6-9-2017

Towards a Better Understanding of On and Off Target Effects of the Lymphocyte-Specific Kinase LCK for the Development of Novel and Safer Pharmaceuticals

Xiaofei Zhang
University of Kentucky, xiaofei.zhang@uky.edu

Amir Kucharski
University of Kentucky

Wibe A. de Jong
Lawrence Berkeley National Laboratory

Sally R. Ellingson
University of Kentucky, sally.ellingson@uky.edu

Towards a Better Understanding of On and Off Target Effects of the Lymphocyte-Specific Kinase LCK for the Development of Novel and Safer Pharmaceuticals

Notes/Citation Information
Published in *Procedia Computer Science*, v. 108C, p. 1222-1231.

Digital Object Identifier (DOI)
https://doi.org/10.1016/j.procs.2017.05.268

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 108C (2017) 1222–1231

Procedia Computer Science

International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland

Towards a better understanding of on and off target effects of the lymphocyte-specific kinase LCK for the development of novel and safer pharmaceuticals

Xiaofei Zhang[1,*], Amir Kucharski[1,†], Wibe A. de Jong[2], and Sally R. Ellingson[1]

[1] University of Kentucky, Lexington, KY, USA

[2] Lawrence Berkeley National Laboratory, Berkeley, CA, USA

xzh289@g.uky.edu, anku223@g.uky.edu, wadejong@lbl.gov, sally@kcr.uky.edu

Abstract

In this work we have developed a multi-tiered computational platform to study protein-drug interactions. At the beginning of the workflow more efficient and less accurate methods are used to enable large libraries of proteins in many conformations and massive chemical libraries to be screened. At each subsequent step in the workflow a subset of input data is investigated with increased accuracy and more computationally expensive methods. We demonstrate the developed workflow with the investigation of the lymphocyte-specific kinase LCK, which is implicated as a drug target in many cancers and also known to have toxic effects when unintentionally targeted. Several LCK states and conformations are investigated using molecular docking and generalized Born and surface area continuum solvation (MM/GBSA). Different variations in the drug screening process provide unique results that may elucidate the biological mechanisms underlying the drug interactions.

Peer-review under responsibility of the scientific committee of the International Conference on Computational Science

1 Introduction

The development of new pharmaceuticals is a lengthy, expensive, and sometimes fatal process when toxic drugs make it to the clinical trial phase. With the number of potential drug targets, the genetic variations that exist in those targets, and the large chemical space of potential drugs, a full understanding of all these possibilities will never be realized experimentally, making computational methods extremely important to improve the drug discovery and development process. Many

[*] Wrote code and ran analysis

[†] Established MM/GBSA protocol and built protein models

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the International Conference on Computational Science
10.1016/j.procs.2017.05.268

computational methods exist to examine protein target and drug interactions, often with a huge trade-off between efficiency and accuracy, with the accuracy still being very dependent on the system being studied. This study looks at a multi-tiered approach of using a more efficient method to do an initial virtual screen of a chemical library and a theoretically more accurate method on a top-scoring subset from the initial screen. The test protein under investigation is lymphocyte-specific kinase LCK, an SRC-family protein implicated as a drug target in many cancers and also known to have toxic effects when unintentionally targeted. We have built models of several conformations, performed a first tier of molecular docking calculations using benchmark data, and done initial development of second tier calculations using molecular mechanics energies combined with generalized Born and surface area continuum solvation (MM/GBSA). This significant research will help alleviate the current economic burden of developing new pharmaceuticals by innovatively utilizing massive computational power and address important public health concerns by providing safer and more affordable drugs.

SRC-family proteins consist of conserved SRC homology (SH) domains starting from the C-terminus, SH1 (kinase, catalytic domain), SH2 (phosphotyrosine recognition region), SH3 (proline-motif recognition region), followed by a unique domain. Crystal structures of the LCK SH1, SH2, SH3, and combined SH2 and SH3 domains exist but not of all three complexed together. Several studies have evaluated the use of homology modeling and molecular docking for kinase drug design (Rockey & Elcock, 2006; Tuccinardi, Botta, Giordano, & Martinelli, 2010). To our knowledge, no in depth computational modelling efforts have been done to generate three-dimensional structures of LCK conformations (active and inactive) including all three SH domains, all of which have been targeted in drug discovery (Lee et al., 2010; Scapin, 2002; Zellefrow et al., 2006). In this study we build models of LCK, test them for usefulness in docking studies using a benchmark set of active and decoy compounds, and set-up an automated workflow to increase accuracy of docking calculations using MM/GBSA (Graves et al., 2008; Greenidge, Kramer, Mozziconacci, & Wolf, 2012; Rastelli, Degliesposti, Del Rio, & Sgobba, 2009; Zhang, Wong, & Lightstone, 2014). The plan is to use a multi-tiered approach of calculations allowing to screen massively large chemical libraries. The first step will be virtual screening with the very efficient but not as accurate molecular docking. Top scoring hits from this step will be used in an MM/GBSA screening. Here we report some initial findings using MM/GBSA as a second tier screening.

