Understanding the Molecular Basis of Tyrosine Kinase Intrinsic Promiscuity Toward Binding Type-I and Type-II Kinase Inhibitors. KIT versus INSR as a Case Study

Hassan M. Shallal and Wade A. Russu

Department of Pharmaceutics and Medicinal Chemistry, Thomas J. Long School of Pharmacy

University of the Pacific, Stockton, California 95211

Address correspondence to: Wade A. Russu, 3601 Pacific Avenue, Stockton, California 95211.

Tel: +1-2099462339. Fax: +1- 2099462160. Email: wrussu@pacific.edu

Abstract

The PDGFR kinase subfamily exhibits a significantly more promiscuous capacity of binding to tested kinase inhibitors, both type-I and type-II, compared to INSR subfamily. Time dependent ensembles of both KIT and INSR apoenzymes, representing PDGFR and INSR subfamilies respectively, were generated using molecular dynamics (MD) experiments in order to seek explanations of the different overall binding attitudes of KIT and INSR. Topologically speaking, DFG-hinge distance tends to be shorter in INSR compared to KIT whereas the DFG-αC-helix distance fluctuates more significantly in INSR than in KIT. In terms of energetics, the binding area tends to be energetically more self-stabilizing in INSR relative to KIT. These results suggest that the binding area in INSR is different from that of KIT in topological, energetic, and dynamic terms with different overall tendency to bind to kinase inhibitors. The relevance of the assumptions, suggested by this study, to kinase drug discovery is also discussed.

Keywords

Kinases, kinase inhibitors, promiscuity, binding area, conformational selection, induced fit

Abbreviations

Binding area (BA), molecular dynamics (MD), Kinase insert domain (KID), high throughput screening (HTS).

Introduction

Throughout the last two decades, the design of kinase targeted therapies, whether small molecules or antibodies, has been one of the most attractive approaches in achieving positive clinical impacts in several kinds of cancer and many other human diseases. Kinases are known to be linked to the pathophysiology, diagnosis and treatment of many kinds of cancer and are continuously implemented in developing novel and rational regimens not only for a given patient population, but also personalized based on an individual patient profile (Parikh, 2010; McDermott, 2009).

Quite often considered as a challenge, the design of a selective kinase inhibitor, one small molecule that inhibits one kinase or a single kinase superfamily more potently than others, has been one of the major driving forces of understanding and solving the x-ray or the NMR-derived structure of many kinases (Morphy, 2010; Smyth, 2009; Knight, 2005). Because ATP acts as the main cofactor for all the kinases performing different biological roles, they evolved many differences in their protein target recognition domains while maintaining conservation of their catalytic binding area which is tailored to bind to ATP and trigger a phosphorylation cascade of their specific protein targets. Accordingly, a small molecule kinase inhibitor, which is competing with ATP, is highly likely able to do so in the vicinity of the catalytic binding site of many rather than certain subset of kinases. Competing with ATP for kinases catalytic binding

areas categorizes a given inhibitor as a type-I kinase inhibitor (Huang, 2009). For ATP and a given type-I kinase inhibitor to bind potently with a given kinase catalytic area, the kinase needs to be chemically phosphorylated with significantly more DFG-in conformational ensembles; the side chain of the aspartic acid, starting the conserved DFG motif, extends towards the catalytic area whereas the side chain of the phenyl alanine, of the same motif, occupies the space between the DFG and the αC-helix. Such a structural situation seems inevitable for ATP and type-I kinase inhibitors to bind reasonably (Kirkland, 2009). When kinases are chemically dephosphorylated, the equilibrium throughout the conformational space starts shifting towards DFG-out rather than DFG-in states. This shift restricts ATP and type-I kinase inhibitors from binding potently to the kinases catalytic area. However, such a change also exposes a hydrophobic area next to the catalytic domain found to be less conserved among different kinases and hence offers a chance for designing more selective ATP non-competitive inhibitors, called type-II kinase inhibitors, which basically lock their target kinases in their DFG-out state. (Eglen, 2010; Jacobs, 2008; Alton, 2008; Bogoyevitch,, 2007; Liu, 2006).

The emergence of biochemically diverse highthroughput screening (HTS) technologies aiming at measuring the binding to or the inhibition of many kinases (phosphorylated or dephosphorylated) by several small molecules has helped generating datasets where patterns can be located and further studied (Milletti, 2010; Ma, 2008; Karaman, 2008; Bamborough, 2008; Federov, 2007). This report aims at explaining a very obvious pattern located in one of these HTS datasets in which dephosphorylated platelet derived growth factor receptor (PDFGR) subfamily members were found to globally as well as strongly bind to several type-I and type-II inhibitors compared to dephosphorylated insulin receptor (INSR) subfamily members (20 Karaman, 2008). In this report, we compared the binding area (BA) of both dephosphorylated KIT and dephosphorylated

INSR, as respective prototypes of PDGFR and INSR subfamilies, via analyzing time generated conformational ensembles using conventional molecular dynamics (MD) simulation experiments. The above comparison revealed interesting topological as well as energetic differences between the BA of KIT and INSR which can explain the difference in their overall promiscuity towards binding to type-I and type-II kinase inhibitors even without the inclusion of any kinase inhibitor in the study.

1. Methods

1.1 Computational models

The protein data bank was used to obtain 1T45 and 1IRK models (Mol, 2004; Hubbard, 1994; Berman, 2000). As a model for the DFG-out conformation of KIT kinase domain, 1T45 was chosen because it doesn't have any missing side chains and it has been recently selected in another MD study (Zou, 2008). All available dephosphorylated KIT x-ray structures have an incomplete kinase insert domain (KID) according to the technical difficulties brought by the bulky and charged KID during crystallization (Kani, 2009). However, the KID is still partially represented by 20 amino acids in 1T45. The 1IRK x-ray structure represents the inactive conformation of INSR. Despite missing the side chains of several amino acids, 1IRK has been used successfully in a recent computational study (Vashisth, 2010). Missing side chains were added to the INSR model using DeepView (Guex, 1997). Since the juxtamembrane region was not consistent in both starting models, it was deleted to allow the systems to be as close as possible in terms of size and studied loops. All the sequences used in the sequence alignment were obtained from Uniprot database and aligned using ClustalX 2.1 software (UniProt C, 2010; Chenna, 2003).

1.2 Molecular dynamics

MD experiments were conducted using Gromacs 4.5.2 software along with GROMOS96 43a1 force field (Scott, 1999; van Gunsteren, 1996; Hess, 2008; Van der Spoel, 2005). GROMOS96 43a1 force field was successfully used in simulating proteins in explicit solvent in several reports (Vreede, 2010; Senet, 2008). The starting structure of INSR was minimized in vacuum using steepest descent method for 1000 steps in order to relax the introduced side chains which were originally missing in the crystal structure (Arfken, 1995). Both INSR and KIT models were subjected to exactly the same computational treatment from this point on to the end of the analysis phase. Hydrogen atoms were treated as either virtual sites or heavy atoms which allowed the integration over a time step of 4 fs; a commonly used strategy to reduce the computational cost with negligible effects on the quality and the robustness of molecular dynamics (Feenstra, 1999; Sadiq, 2010; Chan, 2010; Bjelkmar, 2010). Each system was placed in a dodecahedron box with 1 nm distance between the protein and the edge of the box to allow for using periodic boundary conditions. Each protein was latter solvated using SPC water model known to be compatible with GROMOS96 force field (Berendsen, 1981). Both Na^+ and Cl^- ions were added in quantities enough to neutralize each system as well as maintaining a salt concentration of 0.15 M to simulate physiological and/or experimental buffer conditions.

