Journal of Molecular Graphics and Modelling 57 (2015) 36–48

Contents lists available at ScienceDirect

Journal of Molecular Graphics and Modelling

journal homepage: www.elsevier.com/locate/JMGM

Can structural features of kinase receptors provide clues on selectivity and inhibition? A molecular modeling study

Sarangan Ravichandran*, Brian T. Luke, Jack R. Collins

Advanced Biomedical Computing Center, Frederick National Laboratory for Cancer Research (FNLCR), P.O. Box B, Frederick, MD 21702, USA

ARTICLE INFO

Article history:
Accepted 16 December 2014
Available online 12 January 2015

Keywords:
Kinase drugs
Docking
Activity
Mutation
Cross-reactivity
Biological Pathways

ABSTRACT

Cancer is a complex disease resulting from the uncontrolled proliferation of cell signaling events. Protein kinases have been identified as central molecules that participate overwhelmingly in oncogenic events, thus becoming key targets for anticancer drugs. A majority of studies converged on the idea that ligand-binding pockets of kinases retain clues to the inhibiting abilities and cross-reacting tendencies of inhibitor drugs. Even though these ideas are critical for drug discovery, validating them using experiments is not only difficult, but also in some cases infeasible. To overcome these limitations and to test these ideas at the molecular level, we present here the results of receptor-focused in-silico docking of nine marketed drugs to 19 different wild-type and mutated kinases chosen from a wide range of families. This investigation highlights the need for using relevant models to explain the correct inhibition trends and the results are used to make predictions that might be able to influence future experiments. Our simulation studies are able to correctly predict the primary targets for each drug studied in majority of cases and our results agree with the existing findings. Our study shows that the conformations a given receptor acquires during kinase activation, and their micro-environment, defines the ligand partners. Type II drugs display high compatibility and selectivity for DFG-out kinase conformations. On the other hand Type I drugs are less selective and show binding preferences for both the open and closed forms of selected kinases. Using this receptor-focused approach, it is possible to capture the observed fold change in binding affinities between the wild-type and disease-centric mutations in ABL kinase for Imatinib and the second-generation ABL drugs. The effects of mutation are also investigated for two other systems, EGFR and B-Raf. Finally, by including pathway information in the design it is possible to model kinase inhibitors with potentially fewer side-effects.

1. Introduction

Cancer continues to be a major cause of death among the human diseases. As per the latest statistics [1], new US cancer cases and deaths continues to be a major health concern, and for 2014 they are estimated to be around 1,665,540 and 585,720 respectively. Hence, there is an urgent need for novel cancer drugs with fewer side effects. Protein kinases are enzymes that help carry out the transfer of a phosphate group (phosphorylation) from ATP to substrate proteins [2–4] and play an important role in cancer and signal transduction pathways. So far, approximately 518 kinase genes [5,6] have been identified in the human genome, which amounts to about 1.7% of the total 20-25 K genes. Protein phosphorylation is, in most cases, initiated by extracellular signals and acts like a switch to control a substrate's binding affinity, receptor's enzymatic activity, and the cellular location of their products. These actions can, in turn, influence down-stream gene/protein functions. Kinases also act as effectors and participate in a wide-range of important cellular functions ranging from cell cycle modulation, DNA repair, immunity, growth, and apoptosis. Since kinase activity influences many cellular functions, its regulatory role is usually kept under tight control. Kinase phosphorylation is a reversible reaction, and enzymes called phosphatases catalyze the removal of phosphate groups from substrate proteins and make it ready for the next cycle [4].

Kinase dysregulation, in majority of cases, is the result of unchecked activity resulting in over-expression of proteins or downstream genes, and such a state of increased activity has often been positively correlated with many human diseases such as cancers, metabolic disorders, and infectious diseases. Several events that modify kinases and leave them in elevated active states have been identified [7]. A short list of these events include mutation [8], up-regulation (gene amplification) [9], simultaneous expression

* Corresponding author. Tel.: +1 301 846 1991; fax: +1 301 846 5762.
E-mail address: ravichandrans@mail.nih.gov (S. Ravichandran).

http://dx.doi.org/10.1016/j.jmgm.2014.12.007

of multiple kinases [10], and increased expression of transporters [11].

Kinase dysregulation can be best understood by studying Chronic Myeloid Leukemia (CML). CML is a unique disease caused by unregulated kinase activity of a single malfunctioning gene, BCR-ABL [12], and is commonly invoked to understand the uncontrolled behavior of kinases. This disease is caused by the aberrant genetic material called the Philadelphia chromosome, formed when a section of DNA from chromosome-9 translocates to a region after the BCR gene in chromosome-22. This creates a fusion gene, BCR-ABL, which expresses a protein product also called BCR-ABL [6,12]. Unlike normal ABL, the fusion protein is constitutively active and does not require activation by external signals. Hence, BCR-ABL remains active and participates in cell-cycle pathways, and strongly influences cell-division and growth processes, leaving the affected cell in a highly unstable state.

Other than aberrant genetic material formation, mutations in and around the kinase active site can also influence kinase activity [13]. For example, it had been observed that a small fraction of Philadelphia-positive (ph+) patients who are undergoing Imatinib therapy can relapse and offer resistance to the therapy. This resistance has been attributed to the mutations in the active site that reduces the drug's ability to bind and inhibit activity. Other forms of resistance to treatment have also been documented. Notably, tumor cells tip the balance of regulating phosphorylation–dephosphorylation in their favor by specifically down-regulating the phosphatase to indirectly overcome receptor inhibition [7].

Malfunction arising from a single gene such as BCR-ABL, can in turn affect other down-stream pathway genes to create non-local effects. Since most cancers are associated with dysregulation of multiple genes involved in different pathways, the effect of malfunction can spread more rapidly than the single gene disease to even affect far-away expressed gene products. Constitutional rogue kinase activity or malfunction due to mutations can be controlled by using either small molecule inhibitor drugs or humanized antibodies. The later approach of using a humanized antibody, Trastuzumab, to target ERBB2 receptors for treating breast cancer has shown some potential, but the benefit is not universal [14].

In cancer treatment, the idea of targeting multiple receptors using one or more drugs is not new [3,7,11,15,16]. Imatinib, a drug known for inhibiting BCR-ABL, has also been identified to inhibit platelet-derived growth factor (PDGF), stem-cell factor (SCF) and c-KIT receptors. Imatinib (marketed as Gleevec [http://www.gleevec.com]) is currently being prescribed as a primary drug for treating both CML [12] and GIST [11]. The justification for inhibiting multiple, but related, genes involved in disease proliferation has been clearly illustrated [7] using the tyrosine kinase inhibitor PP121. PP121 was shown to inhibit PI3K and its downstream gene mTOR simultaneously, thus providing additional benefit of blocking both the negative feed loop initiator (mTOR) and the oncogene, PI3K.

Using experimental and modeling data, a class definition for the inhibitors (Type I/II) based on their chemistry and ability to trap the kinases in an active or inactive conformation [17–22] has been proposed. It is not clear whether the inhibitors force the kinases toward a particular conformation or the receptors influence the drug molecules to take a specific binding pose. Recent studies [17,19,21,23] have identified several key issues in kinase inhibition. Of significance is the work by Davis et al. [18,19]; the most elaborate study known to date. These authors used 72 inhibitors to create a selectivity profile for more than 80% of the known 518 kinases. The important points arising from this study [19] are that the inhibitor Type (I/II) does not guarantee selectivity and, most importantly [24], the structural features that are presented by the receptor during the binding often controls selectivity and therefore the choice of the best inhibitor. These findings will most likely have a profound impact on basic research and future kinase drug development.

Table 1
List of kinase target receptors used in this study.

PDB ID	Drug name	EC #	Gene	UniProt	Res. (Å)
1OPJ	Imatinib	2.7.10.2	ABL	P00520	1.75
3UE4	Bosutinib	2.7.10.2	ABL	P00519	2.42
3GVU	Imatinib	2.7.10.2	ABL	P42684	2.05
3CS9	Nilotinib	2.7.10.2	ABL	P00519	2.21
2GQG	Dasatinib	2.7.10.2	ABL	P00519	2.40
1T46	Imatinib	2.7.1.112	KIT	P10721	1.60
3MIY	Sunitinib	2.7.10.2	ITK	Q08881	1.67
3K54	Dasatinib	2.7.10.2	BTK	Q06187	1.94
1XBB	Imatinib	2.7.1.112	Syk	P43405	1.57
2PL0	Imatinib	2.7.10.2	LCK	P06239	2.80
2ZVA	Dasatinib	2.7.10.2	Lyn	P25911	2.60
2OIQ	Imatinib	2.7.10.2	cSRC	P00523	2.07
3G5D	Dasatinib	2.7.10.2	cSRC	P00523	2.20
3GP0	Nilotinib	2.7.11.24	MAPK	Q15759	1.90
3GCS	Sorafenib	2.7.11.24	MAPK	Q16539	2.10
1XKK	Lapatinib	2.7.10.1	EGFR	P00533	2.40
1M17	Erlotinib	2.7.10.1	EGFR	P00533	2.60
2ITZ	Gefitinib	2.7.10.1	EGFR	P00533	2.80
3BBT	Lapatinib	2.7.10.1	ERBB4	Q15303	2.80
1UWH	Sorafenib	2.7.11.1	B-RAF	P15056	2.95

To test these findings, receptor structure focused self- and cross-docking simulations were performed using the experimentally available small molecule kinase drug-bound PDB structures (19 in total) to analyze the respective binding modes and their corresponding estimated affinities. Our results qualitatively agree with the known kinase inhibition profiles and that of Davis et al. [18,19]. Based on our data set, we find that receptor features dictate ligand inhibition and Type II ligands in general are competitive for DFG-out conformation. On the other hand, Type I ligands show less selectivity. Lapatinib appears to behave like a Type II molecule based on its strong affinity for inactive receptor folds but based on its atypical receptor conformation (DFG-out and c-alpha helix conformation) requirements [19], we have decided to place Lapatinib along with Type-II ligands but without a Type classification. Additional simulations carried out on the wild-type and mutated protein system, ABL, confirm that the receptor structural features can qualitatively capture the change in binding affinities [13], and might be useful for identifying potent family-specific inhibitors. The influence of mutations on the binding affinity was further analyzed using EGFR and B-Raf kinase systems and the docking results show the change in the binding affinity and mean binding energy indicative of the negative effect of selected mutations on binding. Finally, using the common pathway information from our gene set, we explore the possibility of single drug or drug combinations that might be useful for advanced stage cancers.