Figure 1: Graphical depiction of entire workflow

2 Methods

Multiple homology models are built to represent the active state, the inactive state, and the DFG-out inactive state. Molecular dynamics is performed on each model and snapshots are selected from the trajectory using RMSD based clustering of each frame. Docking is performed using the Database of Useful Decoys - Enhanced (DUD-E) compounds (Mysinger, Carchia, Irwin, & Shoichet, 2012) and enrichment plots are generated for each conformation. Then select sets of compounds and structures are rescored using MM/GBSA. The entire workflow is shown in Figure 1. Information on how to run each step of the workflow can be found at https://github.com/Xiaofei-Zhang.

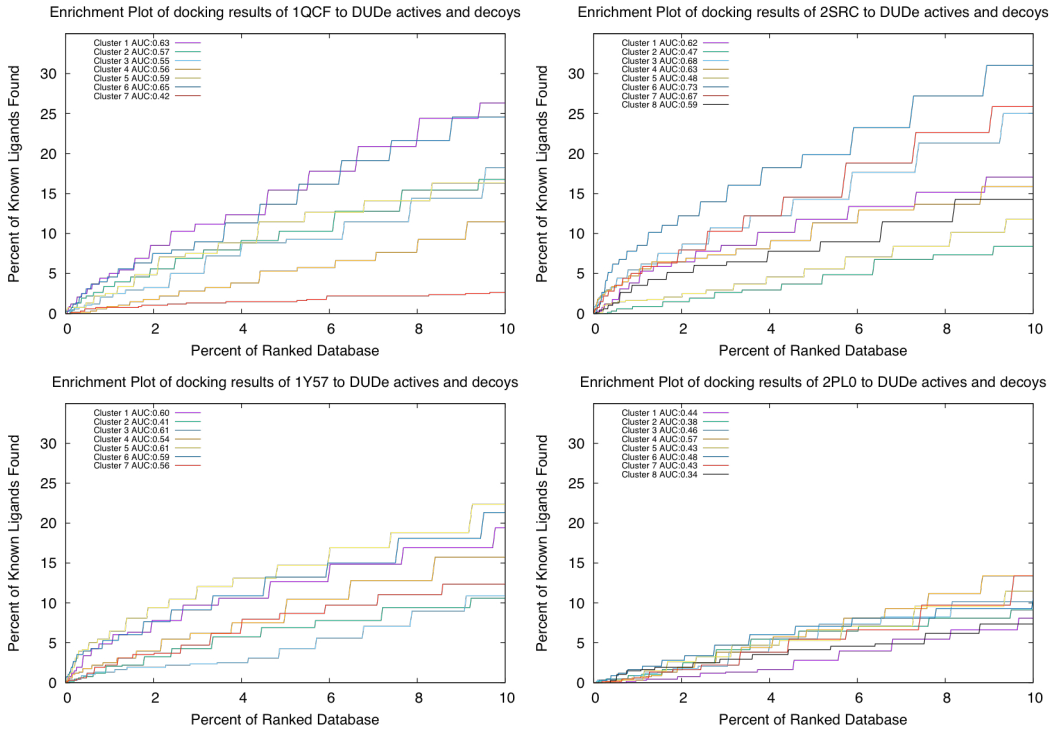

Figure 2: Docking enrichment plots for the DUD-e LCK dataset for each model and selected snapshot. Only early enrichment is shown.

