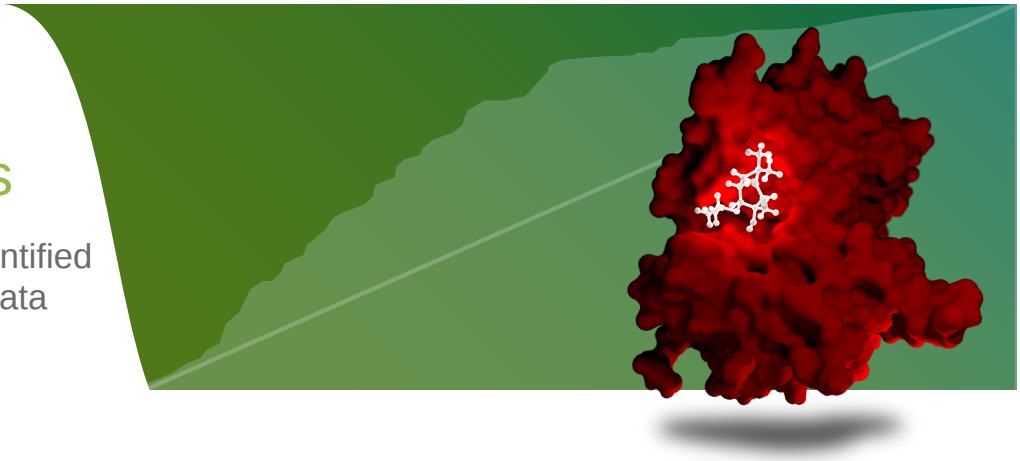
CYCLICA

PROBE$_X$ Proteome Docking: Identifying Off-target Interactions

Validating Off-target Interactions Identified by PROBE$_X$ Against Experimental Data from Kinase Inhibitors



PROBLEM
Identifying a drug's off-target interactions amongst 100,000 structurally characterized proteins *in silico* is a challenging task needing validation against experimental data

TECHNOLOGY
Ligand Express™ tools: PROBE$_X$ Proteome Docking

SOLUTION
A validation study with well-characterized kinase inhibitors demonstrates PROBE$_X$'s predictive capabilities in identifying off-target protein interactions.

INTRODUCTION

Drug development programs attempt to maximize efficacy and minimize adverse effects by developing molecules that interact with an individual, disease-related protein and little else. However, computational studies have indicated that a limited number of distinct binding sites exist, a fact that, when coupled with the large number of proteins found in the human body, implies a drug will interact with many proteins besides its intended target[1]. Identifying which proteins bind a particular drug provides critical information on the drug's pharmacological characteristics and serves to guide its development.

The majority of *in silico* approaches that predict off-target protein-drug interactions rely on so-called "target-fishing" methods[2], which fail to capture the full proteome landscape by only focusing on drug-target proteins represented by chemogenomics information. Cyclica's PROBE$_X$ proteome docking approach is independent of chemogenomics data, instead relying on a structure-based methodology that screens all structurally characterized proteins. To validate PROBE$_X$'s structure-based approach for identifying off-target interactions, we screened a collection of small molecule kinase inhibitors against all structurally-characterized proteins. The results of this screen were compared with known inhibitor-protein interactions extracted from PubChem's BioAssays database[3].

METHODOLOGY

Twenty-six kinase inhibitors with diverse chemical structures and extensive *in vitro* studies on the proteins they bind, both within[4] and outside of the kinome (where kinome refers to the collection of all kinases) were selected for screening with PROBE$_X$ (**Table 1**). PROBE$_X$ matches molecules against the entire structurally-characterized proteome, which consists of over 100,000 structures including 4,871 unique human proteins. In order to benchmark PROBE$_X$'s selectivity and specificity, PubChem's BioAssays database was used to build a set of protein-drug pairs that were experimentally determined to bind(or not) with the 26 kinase inhibitors under investigation. The set consisted of 346 *interacting* and 549 *non-interacting* protein-drug pairs referencing 319 unique structurally-characterized proteins; interactions involving kinases were excluded in order to focus on off-target interactions. Unlike other screening technologies, PROBE$_X$ predicts beyond the kinome; thus, this study focuses exclusively on proteins outside the kinome.