The two solvated systems were energy minimized using steepest descent method; a force cutoff equal to 1000 kJ/mol was assigned to end the minimization. The water molecules and ions were successfully equilibrated around the position-restrained protein in each system by conducting a single 200 ps equilibration phase under an isobaric-isothermal (NPT) ensemble of 300 K and 1 bar. Temperature was maintained by the velocity-rescaling temperature coupling method whereas the pressure was equilibrated using weak coupling with Berendsen's barostat with an

isothermal compressibility of 4.6 $\times$ 10^{-5} bar^{-1} and a coupling constant τ_P of 1 ps (Bussi, 2007; Berendsen, 1984). Equilibration was conducted 4 independent times with velocities randomly generated in each time which provided 4 equilibrated versions of each system, KIT and INSR, with respective 4 different starting velocities to serve as input to the production phase. The production phase was run 4 independent times, each of which was 20 ns long. MD trajectories were produced under similar conditions to those of the equilibration phase except for freeing the protein atoms during production by releasing the position-restraints implemented during equilibration. In both equilibration and production phases, the parallel version of LINCS, P-LINCS, was used to constrain the bond lengths by setting an iteration of 1 and order of 6 because of using virtual sites along with larger time steps (Hess, 1997; Hess, 2008). Electrostatics were treated using PME (Particle-Mesh Ewald) with a short-range cutoff, rcoulomb, equal to 1 nm (Darden, 1993; Essmann, 1995). A grid neighborsearching was implemented using an rlist equal to 1 nm while a twin-range cutoff scheme was considered for the calculation of non-bonded interactions using a long range cutoff equal to 1.4 nm. Each simulation was run in parallel using a cluster of 5 processors, 2.66 MHz, allowing for a rate of 8.213 ns/day for the INSR system and 7.452 ns/day for the KIT system. With such rate, the 8 independent 20 ns simulations, four for each kinase, were completed in less than three days.

1.3 Analysis

The analyses of the trajectories retrieved from the above MD experiments as well as generating the figures shown within this report were performed through the cross implementation of Gromacs 4.5.2 analytical tools, VMD 1.8.7, Carma 1.0 package, and gnuplot 4.4.0, VEGA 2.4.0, and GraphPad prism 5.04 software (Hess, 2008; Glykos, 2006; Pedretti, 2004; Humphrey, 1996). Analysis of any trajectory considered frames generated after the first two nanoseconds all the

way to the last frame in order to average properties only after the system has been allowed to lose the memory of its initial conditions. There are two scenarios implemented in treating independent trajectories of each system, KIT or INSR. In one scenario, the frames of the four independent trajectories of each system were all used to calculate a specific property whereas in the other one, properties were analyzed through individual trajectories. It will be specified throughout the results section which scenario is used.

2. Results & Discussion

Although the general principle stating that type-II inhibitors are usually more selective than type-I, PDGFR subfamily members showed a non-selective behavior towards binding not only to type-II inhibitors, but also to many type-I kinase inhibitors (Karaman, 2008). As shown in Fig. 1, PDGFR subfamily displayed the highest promiscuity towards binding to the tested kinase inhibitors compared to other families whereas INSR displayed the least promiscuous behavior. According to this analysis, it can be argued that subfamilies like INSR and Tec (colored yellow in Fig 1) tend to be generally less promiscuous than PDGFR and Src subfamilies (colored red in Fig 1). It also seems plausible to question, according to the above analysis, whether some tyrosine kinase subfamilies tend to be intrinsically either less or more promiscuous than other subfamilies towards binding small molecule kinase inhibitors, both type-I and type-II.

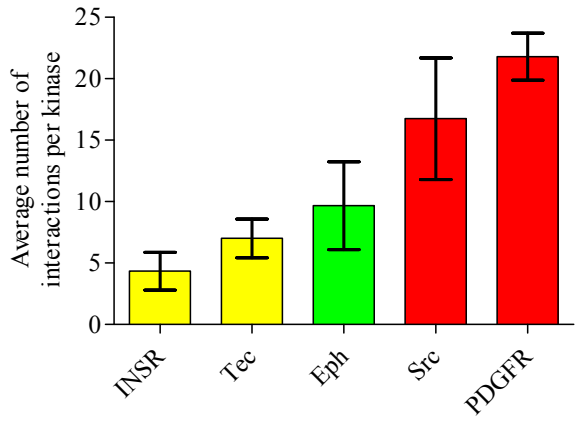

Figure 1 The average positive interaction observed for PDGFR, Src, Eph, Tec, and INSR tyrosine kinase subfamilies as retrieved from a reported HTS dataset [20]. Error bars represent standard deviation from the average value.

Table 1 shows an illustration of the phenomenon of overall promiscuity via comparing the binding of 21 kinase inhibitors against KIT, as a representative of the more promiscuous PDGFR subfamily, versus INSR, as a representative of the less promiscuous INSR subfamily retrieved from the same dataset mentioned above (Karaman, 2008). Dephosphorylated KIT showed positive interactions, $K_d \leq 10$ µM, with the 21 kinase inhibitors included in table 1. This tendency of KIT to bind to many type-II kinase inhibitors whereas INSR was incapable of comparably recognizing any could be justified by the general selectivity of type-II inhibitors. However, even for the four type-I kinase inhibitors recognized by INSR, the INSR-inhibitor K_d tends to be much higher than its respective KIT-inhibitor K_d reflecting an overall weaker ability of the dephosphorylated INSR to recognize not only type-II inhibitors but also type-I inhibitors compared to dephosphorylated KIT.

Mutations of the gatekeeper position, preceding the hinge region, are known to affect the binding to kinase inhibitors; especially type-II (Huang, 2009; Blencke, 2003). Bulky gatekeeper residues impose a steric barrier for type-II inhibitors to access the allosteric hydrophobic area next to the

ATP-binding site (Gorre, 2001). Accordingly, the replacement of electrostatically attractive and sterically reasonable threonine gatekeeper in PDGFR subfamily by methionine present in INSR subfamily members can have a deleterious effect on the binding of type-II inhibitors. It was also recently reported that the number of residues connecting the hinge region and the αD-helix (called the specificity linker) can affect the hydrogen bonding pattern between a given kinase and a given inhibitor (Katayama, 2008). It is commonly known as well as reported recently that the variation of the binding cavity volume would affect the binding of ligands to their protein targets (Saranya, 2009).