2.1 Modeling

The LCK structure was generated using the UniProt primary sequence of human LCK (accession number P06239-1; (Koga et al., 1986; Perlmutter et al., 1988)) as input for the SWISS-MODEL Alignment Tool (Arnold, Bordoli, Kopp, & Schwede, 2006). Only residues 65 (or 67 depending on the PDB template) to 509 were used, as these correspond to the SH regions discussed above. Four models were built as described in the following. 1) Based on PDB structure 1QCF (Schindler et al., 1999). This structure is of HCK and chosen based on sequence identity. This model is of the inactive state. 2) Based on PDB structure 2SRC (Xu, Doshi, Lei, Eck, & Harrison, 1999). This structure is of c-SRC, a related protein in the SRC-like family. This model is also of the inactive state. 3) Based on PDB structure 1Y57 (Cowan-Jacob et al., 2005). This structure is of c-SRC in the active state. 4) Based on PDB structure 1QCF (Schindler et al., 1999) for the structure as a whole and then the coordinates for the kinase domain are swapped with that of 2PL0 (Jacobs, Caron, & Hare, 2008). 2PL0 is a structure of the kinase domain of LCK with imatinib bound. This model is of the inactive state with a DFG-out conformation. Sometimes residues at the C-terminus were excluded, including 507, 508, and 509. The C-terminus was extended using Chimera (Pettersen et al., 2004). Added residues were then assigned the most stable rotameric configurations based on the Dunbrack library in Chimera (Dunbrack, 2002). Protonation states of residues were identified using Protoss (Bietz, Urbaczek, Schulz, & Rarey, 2014).

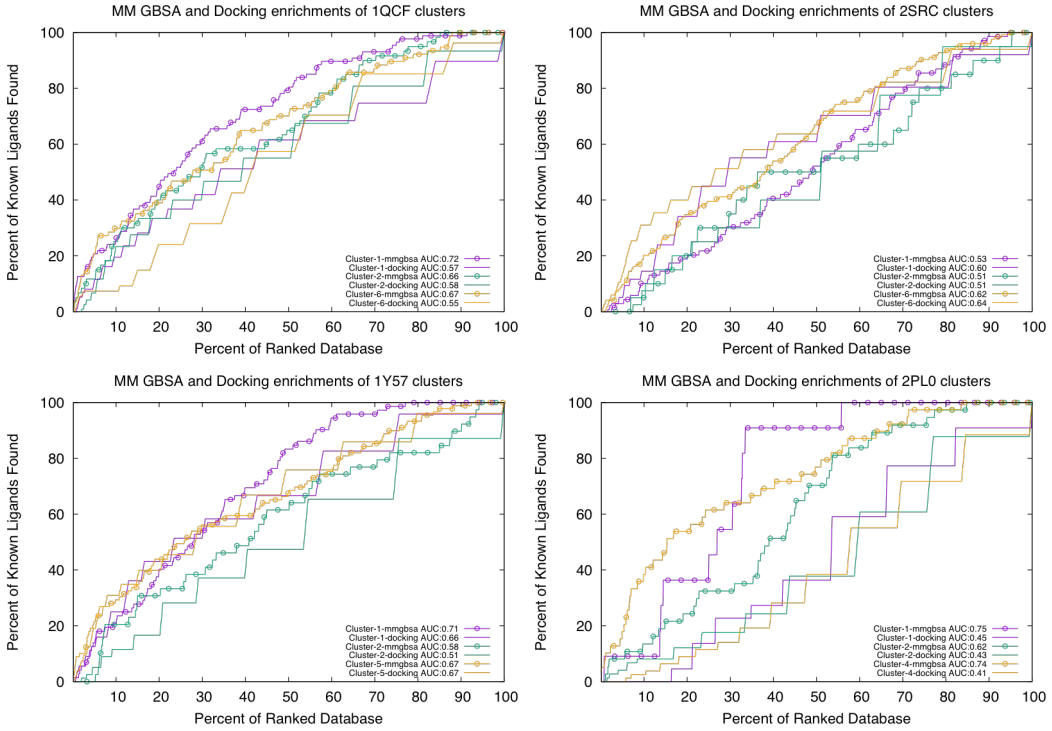

Figure 3: MM/GBSA and docking enrichment plots

2.2 Ensemble Docking

In order to do ensemble docking, molecular dynamics are first done on the resulting LCK models. The homology model was prepared using the combined CHARMM22 proteins/CHARMM27 nucleic acids topology and parameter files (MacKerell Jr et al., 1998). Phosphorylation of TYR 505 was accomplished using the par_all27_prot_na.inp and top_all27_prot_na.inp preparation files (Feng, Philippopoulos, MacKerell, & Lim, 1996). MD simulations were run using NAMD 2.10 (Phillips et al., 2005). Before production runs, the system was subject to 400 steps of minimization, 3000 steps of Berendsen thermostat/barostat equilibration, 100 steps of further minimization, and 3000 steps of Langevin thermostat/barostat equilibration. Then, one 100 ns production run was performed for each structure using Langevin thermodynamics. All time steps were 2 femtoseconds (fs), and frames from the production run were saved every 2 picoseconds (ps). Periodic boundary conditions and Particle Mesh Ewald (PME) electrostatics were used for all above MD procedures with an electrostatics cutoff of 12 Å. Seven to eight snapshots from the 100 ns production trajectory were chosen for each model based on clustering results from GROMOS (Christen et al., 2005).