Following PROBE$_X$ screening, ranked protein-drug predictions were pooled and a receiver operating characteristic (ROC) curve was generated by comparing predictions to the experimental data. ROC curves efficiently depict to what extent a prediction is performing better than random. A straight line passing through the origin represents performance equal to results generated at random. ROC curves approaching the upper-left bounds represent predictions that are significantly better than random (i.e. promising selectivity and specificity). For an optimal score threshold, we calculated a contingency table with prediction metrics (significance, accuracy, sensitivity, and specificity) which is shown as an inset under the ROC curve (**Figure 1**).

Table 1. 26 kinase inhibitors used in this study, representing a range of chemical structures and known protein targets, both within and outside the kinome.

Kinase Inhibitors			
Selumetinib	Dasatinib	Sorafenib	Afatinib
Brivanib	Ruxolitinib	Staurosporine	Crizotinib
Lapatinib	Sunitinib	Gefitinib	Tofacitinib
Tandutinib	Doramapimod	Imatinib	Neratinib
Nilotinib	Lestaurtinib	Masitinib	Erlotinib
Tamatinib	Axitinib	Flavopiridol	SKI-606
PP-242	CI-1040		

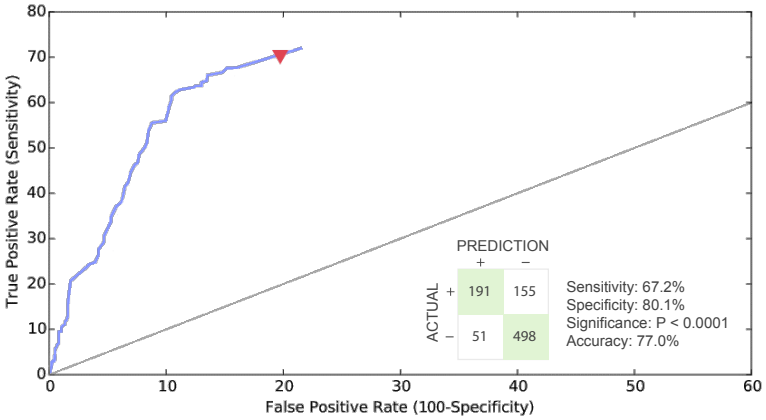


Figure 1. ROC curve for the combined pool of 26 kinase inhibitors. The blue curve represents the selectivity and specificity for off-target protein interaction predictions outside the kinome. The red triangle signifies the threshold value that optimally balances selectivity and specificity.

RESULTS

Figure 1 shows the ROC curve corresponding to $PROBE_X$ predictions of off-kinome interactions. $PROBE_X$ performed significantly (P<0.0001) better than random, with sensitivity (true positive rate) and specificity (true negative rate) of 67.2% and 80.1%, respectively. Accuracy (correct call rate) was 77%. These results are comparable to a representative example from a report by Brylinski and colleagues[5], which demonstrated a range of sensitivities (51-80%) and specificities of (66-89%) when predicting protein-drug interactions within the kinome. It is remarkable that $PROBE_X$ achieves similar accuracy for off-target interactions amongst the entire structurally-characterized proteome as has previously been achieved for a limited set consisting of only kinases.

In order to reduce computational demand, $PROBE_X$ first limits the number of proteins it docks by identifying surfaces complementary to the screened ligand. This intelligent filtration step results in early termination of the ROC curve since a significant portion of proteins are not ranked. Interpretation of the ROC curve reveals that $PROBE_X$'s discriminating step is successful in eliminating false positives (approx. 80%) at the cost of a small number of true positives (approx. 25%).