2. Materials and methods

2.1. Kinase systems

Protein kinase structures co-crystallized with drug molecules were downloaded from RCSB-PDB (http://www.rcsb.org/pdb/) and used for docking simulations. Kinase families studied in this work include ABL, KIT, ITK, BTK, SYK, LCK, LYN, MAP, EGFR, erbB4, and BRAF kinases (see Table 1). The small molecules examined in this study are Dasatinib, Erlotinib, Sunitinib, Gefitinib, Bosutinib, Lapatinib, Sorafenib, Nilotinib, and Imatinib (see Fig. 1). Note that these are marketed small molecule drugs currently being used for the treatment of different types of cancers and tumors. Each kinase family was searched in RCSB-PDB database for the availability of 3D structures that have been co-crystallized with drug molecules

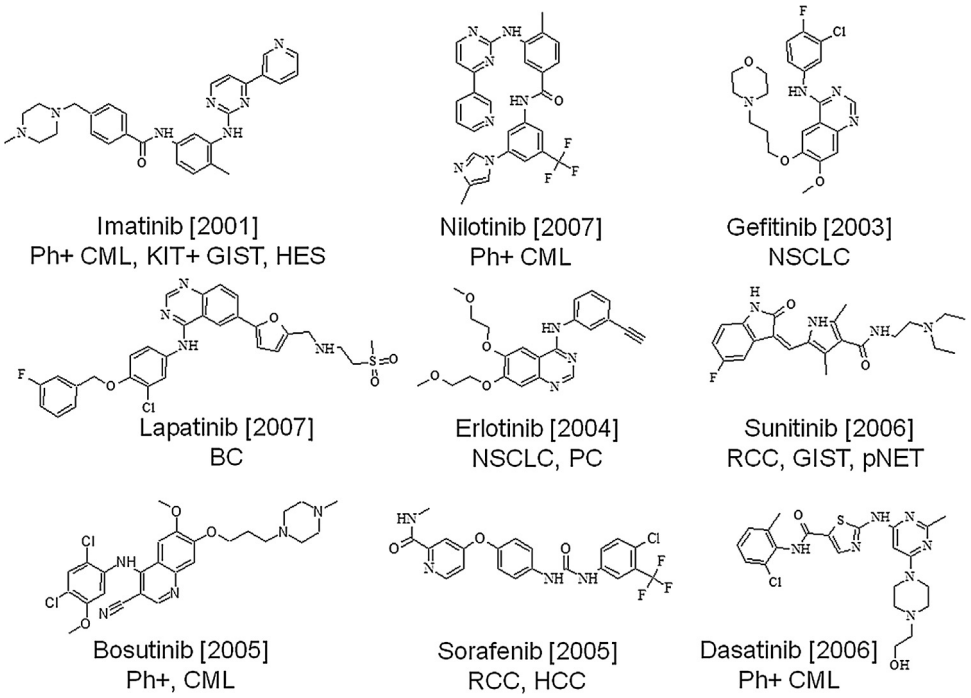

Fig. 1. Chemical drawings of drug molecules studied in this work: chronic myelogenous leukemia (CML); gastrointestinal stromal tumor (GIST); hypereosinophilic syndrome (HES); non-small cell lung carcinoma (NSCLC); breast cancer (BC); lung cancer (LC); renal cell carcinoma (RCC); pancreatic cancer (PC); hepatocellular carcinoma (HCC); pancreatic neuroendocrine tumor (pNET); KIT positive (KIT+); Philadelphia chromosome positive (Ph+). Drug release date is shown after their generic name.

from our list; the best structure with high quality from the list was eventually picked for docking simulations (see Table 1).

2.1.1. Receptor structural information

The kinase receptor families (shown in bold face) used in this study along with their RCSB PDB IDs (shown in parentheses) are: **ABL1/2** (3UE4, 1OPJ, 3UE4, 3GVU, 3CS9) [25–28], **C-KIT** (1T46) [29], **ITK** (3MIY) [30], **BTK** (3k54) [31], **SYK** (1XBB) [32], **LCK** (2PL0) [33], **LYN** (2ZVA) [34], **SRC** (3G5D, 2OIQ) [35,36], **EGFR** (1M17, 1XKK, 2ITZ) [37–39], **ERBB4** (3BBT) [40], MAP (3GP0, 3GCS) and **B-RAF** (1UWH) [41] (Table 1).

2.2. Receptor preparation

The 3D coordinates for the kinase-inhibitor complexes were downloaded from RCSB (http://www.rcsb.org/pdb/). Molprobity (http://molprobity.biochem.duke.edu/) was used for analyzing the PDB structures, adding hydrogens, and applying residue (ASN/GLU/HIS) flips wherever applicable. The output from Molprobity was further analyzed using the SwissPdb Viewer (http://www.expasy.org/spdbv/) and Discovery Studio [ver. 3.5, Accelrys Inc., San Diego, CA]. The protein structures were kept rigid except during the flexible in silico mutation experiments (see Section 2.5), where only the mutated residue was allowed to rotate. AutoDockTools (ADT, version 1.5.6 rc3), a graphical-user-interface for AutoDock, was used to prepare the molecules in AutoDock suitable formats (pdbqt). Gasteiger-PEOE partial charges for molecules were added using ADT. AutoDock input parameter files for grid and docking (GPF, DPF; provided in the Supplemental Pages) were also prepared using ADT. Note that we did not make any attempt to build any missing receptor residue segments that are either away from the binding pocket or not involved in the binding.

2.3. Ligand preparation

Ligand structures (Fig. 1) were downloaded from the NCBI Pubchem database (http://pubchem.ncbi.nlm.nih.gov/) and read into Discovery Studio (Accelrys Inc., San Diego, CA) for further modification and analysis.

2.4. Docking Simulations

AutoDock [42–45] is an automated docking [46] method for identifying the binding modes of ligands with biomolecule receptors. AutoDock Vina [47] is the latest software from the AutoDock family, but uses a different approach for identifying the binding modes and is also significantly faster than AutoDock. In the present study, AutoDock (ver 4.2.3; [45]) and Vina (ver 1.1.2) were used for modeling the binding modes and/or estimating inhibition constants [48] because they both perform unbiased docking and use atomistic details for describing the molecules [49,50]. A brief summary of the AutoDock is provided in the Supplementary Pages S1. Also, these two programs force fields had shown great potential in reproducing crystal-bound conformations with high accuracy. AutoDock (vers. 3 and up) uses a modified force field [44] that allows the user to also predict binding affinities along with the binding free energies.

2.4.1. AutoDock and AutoDockTools (ADT)

AutoDockTools [45] was used for reading PDB files, adding H's and Gasteiger charges, and generating AutoDock input (pdbqt) files. AutoGrid with a grid spacing of 0.375 Å (AutoDock: spacing parameter) was used as a default spacing parameter in our simulations. The grid box size was chosen to contain $100 \times 100 \times 100$ grid points (x, y and z direction) which was found to be large enough to cover the kinase active site and important neighboring residues (see Supplementary figure, Fig. S1). The center of the map was chosen for each system using the co-crystallized ligand center of mass position. ADT was used for generating the grid parameter file (GPF) and the Docking Parameter File (DPF). The AutoDock Genetic Algorithm runs were set as follows: ga_run 200, population size is 300, maximum number of energy evaluations is 2.5×10^7 (# of torsions $\leq$10) or 5.0×10^7 (# torsions >10). We have invoked the AutoDock

post-processing step to carry out conformational clustering on the final docked conformations. Tolerance (rmstol) for the cluster analysis was set to 2.0 Å. Binding energies from the top-cluster were averaged and reported as the mean-binding energy. A detailed list of non-default AutoDock parameter options used in this study is also presented the Supplementary Section S1.

2.4.2. Vina

AutoDock Vina's scoring function and the method is detailed elsewhere [47], here we review only the key points. Vina's scoring function is based on X-score [51] and had been further adjusted using PDBbind [52–54]. Vina's scoring function was developed to take advantage of both knowledge-based schemes and empirical methods. Unlike AutoDock, Vina does not use explicit hydrogen atoms or the partial charges for atoms in the simulations. Optimization in Vina starts with a mutation of the ligand conformation followed by a local optimization. The optimization algorithm used is the Iterated Local Search Global Optimizer [55,56], and the acceptance or rejection follows the Metropolis criterion [13]. Sample of Vina input template *config* file is shown in Supplementary Section S2.

2.4.3. Additional simulation details

The docking simulations were run on a Linux cluster (Intel Xeon; nehalem) with the Ubuntu operating system (12.04.02). Linux C-shell scripts developed locally were used for PBS (portable batch job) submission, post-processing and analysis. After the simulations, the docked ligand conformations were extracted and compared with the crystal-bound conformation to estimate whether the docking had identified the crystal-bound conformation in the top ranking solutions. For this, root-mean-squared-deviation (RMSD) between the docked solutions and crystal-bound conformations were computed using locally developed FORTRAN software (f77 and f90; additional info in Supp. Section S3). The post-analysis and figures were made using Discovery Studio software (Accelrys Inc., San Diego, CA).

2.5. Mutation in silico experiments

To analyze the role of receptor structural features of mutated kinase systems in inhibitor binding, we first focused on the ABL kinases. The reason for choosing this system is because ABL kinases are well-studied and the BCR-ABL related disease, CML, had been classified as a single-gene disorder. Hence, it is relatively easy to study this system compared to other kinases. A 2009 study by Radaelli et al. [13] analyzed several key mutations that occur in this kinase and their impact on the activity by monitoring the IC50 fold changes. We used the data set from this study as our main reference for comparisons. Of all the mutations reported in the above mentioned work [13], we focused our modeling efforts on two highly resistant mutations, T315I and V299L. Mutated systems were prepared from the corresponding wild-type kinase structures: ABL1-Imatinib (PDB: 1OPJ), ABL1-Bosutinib (PDB: 3UE4), ABL1-Dasatinib (PDB: 2GQG) and ABL1-Nilotinib (PDB: 3CS9).

To study the impact of active site residue mutations on drug binding and kinase activity, we examined the existing crystal structures of ABL kinase systems with T315I mutation. Three structures (PDB: 3QRI, 3QRJ, 3QRK) [57] have been reported with the T315I mutation. Upon structure comparisons of our T315I modeled structure with these three crystal structures, we found that with the mutated sequence still retained the WT fold with a similar local conformation around T315. For the V299L mutation, our searches of the RCSB PDB database did not identify any relevant experimental 3D structures. The two residues (V and L) involved in this second mutation are both non-polar, side-chain flexibilities are from low to moderate, and both could make minimal VDW interactions with no possibility of forming H-bonds. So, we believe that ABL should be able to accommodate the V299L point mutation and fold similarly to the wild-type. Starting from the wild-type structures, we introduced single point mutations in positions T315 and V299 separately to create two mutant structures (ABL-T315I and ABL-V299L). The mutated structures were further energy minimized using Discovery Studio (ver. 3.5) to relieve any strain introduced by the mutation. AutoDock has the capability of modeling a limited number of flexible residues. We invoked this option to also carry out both rigid and flexible docking simulations for each mutated system. Finally, we also extended our impact analysis to include two additional kinase systems, EGFR (1M17:T766M; 1XKK:T790M) and B-Raf (1UWH:E500G), to model the impact of selected mutations that occur in the active sites of B-Raf and EGFR [58–60]. Note that 1UWH:E500 occupies position E501 in the Uniprot sequence, P15056. System preparation and the analysis for EGFR and B-Raf and simulations were carried out as described above for the ABL systems.