Active and decoy compounds for LCK were obtained from DUD-E (Mysinger et al., 2012). Docking was performed using the AutoDock Vina (Trott & Olson, 2010) docking engine and VinaMPI (Ellingson, Smith, & Baudry, 2013), a virtual screening tool that allows one to perform a massive number of docking calculations using high performance computing resources. Scripts were used for automated preparation of pdbqt files. The binding site was determined using the position where imatinib, a known inhibitor of LCK is bound in a crystal structure (PDB ID: 2PL0 (Jacobs et al., 2008)). After docking, ranked lists of compounds, in which the top of the list has compounds

predicted to bind the best, are generated using the docking score for each selected snapshot of each cluster.

2.3 MM/GBSA

Docking scores from Vina have a small range and therefore a large number of compounds have the same score. Instead of rescoring an exact percent of each ranked list a cut-off energy value is chosen to have a list of compounds for rescoring that is the top 4-5% of each list. Prior to MM/GBSA calculations, minimizations on the protein-ligand conformations predicted from docking structures were performed using AMBER 12 and AmberTools (Case et al., 2005; Pearlman et al., 1995). Antechamber, parmchk and tleap tools were used in order to generate parameter files for the ligands. Complexes were then solvated in tleap using TIP3P explicit waters and 0.15 M KCl. The system was then subject to 2000 steps of minimization. The protein was held fixed for the first 1000 steps, while everything was mobile in the last 1000 steps. MM/GBSA calculations were then performed on the post-minimization structure for each complex. MM/GBSA is performed on the snapshots that come from the two largest clusters representing conformations in which the protein remained in the longest during MD simulations and from one shorter lived conformation that had a better docking enrichment than other conformations as this conformation may be important for drug binding. Multiple ways of combining MM/GBSA scores from multiple docking poses were evaluated. Parameters were set to output 20 docking poses, but sometimes less than 20 poses are generated if they are not significantly different. Because of the best early enrichment and similar results using other models and clusters, the average of the best 5 scores from processing all generated docking poses is used in all following MM/GBSA calculations.

3 Results

The DUD-e active and decoy datasets were docked to each snapshot extracted from the MD trajectories for each model and enrichment plots are given in Figure 2. The figures only display the top 10% of the ranked compounds as we are interested in improving early enrichment in order to create experimental test sets that are highly enriched with active compounds. The total AUC for each cluster is given in the legend though. It can be seen that the snapshot from the first cluster (longest lived conformation) never has the best AUC. All the models have snapshots that have better than random enrichments. However, the 2PL0 model only has one snapshot that gets better than average enrichment overall.

The MM/GBSA results are given in Figure 3. The MM/GBSA calculations are only done on the top 4-5% of compounds from each cluster. This is because of the computational time to do the calculations and to test whether or not it can be a step in a multi-tiered approach that increases the enrichment on subsets of data slightly enriched in previous steps. A docking curve is included on the plot for comparison. The line given here is an average between the maximum and minimum docking values obtained from placing all actives before and after the decoys with the same energy. The 2PL0 model gets the best improvement using MM/GBSA and 1QCF has consistently better enrichment using MM/GBSA over docking. The performance of MM/GBSA over docking for the 2SRC and 1Y57 models is not as good and at times is actually worse.

To further test the worth of ensemble docking, different binding calculations, and the use of multiple models, we investigated the number of unique actives found using different clusters, models, and methods which can be seen in Figure 4. The Venn diagrams in the left-most box shows the overlap of compounds identified in the top 10% of different clusters for each model and method. It can be seen that identified actives are most often specific to a particular snapshot. A union is done for the list of actives identified in each cluster and the overlap of actives identified within different

models is given in the top right box. Then a union is done for the list of actives identified in each model and the overlap of actives identified with different methods is given in the bottom right.

4 Discussion

Protein kinases, one of the largest families of proteins in higher eukaryotes, have over 900 protein products and contribute to a diversity of cellular processes (Anamika, Garnier, & Srinivasan, 2009). Protein kinases transfer a phosphate group from a bound ATP molecule to another protein substrate and regulate the majority of cellular pathways and signal transduction. Since kinase activity is so integral for normal cellular function, the deregulation of kinases has been implicated in many disease states, especially in cancers. Additionally, due to the high similarity in both sequence and structure between many kinases, kinase selectivity is a huge challenge for drug discovery. This in turn leads to off target effects that may be extremely toxic if drugs additionally interact with kinases that are normally expressed and not implicated in the given disease in which the drug is intended to relieve.