Table 1 The comparison of the binding of KIT and INSR to 21 kinase inhibitors as reported in a HTS dataset [20]. The values given are the K_d in nM. $\Delta\Delta G^o$ (Kcal.mol^{-1}) = R.T. ln [K_d(KIT) / K_d(INSR)] where R = 0.001985 Kcal.mol^{-1}.K^{-1} and T = 300°K.

Kinase inhibitor	Type	KIT	INSR	ΔΔG°
Sunitinib	I	0.37	500	-4.29
Dasatinib	I	0.62	> 10000	> -5.77
SU-14813	I	0.68	1200	-4.45
GW-786034	I	2.8	> 10000	> -4.87
CHIR-258/TKI-258	I	7.5	> 10000	> -4.28
Staurosporine	I	19	110	-1.05
PKC-412	I	220	> 10000	> -2.27
VX-680/MK-0457	I	240	630	-0.57
ZD-6474	I	260	> 10000	> -2.17
JNJ-7706621	I	1800	> 10000	> -1.02
CI-1033	I	7800	> 10000	> -0.15
ABT-869	II	2	> 10000	> -5.07
MLN-518	II	2.7	> 10000	> -4.89
AMG-706	II	3.7	> 10000	> -4.71
PTK-787	II	5.1	> 10000	> -4.51
AST-487	II	5.4	> 10000	> -4.48
Imatinib	II	14	> 10000	> -3.91
AZD-1152HQPA	II	17	> 10000	> -3.79
Sorafenib	II	31	> 10000	> -3.44
BIRB-796	II	170	> 10000	> -2.43
CHIR-265/RAF-265	II	200	> 10000	> -2.33

As illustrated by Fig. 2, the binding area (BA), generally targeted by either type-I or type-II kinase inhibitors, in both KIT and INSR is surrounded by four major structural elements; P-loop (sometimes called G-rich domain), αC-helix, hinge region (including also the gatekeeper residue

and the specificity linker), and the conserved DFG motif. The combination of the two binding sites of both type-I and type-II, throughout the current report, in one BA is intended for simplicity and based on that KIT is more promiscuous towards both type-I and type-II inhibitors. It is clear from the x-ray structure of KIT and INSR that their binding areas have close overall volume (Fig. 2). However, whether the similarity in the binding area volume between KIT and INSR is an intrinsic property that is maintained throughout the majority of their respective conformational spaces or an artifact of comparing two static x-ray structures resulted from crystallizing the two kinases under different conditions is a question that needs to be addressed by generating conformational ensembles via performing a conventional MD simulation experiments.

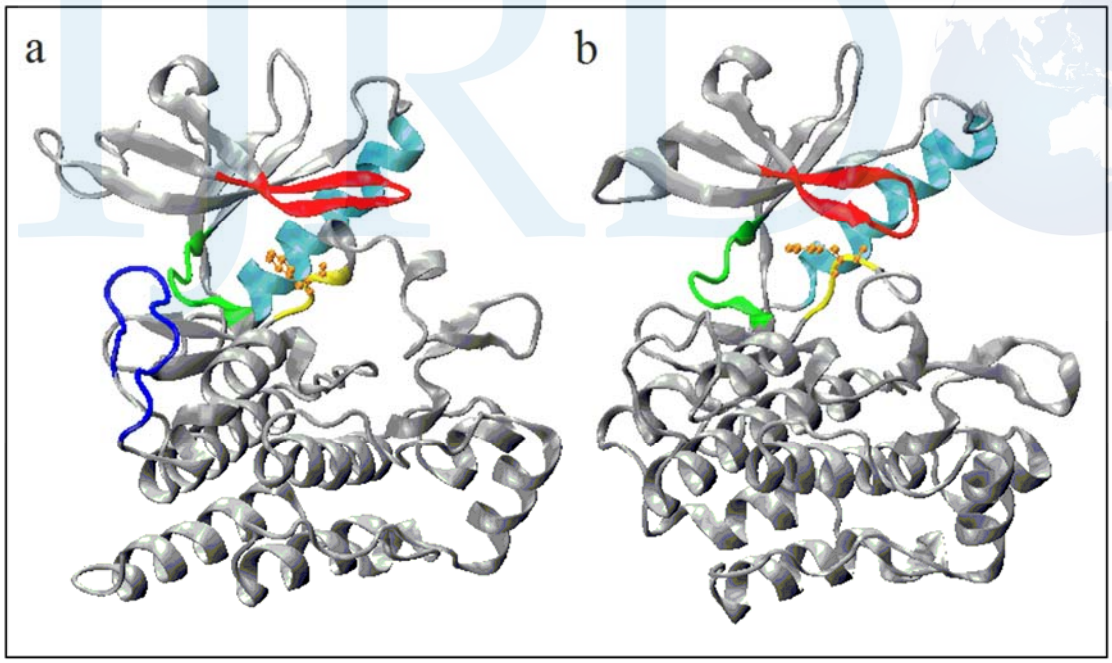

Figure 2 The BA of KIT (IT45) and INSR (1IRK). In both KIT (a) and INSR (b), P-loop is colored red, αC-helix is colored cyan, Hinge (including the gatekeeper and the specificity linker) is colored green, and the conserved DFG motif is colored yellow. The phenyl residue of the DFG motif is colored orange while being represented as balls and sticks. The KID domain in KIT is colored blue. The protein structure is represented by the New Cartoon graphical representation implemented in VMD.

The current report uses MD experiments neither to study the activation process of kinases nor to explain the effect of point mutations on inhibitor binding. It implements MD simulations to generally compare the topological, energetic and dynamic traits of the binding areas of two kinases, one that is more promiscuous towards binding small molecules than the other. So far, such an analysis hasn't been pursued in the discovery of kinase inhibitors or in understanding the molecular basis of kinases' intrinsic promiscuity regardless of the structure of small molecule inhibitors. Although it may be considered as a simplistic approach, we only focused on the BA of each kinase during our analysis in order to yield assumptions relevant to the recognition of the small molecules.

It was decided not to include kinase inhibitors in this study for many reasons. According to table 1, it would be difficult to either choose a prototype inhibitor of the 21 structurally diverse inhibitors showing preferential binding to KIT compared to INSR or to run independent computationally demanding simulations of both kinases with and without many kinase inhibitors. Not only is the latter option not computationally affordable, but it would also require docking and predicting the binding mode of many inhibitors with either INSR or KIT with subsequent introduction of a lot of uncertainty before even starting running the MD simulations. Including any kinase inhibitor in the simulation would introduce more or less bias in the topological and energetic behavior of the two kinases which would interfere with accurately investigating the major aim of this study; why is dephosphorylated INSR generally less capable of recognizing both type-I and type-II inhibitors than dephosphorylated KIT?

For each kinase, four independent 20 ns trajectories were produced, each starting from different initial velocity, in order to maximize the sampling of the conformational space of each kinase and increase the accuracy of the calculated properties in drawing conclusions. The time evolution

of the RMSD of the distances between the backbone atoms in each protein throughout each trajectory is illustrated in Fig. 3. Most of the trajectories start reaching equilibrium around after 2 ns of the production and that is why properties are averaged after the start of the third ns in each one. The RMSD values don't show severe distortions from the starting structure indicating the reliability of the trajectories to perform further analysis.

