
4. Molecular Modeling meeting, Aug. 30-Sept. 1, 2012, Queenstown, New Zealand, Poster presentation, Ashutosh Kumar, Mikako Hirohama, Akihiro Ito, Minoru Yoshida and Kam Y. J. Zhang, "Structure based virtual screening to identify inhibitors of sumoylation activating enzyme E1".
5. Molecular Modeling meeting, Aug. 30-Sept. 1, 2012, Queenstown, New Zealand, Poster presentation, Arnout RD Voet, Francois Berenger, Kam YJ Zhang, "Analysis and application of electrostatic similarities between small molecules and protein ligands for improved SMPPII design using *Elekit"*.
6. QMB Chemical Biology and Drug Discovery Meeting, Aug. 26-27, 2012, Queenstown, New Zealand, Poster presentation, Ashutosh Kumar, Mikako Hirohama, Akihiro Ito, Minoru Yoshida and Kam Y. J. Zhang, "*In silico* screening to identify inhibitors for sumoylation enzyme Ubc9".
7. QMB Chemical Biology and Drug Discovery Meeting, Aug. 26-27, 2012, Queenstown, New Zealand, Poster presentation, Arnout RD Voet, Mikako Hirohama, Akehiro Ito, Minoru Yoshida and Kam YJ Zhang, "Rational design of the first non-peptidic inhibitors of the SUMO1:SBM interaction".
8. JCUP-III, June 7-8, 2012, Tokyo, Japan, Poster presentation, Ashutosh Kumar, Mikako Hirohama, Akihiro Ito, Minoru Yoshida and Kam Y. J. Zhang, "Structure based virtual screening to identify inhibitors of Sumoylation activating enzyme E1".
9. JCUP-III, June 7-8, 2012, Tokyo, Japan, Poster presentation, Muhammad Muddassar and Kam Y. J. Zhang, "Virtual screening of PI3K inhibitors using hybrid approaches".
10. JCUP-III, June 7-8, 2012, Tokyo, Japan, Poster presentation, Kamlesh Sahu, Arnout Voet, Minoru Yoshida and Kam Y. J. Zhang, "Consensus based virtual screening to identify potential G9a-inhibitors".