The SRC family are nonreceptor tyrosine kinases and include c-SRC, LCK, HCK, FYN, BLK, LYN, FGR, YES, and YRK. The lymphocyte-specific protein tyrosine kinase (LCK) is critical in T cell development and activation and is expressed through most of the T cell lifespan and at a somewhat constant rate through their development (Palacios & Weiss, 2004). LCK has implications, as suggested above, as both a drug target and also a serious off-target in which to avoid unintended

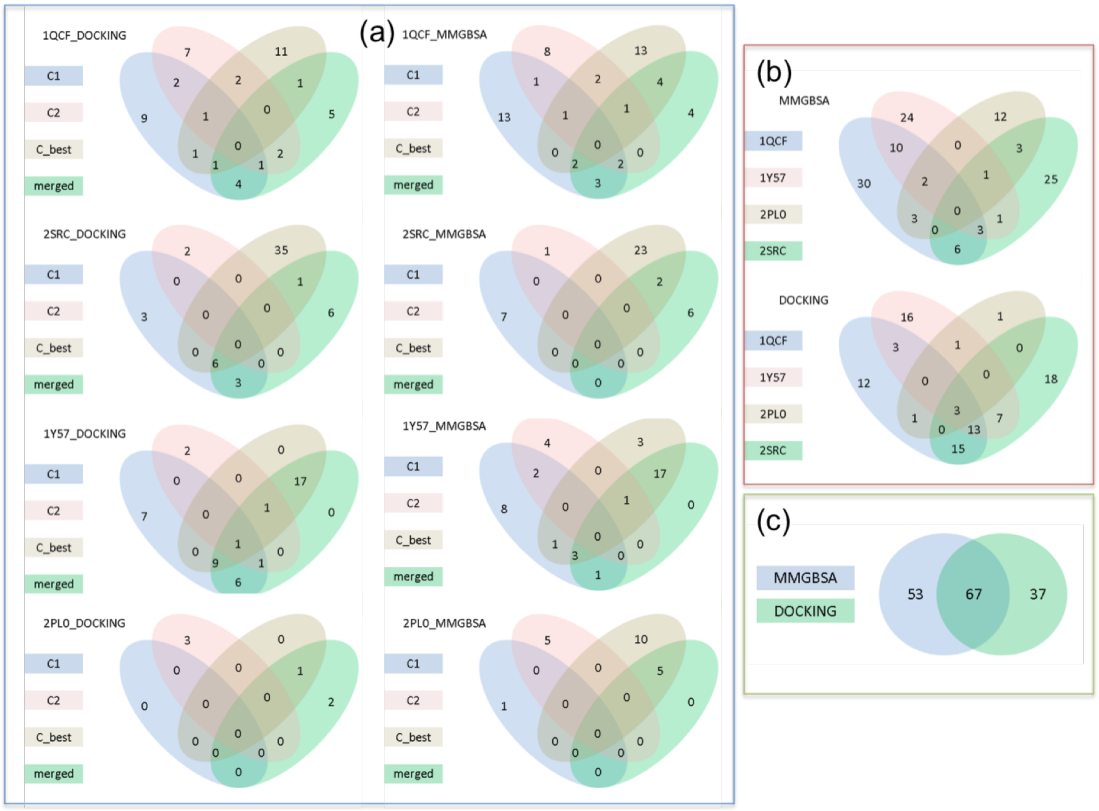

Figure 4: a) Actives identified in top 10% of each snapshot, C1 and C2 are snapshots representing the two largest clusters, C_best is the snapshot with the best docking enrichment, and merged uses the best docking score from any cluster, b) actives identified in each model, and c) actives identified with each method using MM/GBSA.