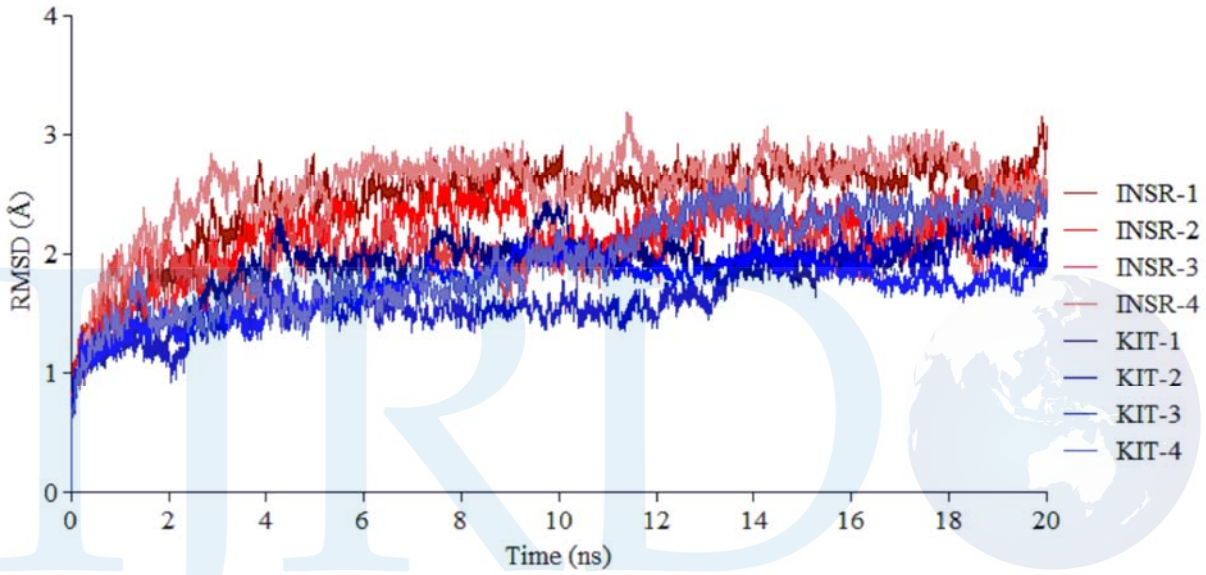

Fig. 3 The RMSD (Å) of the distances between the backbone atoms in each trajectory. INSR trajectories are represented in different grades of red color whereas KIT trajectories are represented as different grades of blue color.

In order to get a general feeling of the topological relations between the residues of the BA in each system, 2-dimensional distance maps were generated using all the frames retrieved from the four independent trajectories of each system so as to avoid any confusion that could be brought via comparing maps generated from eight trajectories (Fig. 4).

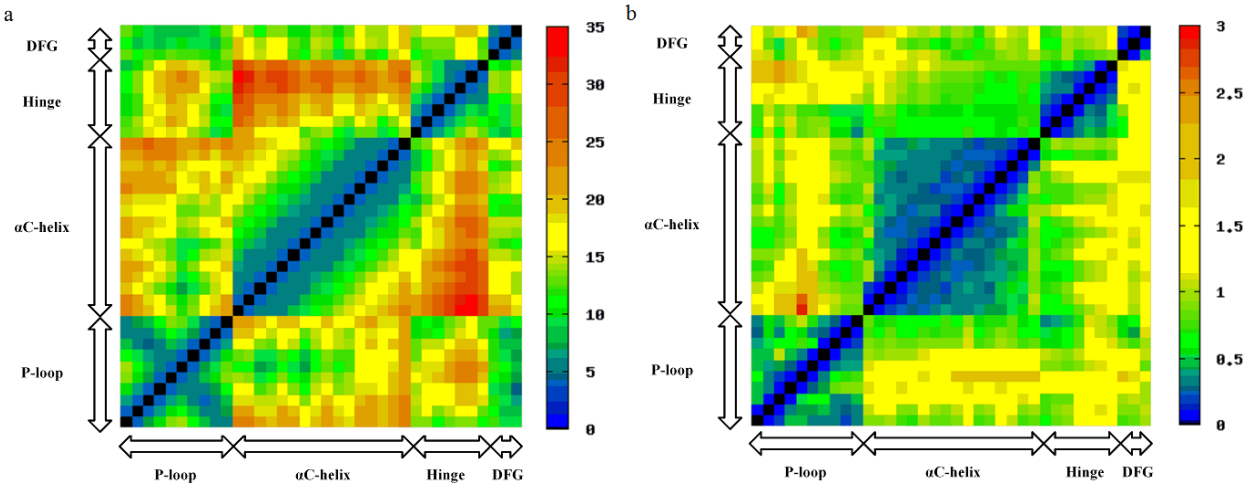

Figure 4 2-dimensional maps indicating: **a)** the average distance in Å between pairs of α-Carbons of the 37 residues comprising the binding area in KIT (above the diagonal) and INSR (below the diagonal) **b)** the RMSD from the average distance in Å between pairs of α-Carbons of the 37 residues comprising the binding area in KIT (above the diagonal) and INSR (below the diagonal). These distance maps were calculated based on analyzing the collective frames of the four independent trajectories.

In terms of the average distances between the α-carbon atoms of the residues included in the BA of each system, KIT and INSR are very similar except for the average distances between the DFG motif and the hinge region which tend to be generally shorter in INSR than in KIT. Comparing the RMSD from the average distances in both systems exposed more fluctuations in the distances between the DFG and both the hinge and the αC-helix in INSR compared to KIT. The last notion markedly indicates that the DFG motif tends to be more flexible and hence exhibits more frequent fluctuations between approaching either the hinge region or the αC-helix in INSR; a topological pattern that is minimized in KIT. In order to confirm the observed differences deduced from the inspection of the 2-dimensional distance maps; the evolution of certain distances was recorded for each individual trajectory. As illustrated by Fig. 5, the distance between the center of mass of the phenyl group of the DFG motif and the center of mass of the hinge region tends to be shorter in INSR than in KIT as was observed from the 2-dimensional distance map of the average distances.