3. Results and discussion

Marketed kinase drug-bound complexes were collected from the RCSB-PDB and the best structure based on resolution and R-factors from each available family were chosen for docking (see Fig. 1 and Table 1). A recent comprehensive experimental study [19] was able to show that the family-specific receptor structural features play a key role in their inhibition, and further went on to explore the possibility of developing unique inhibitors for different kinase families. Using this study as a basis, we performed in silico experiments using receptor-focused models to analyze the role of receptor structural features on kinase inhibition. We then explored how inhibitors respond to the changes in the receptor features in the active site to explain kinase inhibition and cross-reactivity within, and across, different receptor families using pathway information.

The structural transformation of kinases from inactive to active forms involves many mobile regions, including a major movement of an activation loop over several angstroms (see Fig. 2A). In spite of several decades of work, many questions in the area of kinase inhibition still remain unanswered. One straightforward approach to model kinase inhibition is to employ dynamical methods, such as Brownian dynamics [61] or molecular dynamics [62,63], and follow the entire process. Simulations based on these or other novel approaches [64] have been attempted before but the benefits are not universal across different kinase families. Moreover, most of these approaches are computationally intensive and time could be a factor in applying them for large datasets. A reasonable alternative is to model kinase inhibition by splitting the problem and focusing separately on the most recurring active and inactive structures of the drug/inhibitor-bound experimental complexes and use them as a reasonable starting point to understand inhibition. We took this later approach to study kinase inhibition using structure-focused models, followed by force-field based docking simulations and pathway-based analysis.

Kinase structures are often present in a dormant (inactive) state prior to activation. Experimental structures representing both active and inactive forms for most kinase families are readily available from RCSB Protein Data Bank (RCSB-PDB; www.rcsb.org). Since kinase sequences from different families show overall similarity with a common 3D bi-lobe fold, each family exploits the following common sub-regions to create unique binding cavities for activity: (i) Catalytic loops, (ii) Nucleotide binding region, (iii) Hinge, α-helix, and (iv) DFG motifs. Experimental structures provide a glimpse of the flexible regions that play a key role during kinase activation. The differences between active and inactive receptor

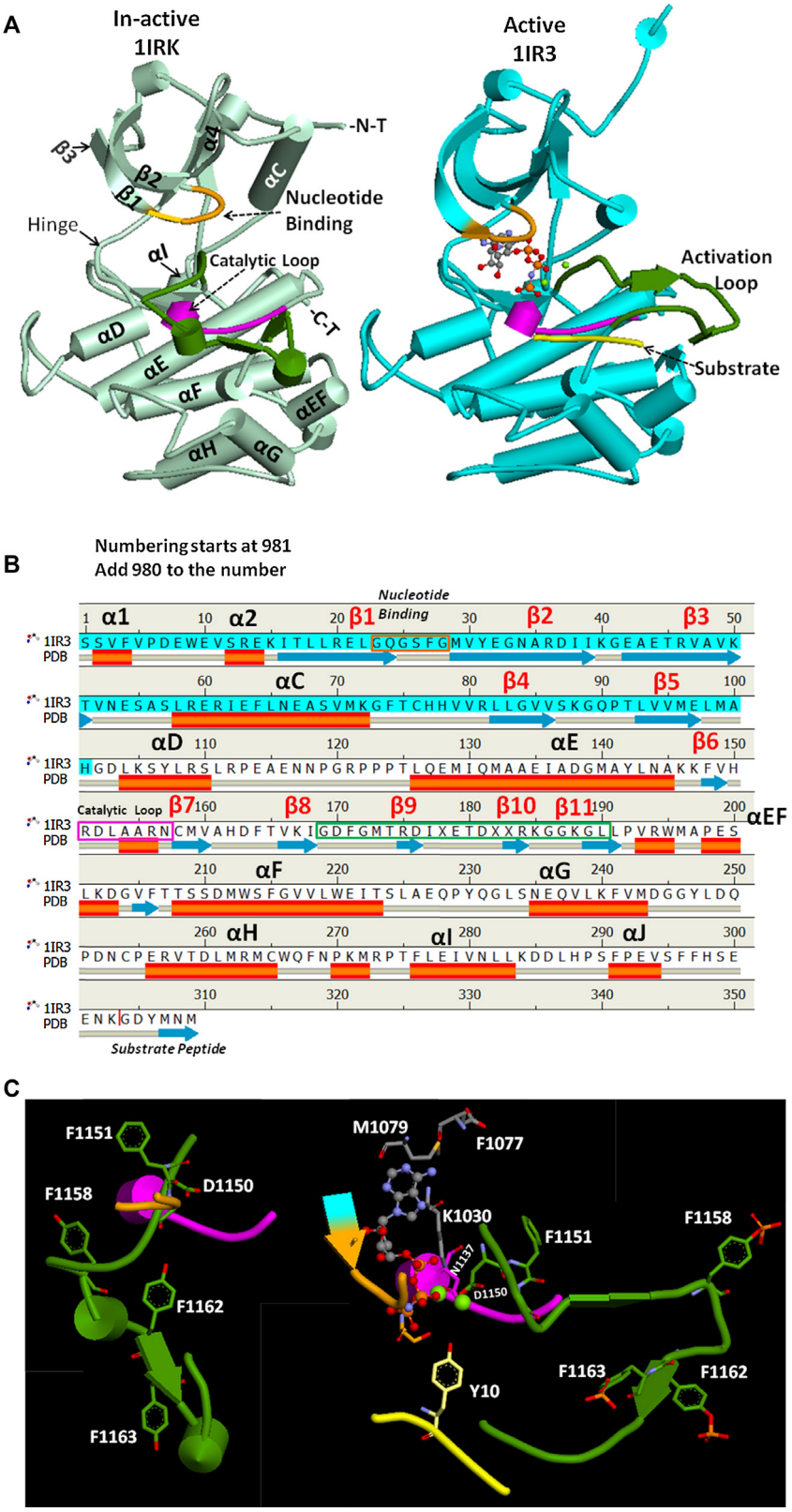

Fig. 2. (A) Kinase 3D architecture differences between the inactive (IA; light green; shown on left) and active (A; cyan; shown on right) forms are highlighted using insulin receptor kinase (IR3) as an example. Key functional domains are shown in different colors for visual clarity. The aligned structures were moved apart to show the differences. (B) Amino acid sequence of the insulin receptor (PDB: 1IR3; active form) is displayed along with the secondary structural information. Important structural domains (see A) are marked. The cyan background shows the range of the smaller N-terminal lobe (see A). Region between β8 and αEF marks the Activation loop (shown using green box). The last six residues identified by a red separator correspond to the substrate peptide. (C) Zoomed-in view of the active site of the insulin receptor forms shown in A. The proteins are shown in the aligned orientation and the color coding is similar to A. Note the difference in the orientation of residues, DF (part of the $D^{1150}F^{1151}G$ motif), and the change in the position of activation loop (green) between the inactive (left) and the active (right) kinases. Phosphorylated PHE residues in the A-loop are shown in stick mode. The yellow color chain marks the substrate peptide. Ligand is shown in ball-and-stick and green balls mark the location of Mg ions.

Table 2
RMSD (Å) values for the AutoDock (ADT)/Vina predicted drug conformation to the cocrystallized drug. The rank (R) of the modeling conformation from AutoDock and Vina are also provided. AutoDock solutions are further clustered (Cluster size, C) using a tolerance limit of 2.5 Å.

Genes	Kinase receptors (PDB ID)	Drugs	ADT(R,C)	Vina (R)
ABL1	1OPJ	Imatinib	2.11 (1,66)	11.61 (7)
ABL1	3UE4	Bosutinib	3.09 (1,7)	3.54 (3)
ABL2	3GVU	Imatinib	1.90 (1,2)	5.05 (7)
ABL1	3CS9	Nilotinib	1.91 (1,126)	5.33 (1)
KIT	1T46	Imatinib	3.59 (1,5)	6.27 (6)
ITK	3MIY	Sunitinib	4.24 (4,21)	4.67 (8)
BTK	3K54	Dasatinib	1.38 (2,18)	0.75 (4)
SYK	1XBB	Imatinib	1.14 (2,14)	0.66 (2)
LCK	2PL0	Imatinib	1.60 (1,32)	3.80 (1)
LYN	2ZVA	Dasatinib	1.36 (2,33)	0.28 (4)
SRC	2OIQ	Imatinib	2.06 (1,87)	2.87 (1)
SRC	3G5D	Dasatinib	0.82 (1,30)	0.89 (5)
MK11	3GP0	Nilotinib	2.87 (4,11)	3.39 (2)
MK14	3GCS	Sorafenib	0.47 (1,176)	0.42 (1)
EGFR	1XKK	Lapatinib	1.21 (1,10)	1.45 (3)
EGFR	1M17	Erlotinib	2.86 (2,63)	1.57 (2)
EGFR	2ITZ	Gefitinib	2.51 (4,7)	3.32 (1)
ErbB4	3BBT	Lapatinib	1.30 (1,6)	1.17 (1)
B-RAF	1UWH	Sorafenib	0.62 (1,140)	0.67 (1)

kinase conformations can be easily understood by comparing the corresponding insulin kinase receptor structures (Fig. 2A–C). For ATP-competitive small molecule kinase inhibitor drugs, a grouping method based on their abilities to recognize partner kinase structural conformations (active or inactive), using ABL kinase system as a reference, has been proposed [17–22]. Using this formalism, Type II ligands prefer to target inactive kinase conformations in which the DFG (one letter amino acid code) motif is present in a "DFG-out" conformation, and is not conducive to either ATP or substrate binding. Due to differences in the held-in activation loop (see Fig. 2A–C) conformations, each inactive conformation can provide a unique interaction site for inhibitors. These sites, in most cases, have been shown to be different for different kinase families, but similar among sub-families, and can be used to explain the varied potencies for different inhibitor drugs. Type I ligands, on the other hand, do not depend on the unique DFG-out conformation for binding (Fig. 2C). Based on our structure-based docking results, it can be seen that the Type I ligands favor active kinase conformations. The active kinase conformations are relatively more open than inactive forms and on average are similar across different kinases.