interactions. In fact, LCK is the kinase agreed upon by several pharmaceutical companies in a consensus minimal toxicity screen. They link LCK activation to T cell activation and LCK inhibition to T cell inhibition and severe (SCID)-like immunodeficiency which is characterized by devastating deficiencies in cellular and humoral immunity (Bowes et al., 2012; Goldman et al., 1998). LCK is a protooncogene (Abraham, Levin, Marth, Forbush, & Perlmutter, 1991) and overexpressed in Burkitt (Jücker, Abts, Eick, Lenoir, & Tesch, 1991), non-Hodgkin's B cell lymphoma (Knethen, Abts, Kube, Diehl, & Tesch, 1997), and lymphocytic leukemias (Majolini et al., 1998). High expression of LCK in chronic lymphocytic leukemia is linked with lymphocytic cell survival and thus studies suggest using LCK specific inhibitors for the treatment of progressive chronic lymphocytic leukemia (Talab, Allen, Thompson, Lin, & Slupsky, 2013). It has also been shown that inhibiting LCK enhances the ability for glucocorticoids to treat lymphoid malignancy, reversing glucocorticoid resistance (Harr et al., 2010). As many anticancer therapeutics are associated with severe adverse reactions and kinases are implicated in many cancers, a better understanding of the interactions of drugs with the entire kinome would give great insight for more efficacious and safe therapies.

Four models of LCK are investigated in this study. Although there are experimental structures for the kinase domain and the SH1 and SH2 domains of LCK, there are no experimental structures of all three of these domains intact. In order to have a more complete structure to obtain more accurate dynamics of the protein and have structures to eventually study alternate binding sites that have been implicated in kinase drug discover, homology models were built using structures of similar proteins with all three domains intact. The 1QCF is of the inactive state using HCK as a template based on sequence identity. The 1Y57 model is of the active state using c-SRC as a template since not many active structures exist for similar proteins. The 2SRC model is also of the inactive state using c-SRC as the template to be consistent with the template for the active model. The 2PL0 model is of the inactive state in a DFG-out conformation. A clustering of public human protein kinase structures has suggested that the catalytically active structure is similar among kinases and that there are two frequently observes inactive states called the "DFG-out" and "C-helix-out" states. In the DFG-out state an Asp sidechain that is part of an Asp-Phe-Gly (DFG) motif is rotated out of the ATP binding site making a larger pocket. In the C-helix-out state the αC helix is shifted away from the ATP binding site making a larger pocket. While other structures of inactive SRC-family kinases adopt the C-helix out conformation, a structure of the LCK kinase only domain bound with imatinib, an effective treatment of chronic myeloid and gastrointestinal stromal tumors (Lee et al., 2010), suggest that the inactive state of LCK may adopt the DFG-out state (Jacobs et al., 2008). This would support the finding that imatinib selectively inhibits LCK among SRC-family kinases (Lee et al., 2010) along with other kinases shown to adopt the DFG-out inactive conformation. However, there are currently no experimental structures for the combined three SH domains for LCK in the active state or locked in the inactive state to help support this.

Using the best enrichment from any cluster and average enrichment over all of the clusters, the order of performance of each model with docking scores is, from best to worst, 2SRC, 1QCF, 1Y57, and 2PL0. This could indicate that the active compounds preferentially target the inactive state of the protein which is normally in the C-helix-out conformation. However, it may sometimes exist in the DFG-out conformation, but less and/or different active compounds bind to this conformation. Most models have some improvement with the MM/GBSA calculations except for 2SRC which actually has decreased performance. Since this model was best performing by docking calculations, it seems that this model is well suited for docking and that methods are definitely model dependent. The greatest improvement by far is with the 2PL0 model. This could be because the active compounds identified in docking truly do bind to the DFG-out conformation and MM/GBSA performed well at rescoring these active compounds to this conformation.

Finally, to test the usefulness of using different conformations, models, and methods, we analyzed how many unique compounds are identified with each variation. It can be seen in Figure 4 that each variation gives information on different active compounds. Most of the identified actives are common

with both methods, but a large number of different actives are identified with each method, with MM/GBSA recovering more. If you use multiple ways of coming up with a set of compounds to test, you will also add to your false positives, and could likely get the same number of positive active compounds by just increasing the percent of compounds experimentally tested from an initial screening. However, using the approach here may provide insight into how the drug is interacting with the protein and provide useful information further down the drug development pipeline.

5 Conclusions

Although the MM/GBSA results do not always out perform the docking results, we show here that different variations in the drug screening process provide unique results which can be used to better understand the biological mechanisms underlying the drug interactions. Of the four models investigated, two of them consistently get better results and two get varying results using MM/GBSA. Future directions of this work include further tuning of the MM/GBSA calculations. Since the MM/GBSA calculations were done on the set of compounds with the best docking score for each cluster, the set of compounds with MM/GBSA scores is not consistent across clusters. Future work will involve calculations to screen a larger, consistent set of compounds to test if combining information across clusters can improve scores. Also, a third tier of increased theoretical accuracy, such as with density functional theory is being discussed.