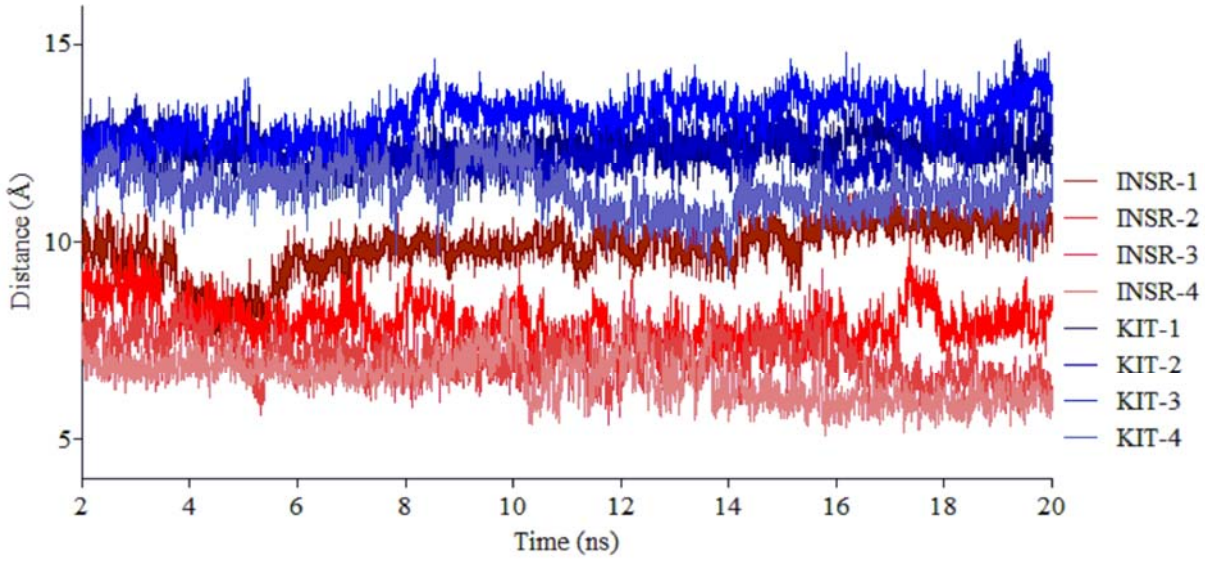

Figure 5 The time evolution of the distance in Å between the center of the mass of the phenyl ring of the DFG motif and the center of mass of the hinge region.

Fig. 6 compares the behavior of the distance between the phenyl group of the DFG motif and the αC-helix in the trajectories of INSR and KIT. It can be clearly observed that the DFG- αC-helix distance exhibits more fluctuations in INSR than in KIT confirming analysis of the 2-dimensional map of the RMSD from the average distances provided above. The implications of these topological differences on the inherent ability of each BA to recognize and then bind to small molecule kinase inhibitors, type-I and type-II, are very straightforward. Shorter distances between the DFG motif and the hinge region should eventually translate into smaller volume available for binding type-I inhibitors exactly as was observed for INSR (table 1). Such shorter distance should also represent a bottleneck rendering the access of type-II inhibitors to their hydrophobic site, between the DFG and the αC-helix, more challenging.

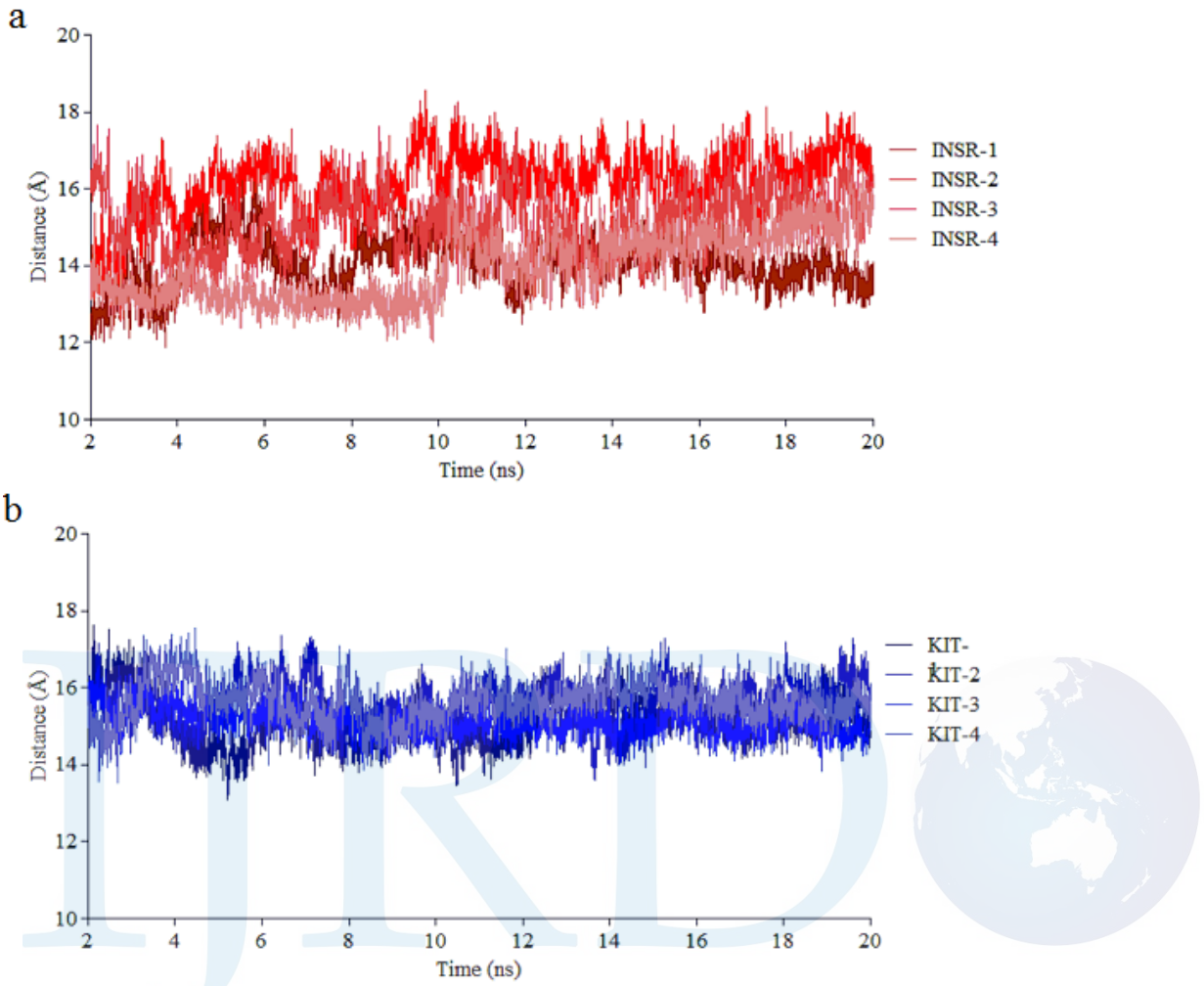

Figure 6 The time evolution of the distance in Å between the center of the mass of the phenyl ring of the DFG motif and the center of mass of the αC-helix in **a)** INSR and **b)** KIT.

Quite often reported in the literature, the rigidity of the host and/or the guest can be correlated to minimized entropic burdens on the recognition process between both and lead to higher affinity interactions (Zhong, 2000; Rekharsky, 2007; Chang, 2007; Mobley, 2009). The last common observation can raise a hypothesis that for two members of the same protein family, the flexibility inherent within each one's BA should be more or less correlated with their respective affinity towards binding small molecule inhibitors. It can be accordingly claimed that the flexibility of the DFG motif, indicated by its significantly fluctuating distance from the αC-helix,

observed in the case of INSR would present a higher entropic cost for the interaction with type-I kinase inhibitors; an effect that tends to be more compromised in the case of KIT with a more rigid BA than in INSR.