Table 2 presents a comparison of AutoDock and Vina docking results against the experimentally co-crystallized conformations using Root Mean Squared Deviation (RMSD) measure (see Fig. 3). Our docking results show that the overwhelming majority (17/19; Table 2) of the top-ranked docked solutions are found to be within ~3 Å from their respective co-crystallized ligand conformations (see Fig. 3) with AutoDock, while only 10 of the 19 Vina conformations show the ligand within 3 Å of the crystal structures. Based on the RMSD measure, and the capability to estimate inhibition constants, we focus on AutoDock computed results in this paper; results of docking simulations using Vina are shown in the Supplementary Material (Table S1). Mean binding energies and the estimated inhibition constants from AutoDock calculations are shown in Tables 3 and 4, respectively. Each row in the tables quantitatively describes how different drug molecules respond to the micro-environment a receptor presents. Each column provides a trend of how a single drug molecule responds to the influence of the shape of the active site and/or the changes in the key residues lining the active site, and the results give an estimate of affinity and cross-reactivity across different receptors. An extended form of Table 3 is shown in Supplementary Table S2.

Our simulations qualitatively capture the overall trends observed by Davis et al. [19] and others [18]. Table 4 lists the estimated inhibition constants [44,45] from docking simulations. Binding energies and inhibition constants show that the Type II or inactive receptor conformation favoring drugs are more tightly bound compared to Type I or DFG-in, favoring ligands (Tables 3 and 4). This can be clearly seen from the estimated inhibition constants between Type II (low nM) predictions compared to a high nM–μM range for Type I ligands. For example, Erlotinib (type I) shows weak binding to ABL1 (~180 nM), which is indicative of less compatibility to inactive fold when compared to Type II favoring Imatinib (4.5 pM). In BRAF (PDB: 1UWH), the predicted inhibition constants for Type II ligands show similar smaller values indicative of strong binding (Lapatinib: 0.12, Nilotinib: 1.84, Sorafenib: 1.69) than for the Type I drug molecules (Sunitinib: 273.30, Erlotinib: 489.68, Dasatinib: 17.34 nm, etc.). Lapatinib has been assigned to both Type I [65] and an undetermined Type [19] ligand. Our simulation results show that Lapatinib exhibits preference for inactive conformations and behaves more like a Type II when compared to Type I molecules. On the other hand, it is also compatible with open Type I folds (Tables 3 and 4). Using ABL as a reference, we can see that Lapatinib's estimated inhibition constant (0.11 nm) is in the low nM range and similar to Imatinib (4.5 pM). This trend for Lapatinib is seen across the different receptors. For instance, with EGFR (inactive conformation, PDB: 1XKK) the mean binding energy is −12.43 kcal/mol compared to −8.57 kcal/mol for the active form (PDB: 2ITZ) (Table 3).

The recently solved structure for the Bosutinib-bound ABL1 (PDB: 3UE4) uncovered an unique inactive kinase conformation not previously observed in ABL kinases, and our results show that its binding site structural features are incompatible with Type II favoring drugs such as Imatinib and others (see Tables 3 and 4). This is reflected in the mean binding energy that is estimated to be around −7.95 kcal/mol for the ABL1-Bosutinib structure (PDB: 3UE4), compared to −14.79 kcal/mol in the ABL1-Imatinib kinase structure (PDB: 1OPJ) (Table 3). Similar behavior is observed for other Type II drug molecules reported in this study when interacting with the PDB: 3UE4 (see Tables 3 and 4) binding pocket. Examining the binding energy values for different receptors, one can see that Bosutinib behaves like a Type I drug, even with a receptor that exhibits an inactive fold, but more experimental evidence is needed before it can be assigned to a specific type. Even though the binding site features of ABL1 (PDB: 3UE4) are unique in this case, it is not surprising since, as alluded earlier, the inactive conformations can arise from different combinations of active loop and c-helix conformations when compared to active forms, and can be unique. Type II inhibitors from our set show higher selectivity for the inactive active site than the active Type I site. For example, Imatinib inhibits LCK (PDB: 2PL0, 0.26 nM) but not SYK (PDB: 1XBB, 318.76 nM), SRC (PDB: 3G5D; 149.21 nM), or LYN (PDB: 2ZVA; 221.87 nM) [33], and the selectivity of Lapatinib for the inactive forms of EGFR are in the low nM range (Table 4). Estimation of binding affinity in modeling is a well-known challenge for force-field based methods and still needs improvement, especially in predicting inhibition constants [48,66]. In spite of difficulties, AutoDock qualitatively identifies the correct trend. This is supported by the comparison with experiment shown in Supplemental Table S6. In this table, the calculated inhibition constants are compared to experimental dissociation constants [19]. Though these are very different quantities, the qualitative trends are often in agreement. A scatter plot of the calculated and experimental results for Bosutinib shows a general monotonic trend (Supplemental Fig. S2), that is reflected in a high Spearman's rank correlation ($\rho = 0.685$, p-value = 0.035). For Dasatinib, the poor binding to SYK that was experimentally observed is also found computationally, as is the relatively poor binding of Imatinib to LYN.

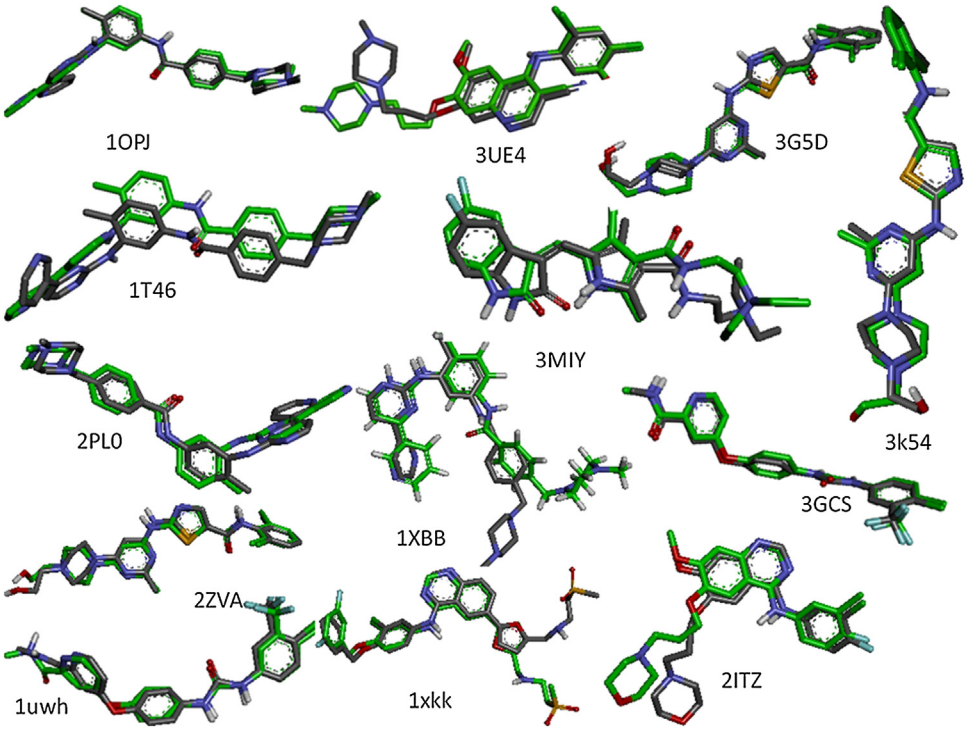

Fig. 3. Comparison of AutoDock predicted and crystal-bound conformations (green color stick) for key kinases studied in this work. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

To explore the mutual impact of the receptor-drug pair on each other during inhibition, we focused our attention on the well-studied drug Imatinib binding to different kinases. In our data set, we have four different Imatinib bound kinase complexes (ABL, KIT, LCK, SRC) that were captured in the inactive conformation. Sequence alignment for this set is displayed in Fig. 4A and the alignment was used as a guide to overlay the crystal structures (1OPJ, 1T46, 2PL0 and 2OIQ) (Fig. 4A and B) using the PDB: 1T46 structure as a template. The co-crystallized ligands after alignment are shown in Fig. 4B and C. One can see that the crystal bound ligands after alignment show that they all bind to their respective receptors using very similar conformations, and the receptor features are outlined in Fig. 4B and C. Comparison of the predicted inhibition constants from our set show that the ABL1 (PDB: 1OPJ, 0.0045 nM) active site provides a very favorable environment for Imatinib and it is therefore expected to bind stronger than to other kinases. For instance, based on the predicted binding affinity, the KIT binding site is less conducive for Imatinib compared to ABL, PDB: 1T46 (2.56 nM). Even though the binding poses are very similar, the predicted binding affinities depend on the micro-environment as shown in the structure-focused sequence alignment, Fig. 4A. For example, the following non-conserved substitutions between 1T46 and 1OPJ could explain the differences in the binding affinity: **Y**672F, **C**673M, **D**677N, **C**788F, **I**808V and **C**809A. Note that the last C to A change occurs a few residues ahead of the catalytically important and well-known DFG motif.

We used the BCR-ABL kinase system from our docking simulations to see whether we could explain the IC50 fold changes

Table 3
AutoDock computed negative Mean-Binding Energies (MBE, kcal/mol). Receptor conformational folds are indicated as Inactive (IA) or Active (A), and the drug types (I,II or none) are also shown in parenthesis following their names.

Mode	Gene	PDB	Dasatinib (I)	Erlotinib (I)	Sunitinib (I)	Gefitinib (I)	Bosutinib	Lapatinib	Sorafenib (II)	Nilotinib (II)	Imatinib (II)
IA	ABL1	1OPJ	10.26	8.75	9.29	8.73	8.08	12.04	11.15	11.56	14.79
IA	ABL1	3UE4	8.83	6.93	6.85	6.95	7.95	9.65	7.29	8.43	9.65
IA	ABL2	3GVU	9.93	6.86	9.06	8.48	7.60	12.05	11.01	11.11	12.53
IA	ABL1	3CS9	9.57	8.03	8.62	8.64	7.10	11.60	9.85	12.60	11.27
IA	KIT	1T46	10.23	8.75	9.08	9.10	8.91	12.08	10.59	10.64	11.11
A	ITK	3MIY	7.97	6.12	6.54	7.30	7.20	9.32	7.85	9.10	9.66
A	BTK	3K54	9.07	6.83	7.31	7.59	7.93	9.54	8.28	9.69	9.16
A	SYK	1XBB	7.39	6.18	6.30	6.54	7.35	7.91	7.24	8.43	8.17
IA	LCK	2PL0	9.32	6.96	7.82	8.54	9.21	11.09	9.73	11.41	12.46
A	LYN	2ZVA	8.89	6.82	6.04	7.17	7.35	8.42	8.39	8.72	8.42
IA	SRC	2OIQ	9.00	7.22	8.13	8.36	9.42	10.87	9.46	11.04	12.00
A	SRC	3G5D	9.83	7.12	6.38	8.04	7.77	9.14	7.03	8.98	8.74
IA	MAPK11	3GP0	9.69	7.03	8.24	7.98	8.81	12.39	10.33	10.26	11.04
IA	MAPK14	3GCS	8.69	7.58	8.11	7.98	9.13	9.93	10.60	10.65	9.86
IA	EGFR	1XKK	9.75	8.06	7.88	7.96	8.22	12.43	7.52	8.80	9.64
A	EGFR	1M17	8.35	6.56	7.01	7.32	7.64	9.31	7.36	9.11	8.89
A	EGFR	2ITZ	8.57	6.57	6.60	7.00	6.95	8.57	7.28	9.23	8.68
IA	ERBB4	3BBT	8.97	8.23	8.45	8.54	8.13	10.94	8.44	9.43	9.53
IA	B-RAF	1UWH	9.81	7.90	8.63	9.26	9.55	13.17	11.31	11.45	10.95