6 Acknowledgements

This research used computational resources at the University of Kentucky's Center for Computational and the National Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. This work was supported by the National Institutes of Health (NIH) National Center for Advancing Translational Science grant KL2TR000116 and 1KL2TR001996-01. This work was supported in part by the U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the BLUFF.

References

Abraham, K. M., Levin, S. D., Marth, J. D., Forbush, K. A., & Perlmutter, R. M. (1991). Thymic tumorigenesis induced by overexpression of p56lck. *Proceedings of the National Academy of Sciences, 88*(9), 3977-3981.

Anamika, K., Garnier, N., & Srinivasan, N. (2009). Functional diversity of human protein kinase splice variants marks significant expansion of human kinome. *BMC genomics, 10*(1), 622.

Arnold, K., Bordoli, L., Kopp, J., & Schwede, T. (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics, 22*(2), 195-201.

Bietz, S., Urbaczek, S., Schulz, B., & Rarey, M. (2014). Protoss: a holistic approach to predict tautomers and protonation states in protein-ligand complexes. *Journal of cheminformatics, 6*(1), 1.

Bowes, J., Brown, A. J., Hamon, J., Jarolimek, W., Sridhar, A., Waldron, G., & Whitebread, S. (2012). Reducing safety-related drug attrition: the use of in vitro pharmacological profiling. *Nature Reviews Drug Discovery, 11*(12), 909-922.

Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., . . . Woods, R. J. (2005). The Amber biomolecular simulation programs. *Journal of computational chemistry, 26*(16), 1668-1688.

Christen, M., Hünenberger, P. H., Bakowies, D., Baron, R., Bürgi, R., Geerke, D. P., . . . Oostenbrink, C. (2005). The GROMOS software for biomolecular simulation: GROMOS05. *Journal of computational chemistry, 26*(16), 1719-1751.

Cowan-Jacob, S. W., Fendrich, G., Manley, P. W., Jahnke, W., Fabbro, D., Liebetanz, J., & Meyer, T. (2005). The crystal structure of a c-Src complex in an active conformation suggests possible steps in c-Src activation. *Structure, 13*(6), 861-871.

Dunbrack, R. L. (2002). Rotamer Libraries in the 21 st Century. *Current opinion in structural biology, 12*(4), 431-440.

Ellingson, S. R., Smith, J. C., & Baudry, J. (2013). VinaMPI: Facilitating multiple receptor high‐throughput virtual docking on high‐performance computers. *Journal of computational chemistry, 34*(25), 2212-2221.

Feng, M.-H., Philippopoulos, M., MacKerell, A. D., & Lim, C. (1996). Structural characterization of the phosphotyrosine binding region of a high-affinity SH2 domain-phosphopeptide complex by molecular dynamics simulation and chemical shift calculations. *Journal of the American Chemical Society, 118*(45), 11265-11277.

Goldman, F. D., Ballas, Z. K., Schutte, B. C., Kemp, J., Hollenback, C., Noraz, N., & Taylor, N. (1998). Defective expression of p56lck in an infant with severe combined immunodeficiency. *Journal of Clinical Investigation, 102*(2), 421.

Graves, A. P., Shivakumar, D. M., Boyce, S. E., Jacobson, M. P., Case, D. A., & Shoichet, B. K. (2008). Rescoring docking hit lists for model cavity sites: predictions and experimental testing. *Journal of molecular biology, 377*(3), 914-934.

Greenidge, P. A., Kramer, C., Mozziconacci, J.-C., & Wolf, R. M. (2012). MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. *Journal of chemical information and modeling, 53*(1), 201-209.

Harr, M., Caimi, P., McColl, K., Zhong, F., Patel, S., Barr, P., & Distelhorst, C. (2010). Inhibition of Lck enhances glucocorticoid sensitivity and apoptosis in lymphoid cell lines and in chronic lymphocytic leukemia. *Cell Death & Differentiation, 17*(9), 1381-1391.

Jacobs, M. D., Caron, P. R., & Hare, B. J. (2008). Classifying protein kinase structures guides use of ligand‐selectivity profiles to predict inactive conformations: Structure of lck/imatinib complex. *Proteins: Structure, Function, and Bioinformatics, 70*(4), 1451-1460.

Jücker, M., Abts, H., Eick, D., Lenoir, G., & Tesch, H. (1991). Over-expression of lck in Burkitt's lymphoma cell lines. *Leukemia, 5*(6), 528.