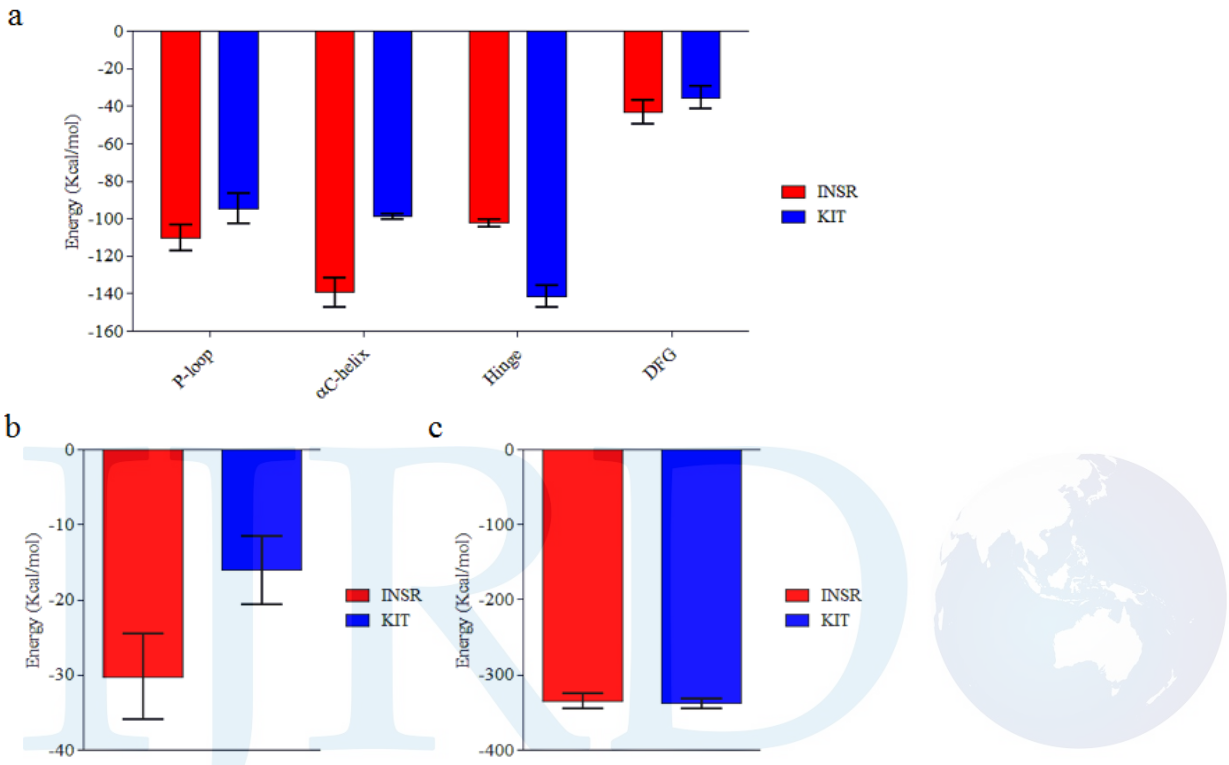

Figure 7 The analysis of the interaction energies between: **a)** the structural elements of the BA and the rest of the protein structure in INSR (red) and KIT (blue) **b)** the structural elements of the binding area in INSR (red) and KIT (blue) **c)** The whole BA and the rest of the protein in INSR (red) and KIT (blue). Error bars represent the standard deviation from the average energy value retrieved from the four independent trajectories.

The trajectory averaged interaction energies of different elements within the BA altogether as well as between the BA and the rest of the protein are plotted in Fig. 7. The interaction energy is the total energy calculated by adding the short-range Coulomb (electrostatic) and both the short and the long-range Lennard-Jones (van Der Waals) energies. The p-loop and the αC-helix have significantly higher interaction (more negative energy) whereas the hinge region has

significantly lower interaction (less negative energy) with the rest of the protein in INSR compared to KIT (Fig. 7a).

Central to the interaction with either type-I or type-II kinase inhibitor is the hinge region which serves as a common anchoring point to be recognized by either type. The observation that the interaction between the hinge region and the rest of the protein tends to be higher in KIT than in INSR prompted analyzing the hydrogen bonding pattern between the hinge region and the rest of the protein in both systems. It is important to notice that hydrogen bonds within the residues of the hinge region itself are not included within the above H-bond analysis. With a distance cutoff of 3.5 Å and an angle cutoff of 30°, 57 different hydrogen bonds were identified between the hinge region and the rest of the INSR with an average 5.991 hydrogen bonds per frame. On the KIT side, 77 different types of hydrogen bonds were identified between the hinge and the rest of the kinase with an average number of 9.903 hydrogen bonds per frame. Such a major difference in the average number of H-bonds can account for the enhanced interaction between hinge and the rest of the protein observed in KIT versus INSR. Twelve out of the additional 20 hydrogen bonds experienced by the hinge region in KIT encounters partnership with certain KID residues, 757-760, as a structural loop totally absent in INSR. Eight out of the above mentioned 12 hydrogen bonds were observed between the KID and the hydroxyl group of Y675 amino acid. Apparently, it seems that the coincident existence of the two hydrogen bonding capable structural features in KIT, neither exists in INSR, is a prerequisite for a network of hydrogen bonds which highly stabilizes the hinge region in KIT. Interestingly, it was found that both the KID and a tyrosine residue in position H5 exist not only in all PDGFR but also in 2 out of 3 members of the also more promiscuous VEGFR subfamily. The overall higher promiscuity of PDGFR and VEGFR subfamilies can therefore be partially correlated to the availability of these

two structural features which lead to an overall rigidification of the hinge region. The higher interaction between the KID and the hinge region also plays an indirect role in expanding the BA of KIT compared to INSR. Higher hydrogen bonding frequency between the KID and the Y675 helps push the hinge region farther away from the DFG motif and hence participates in enlarging the available volume for binding with type-I inhibitor and also allows for the entrance of type-II inhibitors to their target hydrophobic site. Again, the rigidity of the hinge region in KIT should impart less entropic burdens on recognizing both type-I and type-II inhibitors and becomes one of the hallmarks explaining KIT's overall promiscuity. Calculation of the interaction energy within the BA elements in each system also revealed a very interesting feature that differentiates the BA of KIT from INSR; the BA tends to be more energetically stable in INSR relative to KIT by about 15 Kcal/mol (Fig. 7b). The last notion represents a significant enthalpic contribution to the process of ligand binding given the fact that for the tested kinase inhibitors, the range of $\Delta\Delta G°$ comparing the binding affinity to KIT versus INSR is highly likely to lie within the above difference (table 1). These energy measurements indicate that the BA in INSR could be more energetically self-sufficient in contrast to that of KIT. It was also found that the interaction energy between the BA and the rest of the protein in both systems is within the same range (Fig. 7c). The latter notion supports our rationally focused analysis on the BA in both kinases considering recognition of small molecules.