Several of the non-kinase interactors in the PubChem BioAssays database have been structurally-characterized with a ligand bound, allowing comparison of $PROBE_X$ predicted binding modes to reality. One example is that of flavopiridol. In **Figure 2** we compare the predicted binding pose of flavopiridol with that experimentally seen in the crystal structure of glycogen phosphorylase.

The results of this study indicate that $PROBE_X$ is capable of identifying off-target interactions with high specificity from the large set of proteins explored. $PROBE_X$ also provides a high resolution molecular model of the ligand in the putative binding site. This makes $PROBE_X$ a valuable tool for the identification of unknown targets for small molecules, opening up applications in phenotypic screening, investigation of mechanism of action of natural compounds, elucidation of adverse effects, and drug repurposing.

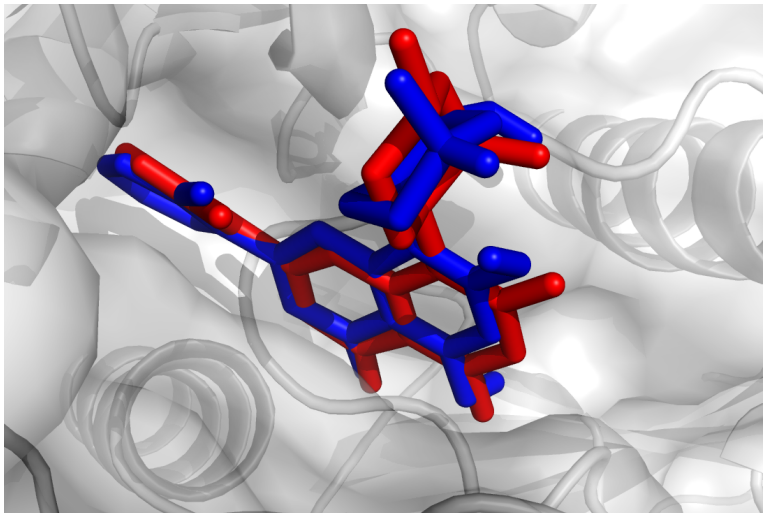


Figure 2. Flavopiridol docked (blue) to the inhibition site of glycogen phosphorylase along with its corresponding experimentally determined binding mode (red).

SUMMARY

A collection of kinase inhibitors was assembled for the validation of $PROBE_X$, a proteome docking technology performing off-target prediction on proteins overlooked by other technologies. Proteome docking results for a set of kinase inhibitors against all non-kinases were compared with experimental results, yielding an ROC curve and prediction metrics. $PROBE_X$ produced predictions substantially better than random (P<0.0001) and with high specificity (80.1%) and sensitivity (67.2%), showcasing its ability to identify probable protein-ligand interactions within the entire proteome. $PROBE_X$ is capable of anticipating off-target interactions for novel drug entities, providing important pharmacological information that can lead to more efficacious and safer drugs.

RESOURCES

1. Gao, M. & Skolnick, J. A Comprehensive Survey of Small-Molecule Binding Pockets in Proteins. *PLoS Comput. Biol.* 9, e1003302 (2013).

2. Rognan, D. Structure-Based Approaches to Target Fishing and Ligand Profiling. *Mol. Inform.* 29, 176–187 (2010).

3. Wang, Y. et al. PubChem's BioAssay Database. *Nucleic Acids Res.* 40, D400–D412 (2012).

4. Davis, M. I. et al. Comprehensive analysis of kinase inhibitor selectivity. *Nat. Biotechnol.* 29, 1046–51 (2011).

5. Brylinski, M. & Skolnick, J. Cross-Reactivity Virtual Profiling of the Human Kinome by X-React KIN: A Chemical Systems Biology *Approach. Mol. Pharm.* 7, 2324–2333 (2010).

CYCLICA INC.

18 King St East, Suite 801
Toronto, Ontario, M5H 1A1, Canada
1-416-304-9201

101 Main Street, Floor 14
Cambridge, Massachusetts,
02014-1519, United States