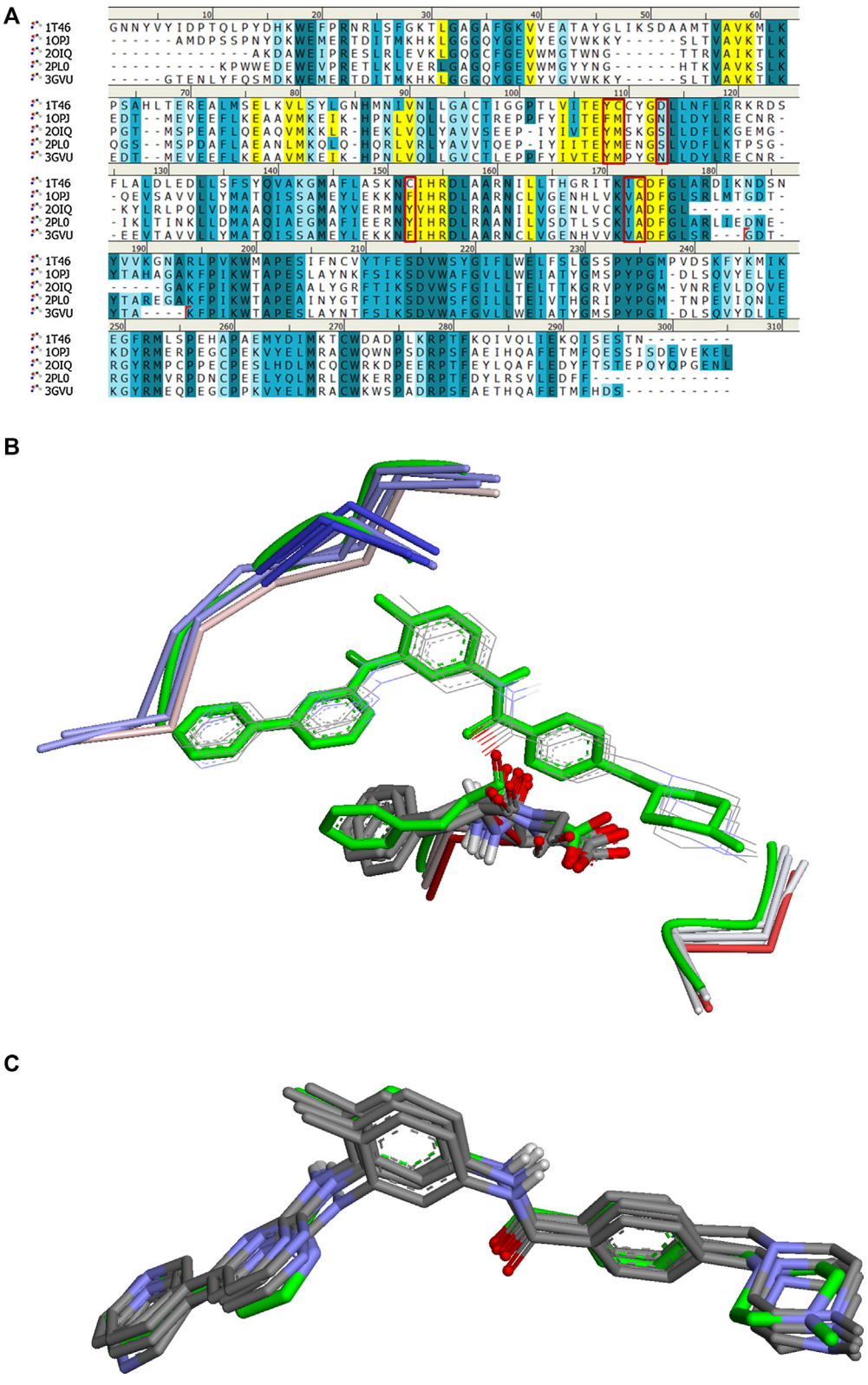

Fig. 4. (A) Sequence alignment of Imatinib-bound inactive kinases, C-KIT: 1T46, ABL1: 1OPJ, SRC: 2OIQ, LCK: 2PL0 and ABL2: 3GVU (sequence identity: 22.3%; sequence similarity: 41.3%). Sequence alignment was used as a guidance to align the corresponding structures and by choosing PDB ID 1T46 as the template. Residues in the 5 Å radius sphere surrounding the 1T46 co-crystallized Imatinib are highlighted in yellow. Key residue changes that could explain the difference in imatinib binding are marked in red boxes. (B) A 5 Å region surrounding the cocrystallized imatinib (see A) is shown in a backbone representation. The partial protein backbone structures are colored blue to red to show the N-terminal to C-terminal direction. The DFG motif is shown in stick form in atom type color. The reference protein (1T46) is identified by green color. RMSD value for each protein with reference to the template structure (1T46) is also shown in Supplemental Table S3. (C) Imatinib molecules from the corresponding PDBs (see A and B) are shown in overlayed form. 1T46 ligand is shown in green. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
AutoDock estimated inhibition constants (k_i; nM). The co-crystallized ligand for the receptor is marked in underlined font. Receptor conformational mode (inactive: IA; active: A) is identified in the Mode column. Potent inhibitor identified by k_i for each receptor (shown in column) is shown in italic font. Types of drugs (I or II), if identified, are shown in parenthesis following their name.

Mode	Gene	PDB	Dasatinib (I)	Erlotinib (I)	Sunitinib (I)	Gefitinib (I)	Bosutinib	Lapatinib	Sorafenib (II)	Nilotinib (II)	Imatinib (II)
IA	ABL1	1OPJ	13.70	179.79	91.30	168.18	446.70	0.11	3.56	3.38	*0.0045*
IA	ABL1	3UE4	64.72	2760.00	6590.00	1730.00	685.71	*27.17*	2510.00	363.72	67.04
IA	ABL2	3GVU	17.33	3300.00	143.45	255.36	965.63	0.31	4.19	*3.70*	*0.27*
IA	ABL1	3CS9	47.79	665.92	297.12	124.98	1220.00	*0.55*	19.55	0.24	0.74
IA	KIT	1T46	27.15	75.13	104.01	60.46	295.47	*0.95*	8.87	5.90	2.56
A	ITK	3MIY	1430.00	4620.00	*8440.00*	1360.00	2750.00	*25.40*	639.49	149.76	20.56
A	BTK	3K54	102.28	2540.00	1060.00	924.91	506.92	53.44	255.68	*37.22*	68.91
A	SYK	1XBB	1030.00	8870.00	10560.00	4210.00	1850.00	*70.83*	3510.00	400.48	318.76
IA	LCK	2PL0	49.03	3120.00	954.04	253.12	175.93	3.62	37.02	3.06	*0.26*
A	LYN	2ZVA	*30.90*	2680.00	10040.00	1460.00	1340.00	*93.06*	255.79	224.16	221.87
IA	SRC	2OIQ	*34.85*	1510.00	201.93	292.06	34.74	4.93	*54.91*	4.32	*0.59*
A	SRC	3G5D	*17.00*	1580.00	3990.00	348.00	574.50	43.89	2400.00	87.54	*149.21*
IA	MAPK11	3GP0	55.90	1950.00	578.42	554.03	185.46	*0.15*	8.80	14.19	6.34
IA	MAPK14	3GCS	65.22	921.85	*375.57*	161.60	203.14	52.97	*8.25*	10.22	43.69
IA	EGFR	1XKK	48.33	741.11	888.34	608.52	583.94	*0.14*	1510.00	*159.71*	25.93
A	EGFR	1M17	221.88	*5580.00*	3080.00	1520.00	1080.00	*26.99*	2150.00	81.45	82.59
A	EGFR	2ITZ	149.14	5300.00	7380.00	*4900.00*	1890.00	*189.85*	2690.00	*85.71*	223.98
IA	ERBB4	3BBT	116.18	178.58	294.00	181.26	268.92	*1.21*	156.94	30.34	23.96
IA	B-RAF	1UWH	17.34	489.68	273.30	75.61	63.81	*0.12*	1.69	1.84	4.50

Table 5A
AutoDock estimated inhibition constants (nM) and their corresponding fold increases obtained from docking simulations for the wild-type (WT) and mutated ABL systems.

System	1OPJ/Imatinib	3UE4/Bosutinib	2GQG/Dasatinib	3CS9/Nilotinib
WT	0.0045	685.71	1.52	0.2376
V299L	0.0147	2690.00	8.15	0.1613
T315I	0.1128	3590.00	103.24	1.8900
WT	1.00	1.00	1.00	1.00
V299L	3.26	3.92	5.36	0.68
T315I	24.95	5.24	67.92	7.95

observed in mutated systems using a simple structure-focused model. The ABL system is chosen because (i) It is linked to diseases, CML and GIST, (ii) It is relatively easy to model the inhibition behavior because the diseases are attributed to the malfunction of a single gene, (iii) It is a well-studied system with sufficient experimental information; and (iv) The ready availability of several experimental structures. The results (Table 5A) show that using this approach, we are able to qualitatively capture the changes due to mutation by focusing on the receptor conformation. The key mutation, T315I, that causes the ABL systems to escape the inhibitory effects of not only Imatinib but also the second-generation drugs is well known [14]. Please note that position T315 is same as T334 in some of the PDB systems described in Table 5A and Table S4, so, from now on for clarity, we will only use T315 to identify this position. Our simulations shows that Imatinib, Bosutinib, Dasatinib and Nilotinib show fold changes of approximately 25, 5, 68, and 8 (see Table 5A) compared to their corresponding wild types. We also followed a second mutation, V299L, which showed a relatively smaller effect on the drug binding compared to the key T315I mutation. Please note that position V299 is same as V318 in some of the PDB systems described in Table 5A and Table S4, so, from now on for clarity, we will only use V299 to identify this position. Our results follow the trends observed for IC50 fold changes [13]. Flexible simulations for these systems follow similar order (see Suppl. Tables S4 and S5). The idea of grouping ligands by Type I/II is definitely interesting, but when it comes to selectivity or identifying new compounds based on known Types is not informative [19,20]. Based on the known Types, one would expect ligands to be compatible only for a specific family or sub-family receptors if the receptors are capable of exhibiting specific active site arrangements during kinase activity. Along these lines, similar binding affinities of ligands to a receptor can provide information about their selectivity and receptor 3D fold. For instance, the predicted inhibition constants of Type II drugs such as Nilotinib, Sorafenib and a Type I drug Dasatinib are similar and indicative of the fact that they are often used as second-generation alternatives for Imatinib. When the receptor sequences are modified by mutations, the binding behaviors of previously known partner ligands start to deviate from their original associated types. For example, as shown by our results for the newly seen ABL receptor inactive fold (PDB: 3UE4), the known ABL Type II drugs that target the inactive fold now become ineffective (1OPJ vs. 3UE4 from Tables 3 and 4). Earlier studies have also shown that different generations of Type II ligands are structurally different from each other and even the newer generation Type II drugs also contain a Type I scaffold [20].