We now use the above analysis to investigate two different mechanisms of molecular recognition among kinases and their small molecule binders; conformational selection and induced-fit theories. All the above topological and energetic MD-based observations can postulate that KIT exists in certain conformational ensembles that are intrinsically more hospitable towards binding small molecules and hence becomes more promiscuous than INSR whose conformational

ensembles are generally less topologically competent and more energetically self-satisfying. That is to say, small molecule kinase inhibitors are highly likely to be able to select for more accommodating conformers in the vicinity of KIT rather than INSR. However, it is known that thermodynamically oriented conformational selection is not the only theory established for molecular recognition. It would be unbiased to also consider the induced fit theory which is another kinetically oriented approach explaining molecular recognition. Such theory simply states that a given protein binding site could further optimize the interaction with a bound ligand via moving its flexible parts and orienting its side chains. Whether one theory may prevail over another or both team together in the control of a given protein-ligand recognition scenario is still a matter of scientific debate (Boehr, 2009). The above study attempted to offer what could be called experimentally (table 1) and computationally (MD analysis) supported assumptions to explain why and how a given kinase could be more promiscuous towards binding to small molecule inhibitors than another. The above binding site oriented MD analysis revealed both thermodynamically as well as kinetically oriented justifications of why and how KIT can interact more favorably with several structurally diverse small kinase inhibitors than INSR. However, the overall experimentally supported higher promiscuity of KIT versus that of INSR may shift in the opposite direction if the kinase inhibitor chemical space is further explored and entirely different kinase binding capable scaffolds become incorporated into future HTS analyses.

The above drawn postulations seem very pertinent to the process of early phase design of kinase inhibitors. For instance, it could be argued that if a research team is interested in discovering selective small molecule kinase inhibitors, it may be easier to discover selective dephosphorylated PDGFR subfamily inhibitors that don't recognize dephosphorylated INSR members. On the other side, it may be more challenging to design selective dephosphorylated

INSR subfamily inhibitors that don't interact with dephosphorylated PDGFR members. As a consequence, exploring other remote sites within the kinase domain may actually offer better opportunities to selectively target INSR or other less promiscuous subfamilies without affecting PDGFR or other more promiscuous families. If targeting the allosteric binding site of the dephosphorylated form usually allows the relatively easier design of selective PDGFR subfamily inhibitors, it may be possible to design selective INSR inhibitors by targeting the same allosteric binding site in the phosphorylated rather than in the dephosphorylated form; a direction that conflicts in principal with the conventional belief in designing selective inhibitors by targeting the dephosphorylated kinase.

Here we have highlighted the potential importance of considering the selectivity problem from the perspective of the kinase target, which translates to careful consideration of target kinase selection in a drug discovery program.

Acknowledgments

The authors would like to deeply thank the POD (Penguin On Demand) staff for providing a great deal of technical support while running the MD simulations using their high performance computing facility.

Conflict of Interest Statement

The authors declare that there is no conflict of interest regarding the publication of this article.

References

Aleksandrov A, Simonson T (2010) Molecular dynamics simulations show that conformational selection governs the binding preferences of imatinib for several tyrosine kinases. Journal of Biological Chemistry 285 (18):13807-13815

Alton GR, Lunney EA (2008) Targeting the unactivated conformations of protein kinases for small molecule drug discovery. Expert Opinion on Drug Discovery 3 (6):595-605

38. Arfken G (1995) The Method of Steepest Descents. In: Mathematical Methods for Physicists. 3rd edn. Academic Press, Orlando, FL, pp 428–436

Bamborough P, Drewry D, Harper G, Smith GK, Schneider K (2008) Assessment of chemical coverage of kinome space and its implications for kinase drug discovery. Journal of Medicinal Chemistry 51 (24):7898-7914

Berman H, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov I, Bourne P (2000) The Protein Data Bank. Nucleic Acids Research 28 (1):235-242

Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J (1981) Interaction models for water in relation to protein hydration. In: Pullman B (ed) Intermolecular Forces. Reidel, Dordrecht), p 331

Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. Journal of Chemical Physics 81 (8):3684-3690

Blencke S, Ullrich A, Daub H (2003) Mutation of threonine 766 in the epidermal growth factor receptor reveals a hotspot for resistance formation against selective tyrosine kinase inhibitors. Journal of Biological Chemistry 278 (17):15435-15440

Bjelkmar Pr, Larsson P, Cuendet M, Hess B, Lindahl E (2010) Implementation of the CHARMM Force Field in GROMACS: Analysis of Protein Stability Effects from Correction Maps, Virtual Interaction Sites, and Water Models. Journal of Chemical Theory and Computation 6 (2):459-466

Boehr DD, Nussinov R, Wright PE (2009) The role of dynamic conformational ensembles in biomolecular recognition. Nature Chemical Biology 5 (11):789-796

Bogoyevitch MA, Fairlie DP (2007) A new paradigm for protein kinase inhibition: blocking phosphorylation without directly targeting ATP binding. Drug Discovery Today 12 (15-16):622-633

Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. The Journal of Chemical Physics 126 (1):014101

Chan DI, Tieleman DP, Vogel HJ (2010) Molecular dynamics simulations of beta-ketoacyl-, beta-hydroxyacyl-, and trans-2-enoyl-acyl carrier proteins of Escherichia coli. Biochemistry 49 (13):2860-2868. doi:10.1021/bi901713r

Chang CE, Chen W, Gilson MK (2007) Ligand configurational entropy and protein binding. Proc Natl Acad Sci U S A 104 (5):1534-1539. doi:0610494104 [pii] 10.1073/pnas.0610494104

Chenna R, Sugawara H, Koike T, Lopez R, Gibson T, Higgins D, Thompson J (2003) Multiple sequence alignment with the Clustal series of programs. Nucl Acids Res 31 (13):3497-3500

Darden T, York D, Pedersen L (1993) Particle mesh {Ewald: An N.log(N)} method for {Ewald} sums in large systems. J Chem Phys 98 (12):10089-10092

Eglen RM, Reisine T (2010) Human kinome drug discovery and the emerging importance of atypical allosteric inhibitors. Expert Opinion on Drug Discovery 5 (3):277-290

Essmann U, Perera L, Berkowitz M, Darden T, Lee H, Pedersen L (1995) A smooth particle mesh Ewald method. The Journal of Chemical Physics 103 (19):8577-8593

Fedorov O, Marsden B, Pogacic V, Rellos P, MÃ¼ller S, Bullock AN, Schwaller J, SundstrÃ¶m M, Knapp S (2007) A systematic interaction map of validated kinase inhibitors with Ser/Thr kinases. Proceedings of the National Academy of Sciences of the United States of America 104 (51):20523-20528

Feenstra KA, Hess B, Berendsen H (1999) Improving efficiency of large time-scale molecular dynamics simulations of hydrogen-rich systems. Journal of Computational Chemistry 20 (8):786-798

Glykos NM (2006) Software news and updates carma: A molecular dynamics analysis program. Journal of Computational Chemistry 27 (14):1765-1768

Gorre ME, Mohammed M, Ellwood K, Hsu N, Paquette R, Nagesh Rao P, Sawyers CL (2001) Clinical resistance to STI-571 cancer therapy caused by BCR-ABL gene mutation or amplification. Science 293 (5531):876-880

Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18 (15):2714-2723

Hess B, Bekker H, Berendsen H, Fraaije J (1997) LINCS: A linear constraint solver for molecular simulations. Journal of Computational Chemistry 18 (12):1463-1472

Hess B, Kutzner C, van der Spoel D, Lindahl E (2008) GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. Journal of Chemical Theory and Computation 4 (3):435-447

Hess B (2008) P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. J Chem Theory Comput 4 (1):116-122

Huang D, Zhou T, Lafleur K, Nevado C, Caflisch A (2009) Kinase selectivity potential for inhibitors targeting the ATP binding site: A network analysis. Bioinformatics 26 (2):198-204

Hubbard SR, Wei L, Ellis L, Hendrickson WA (1994) Crystal structure of the tyrosine kinase domain of the human insulin receptor. Nature 372 (6508):746-754. doi:10.1038/372746a0

Humphrey W, Dalke A, Schulten K (1996) VMD: Visual molecular dynamics. Journal of Molecular Graphics 14 (1):33-38

Jacobs MD, Caron PR, Hare BJ (2008) Classifying protein kinase structures guides use of ligand-selectivity profiles to predict inactive conformations: Structure of lck/imatinib complex. Proteins: Structure, Function and Genetics 70 (4):1451-1460

Kani RS (2009) Structure-Based Design and Characterization of Axitinib. In: Kinase Inhibitor Drugs. John Wiley & Sons Inc., Hoboken, New Jersey, p 171

Karaman MW, Herrgard S, Treiber DK, Gallant P, Atteridge CE, Campbell BT, Chan KW, Ciceri P, Davis MI, Edeen PT, Faraoni R, Floyd M, Hunt JP, Lockhart DJ, Milanov ZV, Morrison MJ, Pallares G, Patel HK, Pritchard S, Wodicka LM, Zarrinkar PP (2008) A quantitative analysis of kinase inhibitor selectivity. Nature Biotechnology 26 (1):127-132

Katayama N, Orita M, Yamaguchi T, Hisamichi H, Kuromitsu S, Kurihara H, Sakashita H, Matsumoto Y, Fujita S, Niimi T (2008) Identification of a key element for hydrogen-bonding patterns between protein kinases and their inhibitors. Proteins: Structure, Function and Genetics 73 (4):795-801

Kirkland LO, McInnes C (2009) Non-ATP competitive protein kinase inhibitors as anti-tumor therapeutics. Biochemical Pharmacology 77 (10):1561-1571

Knight ZA, Shokat KM (2005) Features of selective kinase inhibitors. Chemistry and Biology 12 (6):621-637

Liu Y, Gray NS (2006) Rational design of inhibitors that bind to inactive kinase conformations. Nature Chemical Biology 2 (7):358-364

Ma H, Deacon S, Horiuchi K (2008) The challenge of selecting protein kinase assays for lead discovery optimization. Expert Opinion on Drug Discovery 3 (6):607-621

McDermott U, Settleman J (2009) Personalized cancer therapy with selective kinase inhibitors: An emerging paradigm in medical oncology. Journal of Clinical Oncology 27 (33):5650-5659

Milletti F, Vulpetti A (2010) Predicting polypharmacology by binding site similarity: From kinases to the protein universe. Journal of Chemical Information and Modeling 50 (8):1418-1431

Mobley DL, Dill KA (2009) Binding of small-molecule ligands to proteins: "what you see" is not always "what you get". Structure 17 (4):489-498. doi:S0969-2126(09)00092-6 [pii]10.1016/j.str.2009.02.010

Mol CD, Dougan DR, Schneider TR, Skene RJ, Kraus ML, Scheibe DN, Snell GP, Zou H, Sang BC, Wilson KP (2004) Structural basis for the autoinhibition and STI-571 inhibition of c-Kit tyrosine kinase. J Biol Chem 279 (30):31655-31663

Morphy R (2010) Selectively nonselective kinase inhibition: striking the right balance. Journal of Medicinal Chemistry 53 (4):1413-1437

Parikh K, Peppelenbosch MP (2010) Kinome profiling of clinical cancer specimens. Cancer Research 70 (7):2575-2578

Pedretti A, Villa L, Vistoli G (2004) VEGA - An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. Journal of Computer-Aided Molecular Design 18 (3):167-173

Rekharsky MV, Mori T, Yang C, Ko YH, Selvapalam N, Kim H, Sobransingh D, Kaifer AE, Liu S, Isaacs L, Chen W, Moghaddam S, Gilson MK, Kim K, Inoue Y (2007) A synthetic host-guest system achieves avidin-biotin affinity by overcoming enthalpy-entropy compensation. Proc Natl Acad Sci U S A 104 (52):20737-20742. doi:0706407105 [pii]10.1073/pnas.0706407105

Sadiq SK, De Fabritiis G (2010) Explicit solvent dynamics and energetics of HIV-1 protease flap opening and closing. Proteins 78 (14):2873-2885. doi:10.1002/prot.22806

Saranya N, Selvaraj S (2009) Variation of protein binding cavity volume and ligand volume in protein-ligand complexes. Bioorganic and Medicinal Chemistry Letters 19 (19):5769-5772

Scott W, Hunenberger P, Tironi I, Mark A, Billeter S, Fennen J, Torda A, Huber T, Kruger P, van Gunsteren W (1999) The GROMOS Biomolecular Simulation Program Package. The Journal of Physical Chemistry A 103 (19):3596-3607

Senet P, Maisuradze G, Foulie C, Delarue P, Scheraga H (2008) How main-chains of proteins explore the free-energy landscape in native states. Proceedings of the National Academy of Sciences 105 (50):19708-19713

Smyth LA, Collins I (2009) Measuring and interpreting the selectivity of protein kinase inhibitors. Journal of Chemical Biology 2 (3):131-151

UniProt C (2010) The Universal Protein Resource (UniProt) in 2010. Nucleic Acids Research 38 (suppl 1):D142-D148

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJC (2005) GROMACS: Fast, flexible, and free. J Comput Chem 26 (16):1701-1718

van Gunsteren WF, Billeter SR, Eising AA, HÃ¼nenberger PH, KrÃ¼ger P, Mark AE, Scott WRP, Tironi IG (1996) Biomolecular Simulation: The {GROMOS96} manual and userguide. Hochschuleverlag AG an der ETH Zurich,

Vashisth H (2010) Molecular Simulation Studies of the Insulin Receptor Family. Drexel University

Vreede J, Juraszek J, Bolhuis P (2010) Predicting the reaction coordinates of millisecond light-induced conformational changes in photoactive yellow protein. Proceedings of the National Academy of Sciences 107 (6):2397-2402

Zhong D, Douhal A, Zewail AH (2000) Femtosecond studies of protein-ligand hydrophobic binding and dynamics: Human serum albumin. Proceedings of the National Academy of Sciences of the United States of America 97 (26):14056-14061

Zou J, Wang YD, Ma FX, Xiang ML, Shi B, Wei YQ, Yang SY (2008) Detailed conformational dynamics of juxtamembrane region and activation loop in c-Kit kinase activation process. Proteins 72 (1):323-332. doi:10.1002/prot.21928