Knethen, A. V., Abts, H., Kube, D., Diehl, V., & Tesch, H. (1997). Expression of p56lck in B-cell neoplasias. *Leukemia & lymphoma, 26*(5-6), 551-562.

Koga, Y., Caccia, N., Toyonaga, B., Spolski, R., Yanagi, Y., Yoshikai, Y., & Mak, T. W. (1986). A human T cell‐specific cDNA clone (YT16) encodes a protein with extensive homology to a family of protein‐tyrosine kinases. *European journal of immunology, 16*(12), 1643-1646.

Lee, K., Ouwehand, I., Giannini, A., Thomas, N., Dibb, N., & Bijlmakers, M. (2010). Lck is a key target of imatinib and dasatinib in T-cell activation. *Leukemia, 24*(4), 896-900.

MacKerell Jr, A. D., Bashford, D., Bellott, M., Dunbrack Jr, R. L., Evanseck, J. D., Field, M. J., . . . Ha, S. (1998). All-atom empirical potential for molecular modeling and dynamics studies of proteins†. *The journal of physical chemistry B, 102*(18), 3586-3616.

Majolini, M. B., D'elios, M. M., Galieni, P., Boncristiano, M., Lauria, F., Del Prete, G., . . . Baldari, C. T. (1998). Expression of the T-cell–specific tyrosine kinase Lck in normal B-1 cells and in chronic lymphocytic leukemia B cells. *Blood, 91*(9), 3390-3396.

Mysinger, M. M., Carchia, M., Irwin, J. J., & Shoichet, B. K. (2012). Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking. *Journal of medicinal chemistry, 55*(14), 6582-6594.

Palacios, E. H., & Weiss, A. (2004). Function of the Src-family kinases, Lck and Fyn, in T-cell development and activation. *Oncogene, 23*(48), 7990-8000.

Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., DeBolt, S., . . . Kollman, P. (1995). AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. *Computer Physics Communications, 91*(1), 1-41.

Perlmutter, R. M., Marth, J. D., Lewis, D. B., Peet, R., Ziegler, S. F., & Wilson, C. B. (1988). Structure and expression of Ick transcripts in human lymphoid cells. *Journal of cellular biochemistry, 38*(2), 117-126.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., & Ferrin, T. E. (2004). UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of computational chemistry, 25*(13), 1605-1612.

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., . . . Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of computational chemistry, 26*(16), 1781-1802.

Rastelli, G., Degliesposti, G., Del Rio, A., & Sgobba, M. (2009). Binding estimation after refinement, a new automated procedure for the refinement and rescoring of docked ligands in virtual screening. *Chemical biology & drug design, 73*(3), 283-286.

Rockey, W. M., & Elcock, A. H. (2006). Structure selection for protein kinase docking and virtual screening: homology models or crystal structures? *Current Protein and Peptide Science, 7*(5), 437-457.

Scapin, G. (2002). Structural biology in drug design: selective protein kinase inhibitors. *Drug discovery today, 7*(11), 601-611.

Schindler, T., Sicheri, F., Pico, A., Gazit, A., Levitzki, A., & Kuriyan, J. (1999). Crystal structure of Hck in complex with a Src family–selective tyrosine kinase inhibitor. *Molecular cell, 3*(5), 639-648.

Talab, F., Allen, J. C., Thompson, V., Lin, K., & Slupsky, J. R. (2013). LCK is an important mediator of B-cell receptor signaling in chronic lymphocytic leukemia cells. *Molecular Cancer Research, 11*(5), 541-554.

Trott, O., & Olson, A. J. (2010). AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry, 31*(2), 455-461.

Tuccinardi, T., Botta, M., Giordano, A., & Martinelli, A. (2010). Protein kinases: docking and homology modeling reliability. *Journal of chemical information and modeling, 50*(8), 1432-1441.

Xu, W., Doshi, A., Lei, M., Eck, M. J., & Harrison, S. C. (1999). Crystal structures of c-Src reveal features of its autoinhibitory mechanism. *Molecular cell, 3*(5), 629-638.

Zellefrow, C. D., Griffiths, J. S., Saha, S., Hodges, A. M., Goodman, J. L., Paulk, J., . . . Schepartz, A. (2006). Encodable activators of Src family kinases. *Journal of the American Chemical Society, 128*(51), 16506-16507.

Zhang, X., Wong, S. E., & Lightstone, F. C. (2014). Toward fully automated high performance computing drug discovery: a massively parallel virtual screening pipeline for docking and molecular mechanics/generalized Born surface area rescoring to improve enrichment. *Journal of chemical information and modeling, 54*(1), 324-337.