This mutation analysis was expanded by including two additional kinase systems, EGFR and B-Raf. The PDB structures used for this analysis were EGFR: 1M17/1XKK and B-Raf: 1UWH. Experimental studies on EGFR identify the gatekeeper residue mutation, T790M, as a secondary mutation observed in NSCLC cell lines which often lead to resistance for the binding of either Gefitinib or Erlotinib [58–60]. The gatekeeper mutation and its structural impact on the activity has also been addressed by Yun et al. [67]. They observed that T790M mutant binding pockets for both the active and inactive conformations, have enough space to accommodate inhibitors, and, using their co-crystallized complexes, showed that the inhibitor binding poses are very similar [67]. Our mutation experiments using the inactive form of EGFR (1XKK) show similar or even better binding affinities, but the active ones (1M17) are negatively influenced by the mutation (Table 5B). Finally, the simulations for the B-Raf mutation E500G show the loss of glutamic acid interactions when replaced with glycine. This mutation has been observed in Cardio-facio-cutaneous (CFC) syndrome patients [68] and, as our simulations show, could arise from the weakened interactions from either loss of flexibility, or H-bonds, or increase in hydrophobicity (Table 5B).

The drug targets identified from docking simulations (see Tables 3 and 4) were compared and found to show (Table 6) good overall agreement with known experimental observations [18,19]. For Dasatinib, a Type I ligand, primary and secondary targets have been identified by Davis et al. [18,19] as ABL1 and SRC. Based on our simulations, estimated inhibition constants are 13.70 and 17.0 nM, respectively. Using these estimates we can assign ABL1 and SRC as primary and secondary targets (see Table 6). Similar observations

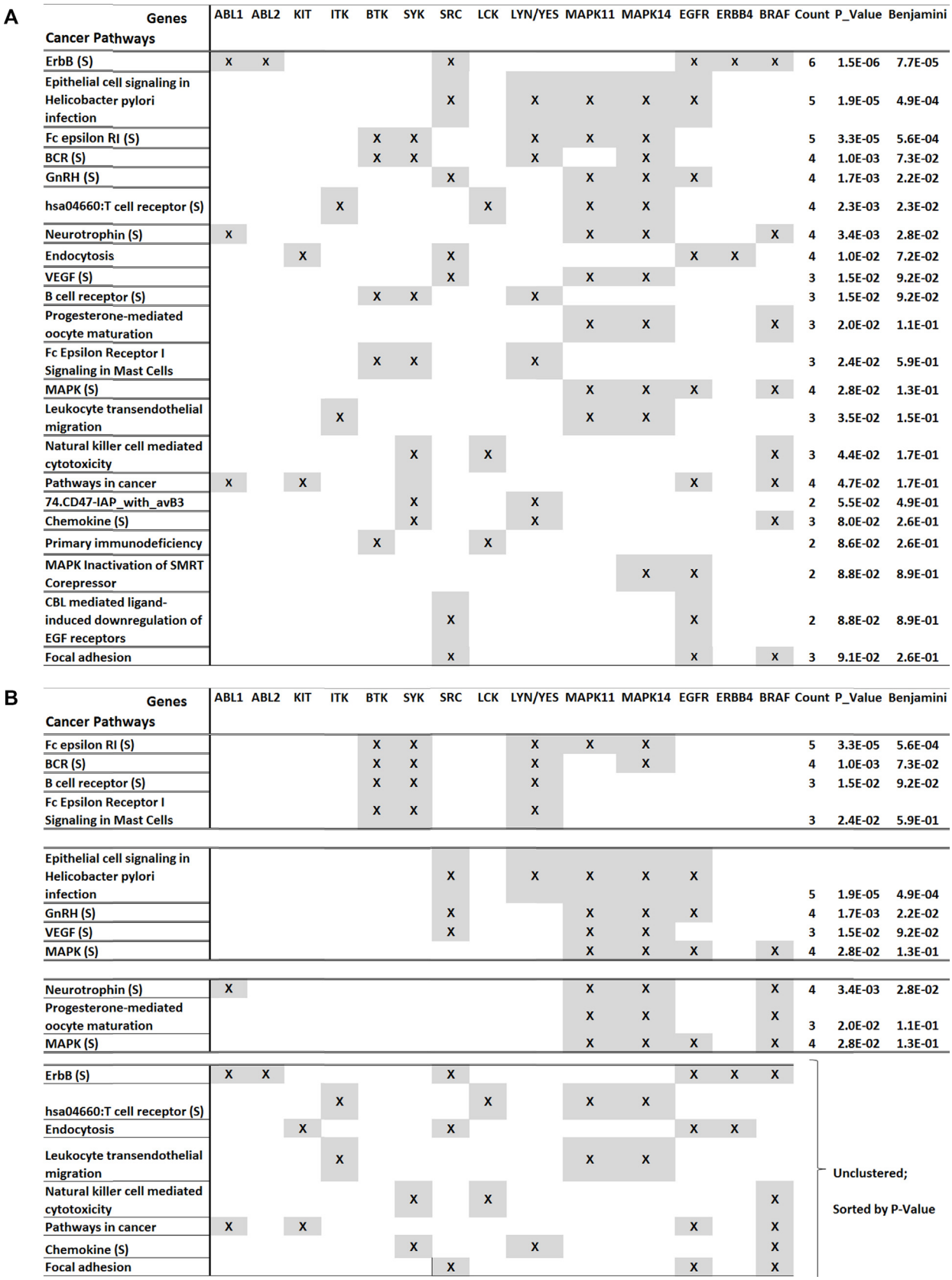

Fig. 5. (A) Pathway related information for the kinase genes studied in this work. DAVID web server (http://david.abcc.ncifcrf.gov/) was used for the calculation. Threshold option was set to count = 2, ease = 0.1. Please note that (S) denotes signaling pathway. (B) Pathways shown in A were clustered using kappa similarity with the following options: Similarity threshold = 0.50, classification: initial group membership = 3, final group membership = 3, multiple linkage threshold = 0.50. Enrichment was set to Ease with thresholds = 1.0. DAVID web server (http://david.abcc.ncifcrf.gov/) was used for the calculation. Please note that (S) denotes signaling pathway.

Table 5B
AutoDock estimated inhibition constants (nM) for the wild-type (WT) and mutated EGFR and B-Raf systems. Mutation numbering is based on the corresponding PDB structure. Note the 3-character (ex. STI) codes are the PDB assigned ligand IDs.

Gene	System		1N1	AQ4	B49	IRE	DB8	FMM	BAX	NIL	STI
	PDB		Dasatinib	Erlotinib	Sunitinib	Gefitinib	Bosutinib	Lapatinib	Sorafenib	Nilotinib	Imatinib
BRAF	1UWH	WT	17.34	489.68	273.30	75.61	63.81	0.12	1.69	1.84	4.50
		E500G	24.48	399.73	550.39	156.46	33.26	2.67	21.09	9.58	18.15
EGFR	1M17	WT	221.88	5580.00	3080.00	1520.00	1080.00	26.99	2150.00	81.45	82.59
		T766M	5470.00	8370.00	10840.00	3100.00	7720.00	134590.00	2690.00	349400.00	143.00
	1XKK	WT	48.33	741.11	888.34	608.52	583.94	0.14	1510.00	159.71	25.93
		T790M	15.69	188.58	2000.00	733.05	323.07	1.76	844.81	66.38	96.38

Table 6
Comparison of kinase drug targets from Davis et al. (Nature Biotech, V29, 1046, 2011) with docking simulations.

Drugs (type)	Primary	Secondary and others	Docking
Dasatinib(I)	ABL1	SRC	ABL1, SRC
Erlotinib(I)	EGFR		KIT, erbB4, EGFR
Sunitinib(I)	KIT	VEGFR2	ABL1, KIT
Gefitinib(I)	EGFR		KIT, BRAF
Lapatinib	ERBB2	EGFR	ABL1, BRAF, EGFR
Sorafenib(II)	VEGFR2	BRAF, FLT3	BRAF, ABL1
Bosutinib	ABL1	SRC	SRC, BRAF
Nilotinib(II)	ABL1	KIT	ABL1, BRAF SRC, KIT
Imatinib(II)	ABL1	KIT	ABL1/2, KIT, LCK

for other systems had also been reported in other studies as well [19,69]. The approximations involved in structure-based modeling, unfortunately, also lead to some discrepancies. For example, the predicted binding affinities for Lapatinib, a drug molecule that is known to mostly target inactive folds, agrees qualitatively but not quantitatively with the experiments [19]. The differences could be due to the approximations involved in modeling solvents and the dynamical effects. Since biological effects such as bioavailability are not inherently modeled, the binding ability for a ligand will be true only if the drug can successfully reach the active site. For example, bioavailability could be the reason why for some systems (EGFR-Gefitinib, PDB: 2ITZ), the co-crystallized ligand is not identified as the potent binder. Our future efforts will focus to include pathway information to address these issues.

Drug resistance arising from non-mutation mechanisms, as observed in EGFR systems, are also a major concern for both patient treatment and drug discovery. A 2010 study by Knight et al. [7] elegantly proposed a polypharmacology-based cautious approach on how to choose effective therapeutic targets that might lead to less toxic side-effects, but still could lead to complete inhibition. In advanced stage cancers, targeting multiple but related pathways appears to be a reasonable and necessary option. In an attempt to identify the pathway represented by kinase kinase receptor set, we used the functional annotation tool DAVID (http://david.abcc.ncifcrf.gov) to carry out an enrichment analysis. The list of common pathways for our gene set and the top clusters based on functional annotation are shown in Fig. 5A and B respectively. The top cluster (Fig. 5B) from our annotation results is the Fc epsilon RI signaling pathway. This pathway involves several signaling molecules and the participating mediators often contribute to inflammatory response events. This signaling pathway involves BTK, SYK, LYN and MAPK kinases from our set. As one can see from the KEGG pathway chart (KEGG: hsa04664; http://www.genome.jp/dbget-bin/www_bget?hsa04664) LYN is present upstream of BTK, SYK and MAPK. Our results indicate that Dasatinib or Lapatinib can be used as a drug to effectively target all of these kinases. Over-expression of ABL kinases have been observed in several breast cancer cases and often contributes to over-active cell growth [70]. The most common pathway (Fig. 5A) in our data set is the ErbB pathway. A KEGG pathway map (KEGG: hsa04012) shows that ABL genes are present downstream of EGFR. Experimental studies show that in breast cancer patients ABL activity seems to protect the activated EGFR from dephosphorylation by phosphatase [71,72]. So instead of inhibiting EGFR alone, it would be therapeutically rewarding to design a treatment that also includes ABL. Recent experiment by Yuan-Hung Lo et al. [24] tested this hypothesis and found evidence that a combined therapy that uses the dual kinase inhibitor Tykerb (Lapatinib) along with Imatinib showed improved therapeutic benefit and responsiveness for breast cancer treatment than using only Lapatinib. Our simulation (Table 4) results show that Lapatinib and Imatinib are compatible with the active-sites of ABL and EGFR, and the results qualitatively agree with the experimental observations. For example, our simulations show that Imatinib and Lapatinib being the potent inhibitors for ABL kinases (e.g. 3GVU: 0.27/0.31) and EGFR (1XKK: 25.93/0.14), respectively (Table 4). The additive effect of these drugs is beyond the scope of this study and needs experimental evaluation.

4. Conclusions

This study examined the interaction of nine commercially available kinase inhibitors with different kinases and selected mutants. These results show that the overwhelming majority of the docking experiments (90%, 17/19, RMSD <3 Å, see Table 2) were able to correctly identify the experimentally observed binding mode within the top 4 ranked conformations. The estimated inhibition constants also qualitatively follow the experimentally observed trend, thus highlighting the need for incorporating the characteristic conformation of the receptor (active or inactive) as a key early step in modeling kinase inhibition. One of the goals of this work is to see whether insight into cross-reactivity can be gained. Our results show that drug molecules that belong to Type II (e.g. Imatinib, see Tables 3 and 4) are more selective of the receptor structural features during binding when compared to Type I (e.g. Erlotinib, see Tables 3 and 4) ligands. This appears to follow from the fact that the active kinase conformations are more open and similar to each other when compared to the inactive ones. On the other hand, inactive active site functional group arrangements are different from each other and contributes to the selectivity of the drugs that favor these conformations [19]. To further explore the importance of the effect of kinase conformations on the selectivity, two key mutations, T315I and V299L, of the ABL gene (Table 5A) were examined. The docking results were able to qualitatively capture the effects of these key mutations for the most common first and second generation drugs used for treating CML patients with these mutations. Mutation effects were further examined by analyzing two additional systems, EGFR and B-Raf. The results show (Table 5B) that wild-type inhibitors lose key binding interactions upon receptor mutation; the weakening effect is reflected in the increase of the inhibition constants.

Kinase conformations retain clues on its selectivity for the inhibitors; hence a specific drug can be an effective inhibitor for two different receptors as long as these proteins can provide the

common conformation (e.g. inactive or active form) with similar active site residues. Also, the ligand Types can be used to understand cross-reactivity, but it is kinase family-specific and there are exceptions.

It is well-known that modeling kinase inhibition is a challenge. A recent study using a large data set was able to identify key issues such as structural features and the roles of ligand Types in kinase inhibition [19]. Using these observations as basis, a receptor-focused modeling and docking approach can correctly predict the kinase profile across a wide-range of kinase families. Ligand Types are often characterized based on their abilities to target the receptor ATP binding pocket when the activation loop is either present in the compact inactive or active open conformations. When the ligand contains an extra functional group(s), they fail to behave like the parent ligands [20]; hence the focus should be on the receptor. Even though there are experimental methods that are successful in identifying the common binding mode of kinase inhibitors, the concept of dynamic equilibrium during kinase activation cannot be ignored [20]. As elegantly pointed out by Chene [66] and others, the influence of pharmacokinetics [20] and slow off-rates continue to daunt [38] the prediction efforts. Specifically, ligands that target active kinase forms prefer to bind near the edge of the active site region and have portions of their molecule exposed to the solvent. In such cases, for example the simulation systems like **Gene**/PDB_ID (**BTK**/3k54, **SYK**/1XBB, **LYN**/2ZVA, **EGFR**/2ITZ), solvation effects may dominate the binding and hence an accurate description of these effects have to be included in the modeling to properly estimate binding affinities. As noted by Kitchen et al. [46], in such cases scoring and further ranking the docking solutions becomes difficult. In spite of these issues, by focusing on the structural features, these docking simulations were able to correctly predict binding modes and the trends in kinase inhibition. Future efforts should focus on three fronts: (a) Developing a receptor-based pharmacophore with spatial constraints to include receptor residues as exclusion areas to improve the estimation of binding affinities; (b) Inclusion of the impact of conformational changes of key residues in the active site on the estimation of inhibition constant; and (c) A combined approach of using analytic learning methods along with docking simulations to model cross-reactivity and toxicity [73].

Acknowledgements

This work was supported in part by funds from the National Cancer Institute/National Institutes of Health contract No. HHSN261200800001E. The content of this publication does not necessarily reflect the views of policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. There are no financial conflicts. The authors wish to thank the anonymous reviewer for the helpful comments.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jmgm.2014.12.007.

References

[1] R. Siegel, J. Ma, Z. Zou, A. Jemal, Cancer statistics, 2014, CA Cancer J. Clin. 64 (2014) 9–29.
[2] L.N. Johnson, Protein kinase inhibitors: contributions from structure to clinical compounds, Q. Rev. Biophys. 42 (2009) 1–40.
[3] J.T. Metz, E.F. Johnson, N.B. Soni, P.J. Merta, L. Kifle, P.J. Hajduk, Navigating the kinome, Nat. Chem. Biol. 7 (2011) 200–202.
[4] L.N. Johnson, R.J. Lewis, Structural basis for control by phosphorylation, Chem. Rev. 101 (2001) 2209–2242.
[5] G. Manning, D.B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, The protein kinase complement of the human genome, Science 298 (2002) 1912–1934.
[6] G. Manning, Genomic overview of protein kinases, in: WormBook: the online review of *C. elegans* biology, 2005, pp. 1–19.
[7] Z.A. Knight, H. Lin, K.M. Shokat, Targeting the cancer kinome through polypharmacology, Nat. Rev. Cancer 10 (2010) 130–137.
[8] E. Tamborini, L. Bonadiman, A. Greco, V. Albertini, T. Negri, A. Gronchi, et al., A new mutation in the KIT ATP pocket causes acquired resistance to imatinib in a gastrointestinal stromal tumor patient, Gastroenterology 127 (2004) 294–299.
[9] J.A. Engelman, K. Zejnullahu, T. Mitsudomi, Y. Song, C. Hyland, J.O. Park, et al., MET amplification leads to gefitinib resistance in lung cancer by activating ERBB3 signaling, Science 316 (2007) 1039–1043.
[10] J.M. Stommel, A.C. Kimmelman, H. Ying, R. Nabioullin, A.H. Ponugoti, R. Wiedemeyer, et al., Coactivation of receptor tyrosine kinases affects the response of tumor cells to targeted therapies, Science 318 (2007) 287–290.
[11] H. Daub, K. Specht, A. Ullrich, Strategies to overcome resistance to targeted protein kinase inhibitors, Nat. Rev. Drug Discov. 3 (2004) 1001–1010.
[12] B.J. Druker, M. Talpaz, D.J. Resta, B. Peng, E. Buchdunger, J.M. Ford, et al., Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia, N. Engl. J. Med. 344 (2001) 1031–1037.
[13] S. Redaelli, R. Piazza, R. Rostagno, V. Magistroni, P. Perini, M. Marega, et al., Activity of bosutinib, dasatinib, and nilotinib against 18 imatinib-resistant BCR/ABL mutants, J. Clin. Oncol. 27 (2009) 469–471.
[14] U.M. Martens, Small Molecules in Oncology, Springer, Heidelberg/London, 2010.
[15] B. Apsel, J.A. Blair, B. Gonzalez, T.M. Nazif, M.E. Feldman, B. Aizenstein, et al., Targeted polypharmacology: discovery of dual inhibitors of tyrosine and phosphoinositide kinases, Nat. Chem. Biol. 4 (2008) 691–699.
[16] L. Gossage, T. Eisen, Targeting multiple kinase pathways: a change in paradigm, Clin. Cancer Res. 16 (2010) 1973–1978.
[17] L.M. Wodicka, P. Ciceri, M.I. Davis, J.P. Hunt, M. Floyd, S. Salerno, et al., Activation state-dependent binding of small molecule kinase inhibitors: structural insights from biochemistry, Chem. Biol. 17 (2010) 1241–1249.
[18] M.W. Karaman, S. Herrgard, D.K. Treiber, P. Gallant, C.E. Atteridge, B.T. Campbell, et al., A quantitative analysis of kinase inhibitor selectivity, Nat. Biotechnol. 26 (2008) 127–132.
[19] M.I. Davis, J.P. Hunt, S. Herrgard, P. Ciceri, L.M. Wodicka, G. Pallares, et al., Comprehensive analysis of kinase inhibitor selectivity, Nat. Biotechnol. 29 (2011) 1046–1051.
[20] Y. Liu, N.S. Gray, Rational design of inhibitors that bind to inactive kinase conformations, Nat. Chem. Biol. 2 (2006) 358–364.
[21] M. Xu, L. Yu, B. Wan, L. Yu, Q. Huang, Predicting inactive conformations of protein kinases using active structures: conformational selection of type-II inhibitors, PLoS ONE 6 (2011) e22644.
[22] I. Kufareva, R. Abagyan, Type-II kinase inhibitor docking, screening, and profiling using modified structures of active kinase states, J. Med. Chem. 51 (2008) 7921–7932.
[23] T. Anastassiadis, S.W. Deacon, K. Devarajan, H. Ma, J.R. Peterson, Comprehensive assay of kinase catalytic activity reveals features of kinase inhibitor selectivity, Nat. Biotechnol. 29 (2011) 1039–1045.
[24] Y.H. Lo, P.C. Ho, H.J. Zhao, S.C. Wang, Inhibition of c-ABL sensitizes breast cancer cells to the dual ErbB receptor tyrosine kinase inhibitor lapatinib (GW572016), Anticancer Res. 31 (2011) 789–795.
[25] N.M. Levinson, S.G. Boxer, Structural and spectroscopic analysis of the kinase inhibitor bosutinib and an isomer of bosutinib binding to the Abl tyrosine kinase domain, PLOS ONE 7 (2012) e29828.
[26] E. Weisberg, P.W. Manley, W. Breitenstein, J. Bruggen, S.W. Cowan-Jacob, A. Ray, et al., Characterization of AMN107, a selective inhibitor of native and mutant Bcr-Abl, Cancer Cell 7 (2005) 129–141.
[27] B. Nagar, O. Hantschel, M.A. Young, K. Scheffzek, D. Veach, W. Bornmann, et al., Structural basis for the autoinhibition of c-Abl tyrosine kinase, Cell 112 (2003) 859–871.
[28] E. Salah, E. Ugochukwu, A. Barr, P. Mahagan, B. Shrestha, P. Savitsky, S. Knapp, The Crystal Structure of human ABL2 in complex with GLEEVEC (to be published).
[29] C.D. Mol, D.R. Dougan, T.R. Schneider, R.J. Skene, M.L. Kraus, D.N. Scheibe, et al., Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase, J. Biol. Chem. 279 (2004) 31655–31663.
[30] A.K. Kutach, A.G. Villasenor, D. Lam, C. Belunis, C. Janson, S. Lok, et al., Crystal structures of IL-2-inducible T cell kinase complexed with inhibitors: insights into rational drug design and activity regulation, Chem. Biol. Drug Des. 76 (2010) 154–163.
[31] D.J. Marcotte, Y.T. Liu, R.M. Arduini, C.A. Hession, K. Miatkowski, C.P. Wildes, et al., Structures of human Bruton's tyrosine kinase in active and inactive conformations suggest a mechanism of activation for TEC family kinases, Protein Sci. 19 (2010) 429–439.
[32] S. Atwell, J.M. Adams, J. Badger, M.D. Buchanan, I.K. Feil, K.J. Froning, et al., A novel mode of Gleevec binding is revealed by the structure of spleen tyrosine kinase, J. Biol. Chem. 279 (2004) 55827–55832.
[33] M.D. Jacobs, P.R. Caron, B.J. Hare, Classifying protein kinase structures guides use of ligand-selectivity profiles to predict inactive conformations: structure of lck/imatinib complex, Proteins 70 (2008) 1451–1460.

[34] N.K. Williams, I.S. Lucet, S.P. Klinken, E. Ingley, J. Rossjohn, Crystal structures of the Lyn protein tyrosine kinase domain in its Apo- and inhibitor-bound state, J. Biol. Chem. 284 (2009) 284–291.
[35] M. Getlik, C. Grutter, J.R. Simard, S. Kluter, M. Rabiller, H.B. Rode, et al., Hybrid compound design to overcome the gatekeeper T338 M mutation in cSrc, J. Med. Chem. 52 (2009) 3915–3926.
[36] M.A. Seeliger, B. Nagar, F. Frank, X. Cao, M.N. Henderson, J. Kuriyan, c-Src binds to the cancer drug imatinib with an inactive Abl/c-Kit conformation and a distributed thermodynamic penalty, Structure 15 (2007) 299–311.
[37] J. Stamos, M.X. Sliwkowski, C. Eigenbrot, Structure of the epidermal growth factor receptor kinase domain alone and in complex with a 4-anilinoquinazoline inhibitor, J. Biol. Chem. 277 (2002) 46265–46272.
[38] E.R. Wood, A.T. Truesdale, O.B. McDonald, D. Yuan, A. Hassell, S.H. Dickerson, et al., A unique structure for epidermal growth factor receptor bound to GW572016 (Lapatinib): relationships among protein conformation, inhibitor off-rate, and receptor activity in tumor cells, Cancer Res. 64 (2004) 6652–6659.
[39] C.H. Yun, T.J. Boggon, Y. Li, M.S. Woo, H. Greulich, M. Meyerson, et al., Structures of lung cancer-derived EGFR mutants and inhibitor complexes: mechanism of activation and insights into differential inhibitor sensitivity, Cancer Cell 11 (2007) 217–227.
[40] C. Qiu, M.K. Tarrant, S.H. Choi, A. Sathyamurthy, R. Bose, S. Banjade, et al., Mechanism of activation and inhibition of the HER4/ErbB4 kinase, Structure 16 (2008) 460–467.
[41] P.T. Wan, M.J. Garnett, S.M. Roe, S. Lee, D. Niculescu-Duvaz, V.M. Good, et al., Mechanism of activation of the RAF-ERK signaling pathway by oncogenic mutations of B-RAF, Cell 116 (2004) 855–867.
[42] D.S. Goodsell, G.M. Morris, A.J. Olson, Automated docking of flexible ligands: applications of AutoDock, J. Mol. Recognit. 9 (1996) 1–5.
[43] G.M. Morris, D.S. Goodsell, R. Huey, A.J. Olson, Distributed automated docking of flexible ligands to proteins: parallel applications of AutoDock 2.4, J. Comput. Aided Mol. Des. 10 (1996) 293–304.
[44] G.M. Morris, D.S. Goodsell, R.S. Halliday, R. Huey, W.E. Hart, R.K. Belew, et al., Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function, J. Comput. Chem. 19 (1998) 1639–1662.
[45] G.M. Morris, R. Huey, W. Lindstrom, M.F. Sanner, R.K. Belew, D.S. Goodsell, et al., AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility, J. Comput. Chem. 30 (2009) 2785–2791.
[46] D.B. Kitchen, H. Decornez, J.R. Furr, J. Bajorath, Docking and scoring in virtual screening for drug discovery: methods and applications, Nat. Rev. Drug Discov. 3 (2004) 935–949.
[47] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading, J. Comput. Chem. 31 (2010) 455–461.
[48] M.K. Gilson, H.X. Zhou, Calculation of protein–ligand binding affinities, Annu. Rev. Biophys. Biomol. Struct. 36 (2007) 21–42.
[49] C. Hetenyi, D. van der Spoel, Efficient docking of peptides to proteins without prior knowledge of the binding site, Protein Sci. 11 (2002) 1729–1737.
[50] C. Hetenyi, D. van der Spoel, Blind docking of drug-sized compounds to proteins with up to a thousand residues, FEBS Lett. 580 (2006) 1447–1450.
[51] R. Wang, L. Lai, S. Wang, Further development and validation of empirical scoring functions for structure-based binding affinity prediction, J. Comput. Aided Mol. Des. 16 (2002) 11–26.
[52] R.X. Wang, X.L. Fang, Y.P. Lu, S.M. Wang, The PDBbind database: collection of binding affinities for protein–ligand complexes with known three-dimensional structures, J. Med. Chem. 47 (2004) 2977–2980.
[53] R. Wang, X. Fang, Y. Lu, S. Wang, The PDBbind database: collection of binding affinities for protein–ligand complexes with known three-dimensional structures, J. Med. Chem. 47 (2004) 2977–2980.
[54] R. Wang, X. Fang, Y. Lu, C.Y. Yang, S. Wang, The PDBbind database: methodologies and updates, J. Med. Chem. 48 (2005) 4111–4119.
[55] J. Baxter, Local optima avoidance in depot location, J. Oper. Res. Soc. 32 (1981) 815–819.
[56] Hybrid Metaheuristics: An Emerging Approach to Optimization, Springer-Verlag, Berlin, Heidelberg, 2008.
[57] W.W. Chan, S.C. Wise, M.D. Kaufman, Y.M. Ahn, C.L. Ensinger, T. Haack, et al., Conformational control inhibition of the BCR-ABL1 tyrosine kinase, including the gatekeeper T315I mutant, by the switch-control inhibitor DCC-2036, Cancer Cell 19 (2011) 556–568.
[58] W. Pao, V.A. Miller, K.A. Politi, G.J. Riely, R. Somwar, M.F. Zakowski, et al., Acquired resistance of lung adenocarcinomas to gefitinib or erlotinib is associated with a second mutation in the EGFR kinase domain, PLoS Med. 2 (2005) e73.
[59] M.N. Balak, Y. Gong, G.J. Riely, R. Somwar, A.R. Li, M.F. Zakowski, et al., Novel D761Y and common secondary T790M mutations in epidermal growth factor receptor-mutant lung adenocarcinomas with acquired resistance to kinase inhibitors, Clin. Cancer Res. 12 (2006) 6494–6501.
[60] T. Kosaka, Y. Yatabe, H. Endoh, K. Yoshida, T. Hida, M. Tsuboi, et al., Analysis of epidermal growth factor receptor gene mutation in patients with non-small cell lung cancer and acquired resistance to gefitinib, Clin. Cancer Res. 12 (2006) 5764–5769.
[61] R.C. Wade, B.A. Luty, E. Demchuk, J.D. Madura, M.E. Davis, J.M. Briggs, et al., Simulation of enzyme-substrate encounter with gated active sites, Nat. Struct. Biol. 1 (1994) 65–69.
[62] Y. Shan, E.T. Kim, M.P. Eastwood, R.O. Dror, M.A. Seeliger, D.E. Shaw, How does a drug molecule find its target binding site, J. Am. Chem. Soc. 133 (2011) 9181–9183.
[63] J.D. Durrant, J.A. McCammon, Molecular dynamics simulations and drug discovery, BMC Biol. 9 (2011) 71.
[64] W. Chen, M.K. Gilson, S.P. Webb, M.J. Potter, Modeling protein–ligand binding by mining minima, J. Chem. Theor. Comput. 6 (2010) 3540–3557.
[65] F. Zuccotto, E. Ardini, E. Casale, M. Angiolini, Through the gatekeeper door: exploiting the active kinase conformation, J. Med. Chem. 53 (2010) 2681–2694.
[66] P. Chene, Can biochemistry drive drug discovery beyond simple potency measurements, Drug Discov. Today 17 (2012) 388–395.
[67] C.H. Yun, K.E. Mengwasser, A.V. Toms, M.S. Woo, H. Greulich, K.K. Wong, et al., The T790M mutation in EGFR kinase causes drug resistance by increasing the affinity for ATP, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 2070–2075.
[68] T. Niihori, Y. Aoki, Y. Narumi, G. Neri, H. Cave, A. Verloes, et al., Germline KRAS and BRAF mutations in cardio-facio-cutaneous syndrome, Nat. Genet. 38 (2006) 294–296.
[69] A.K. Ghose, T. Herbertz, D.A. Pippin, J.M. Salvino, J.P. Mallamo, Knowledge based prediction of ligand binding modes and rational inhibitor design for kinase drug discovery, J. Med. Chem. 51 (2008) 5149–5171.
[70] H.J. Zhao, F. Ou-Yang, I.F. Chen, M.F. Hou, S.S.F. Yuan, H.L. Chang, et al., Enhanced resistance to tamoxifen by the c-ABL proto-oncogene in breast cancer, Neoplasia 12 (2010), 214-U15.
[71] R.B. Birge, J.E. Fajardo, B.J. Mayer, H. Hanafusa, Tyrosine-phosphorylated epidermal growth factor receptor and cellular p130 provide high affinity binding substrates to analyze Crk-phosphotyrosine-dependent interactions in vitro, J. Biol. Chem. 267 (1992) 10588–10595.
[72] B. Tanos, A.M. Pendergast, Abl tyrosine kinase regulates endocytosis of the epidermal growth factor receptor, J. Biol. Chem. 281 (2006) 32714–32723.
[73] K.Y. Hsin, S. Ghosh, H. Kitano, Combining machine learning systems and multiple docking simulation packages to improve docking prediction reliability for network pharmacology, PLOS ONE 8 (2013) e83922.