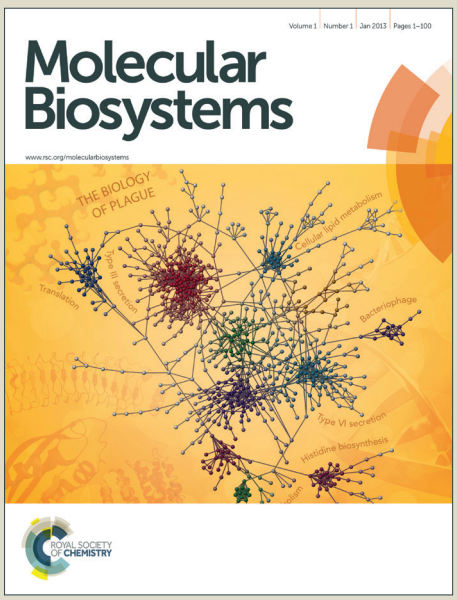
Molecular BioSystems

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/molecularbiosystems

Journal Name

ARTICLE

Small-world networks of residue interactions in the Abl kinase complexes with cancer drugs: Topology of allosteric communication pathways can determine drug resistance effects

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

A. Tse[a] and G. M. Verkhivker[a,b,†]

The human protein kinases play a fundamental regulatory role in orchestrating functional processes in complex cellular networks. Understanding how conformational equilibrium between functional kinase states can be modulated by ligand binding or mutations is critical for quantifying molecular basis of allosteric regulation and drug resistance. In this work, molecular dynamics simulations of the Abl kinase complexes with cancer drugs (Imatinib and Dasatinib) were combined with structure-based network modeling to characterize dynamics of the residue interaction networks in these systems. The results have demonstrated that structural architecture of kinase complexes can produce a small-world topology of the interaction networks. Our data have indicated that specific Imatinib binding to a small number of highly connected residues could lead to network-bridging effects and allow for efficient allosteric communication, which is mediated by a dominant pathway sensitive to the unphosphorylated Abl state. In contrast, Dasatinib binding to the active kinase form may activate a broader ensemble of allosteric pathways that are less dependent on the phosphorylation status of Abl and provide a better balance between the efficiency and resilience of signaling routes. Our results have unveiled how differences in the residue interaction networks and allosteric communications of the Abl kinase complexes can be directly related to drug resistance effects. This study offers a plausible perspective on how efficiency and robustness of the residue interaction networks and allosteric pathways in kinase structures may be associated with protein responses to drug binding.

Introduction

Protein kinase genes are signalling switches with a conserved catalytic domain that are regulated via phosphorylation of the activation loops, through autoinhibition or by allosteric activation that enable the kinase domain to adopt a catalytically competent conformation[1-10]. Recent studies of the structure and regulation of the Abl kinase domains[11-13] have emphasized that kinase activation and various functions can be orchestrated and supported by conformational changes in the key functional regions of the catalytic domain: the Asp-Phe-Gly (DFG) motif, the glycine-rich P-loop, the regulatory αC-helix, and the activation loop (A-loop). A dynamic equilibrium between distinct conformational states is central for kinase regulation and has been exploited in drug discovery of type 1 inhibitors (Dasatinib) that target the active DFG-in conformation of the kinase domain[14], and type 2 inhibitors (Imatinib) that recognize the inactive DFG-out kinase conformation[15-17]. Systematic analyses of type 2 inhibitors that stabilize an Imatinib-like inactive conformation have revealed that Abl-selective ligands preferentially bind to the unphosphorylated kinase form[18]. This study has also pointed out to a relationship between Imatinib sensitivity to the P-loop mutations and binding preferences for the unphosphorylated state, suggesting that ligand-mediated allosteric coupling between the P-loop and A-loop regions could modulate binding preferences of specific inhibitors. Dasatinib is a type 1 inhibitor that targets the catalytically competent active ABL conformation[14] and has a broad spectrum of activity against the SRC, CSK, TEC, and EPH families of tyrosine kinases[19-23]. NMR studies have shown that Dasatinib binds an active-like conformation of an unphosphorylated Abl and suggested that DFG-in conformation may be an obligatory kinase form for a productive inhibitor binding. Recent functional experiments have demonstrated that Dasatinib binding may be phosphorylation state-independent but exhibit conformation-specific binding preferences towards active Abl state[24]. Computational studies have investigated molecular mechanisms of protein kinases and structural effects of drug resistance mutations[25-34]. Molecular dynamics (MD) simulations and free energy calculations have shown that conformational selection and the stability difference between the inactive DFG-out and active DFG-in conformations may be the primary determinant underlying Imatinib specificity[30,31]. MD simulations and free energy calculations combined with

[a] Graduate Program in Computational and Data Sciences, Department of Computational Sciences, Schmid College of Science and Technology, Chapman University, Orange, CA 92866, USA
[b] Chapman University School of Pharmacy, Irvine, CA 92618, USA
† Corresponding author. Email: verkhivk@chapman.edu

isothermal titration calorimetry have quantified energetics of conformational transitions in Abl kinase[32], confirming that a more favorable kinetic accessibility and thermodynamic stability of the DFG-out conformation in Abl may determine Imatinib selectivity. A systematic investigation based on extensive molecular simulations has revealed that Dasatinib shows a marked preference for binding to the active state of the Abl kinase[35]. Understanding of mechanisms by which conformational equilibrium between functional kinase states can be modulated upon inhibitor binding is critical for quantifying molecular basis of allosteric regulation and drug resistance. Despite wealth of structural and computational studies, the dynamic nature of functional kinase complexes with inhibitors can often hinder the molecular details underlying ligand-induced modulation of the residue interaction networks and allosteric effects. Quantifying allosteric changes due to cooperative interactions and collective influence of multiple residues remains a challenging task in biophysical simulations of macromolecular complexes.

Protein structure network (PSN) analysis using a graph-based representation, where the nodes represent residues and the edges their contacts, can yield a convenient description of protein dynamics and stability[36-40] by simultaneously capturing the effects of global topological rearrangements and local side-chain interactions into computational analyses. Protein structure topologies could often produce small-world networks which balance a high local connectivity of residue nodes with a smaller number of long-range interactions, giving rise to a high degree of interaction cooperativity. Small-world network models of proteins and topology-based network parameters describing node centrality (degree, closeness, and betweenness) have been exploited to predict protein-protein interactions[41,42], protein-DNA interfaces[43] , ligand binding sites[44-46], and catalytic residues in enzymes[47]. These studies have linked small-world organization of protein structure networks with structural stability and high connectivity of functional residues, particularly indicating that residues involved in short path length communications could mediate signaling[48]. Graph-based protein networks that incorporated topology-based residue connectivity and contact maps of residues cross-correlations obtained from MD simulations have provided important insights into structural mechanisms underlying allosteric interactions and communication pathways in various protein systems[49-51]. The network organization is also determined by the average degree correlation between nodes, so that complex networks may be either disassortative, where the links between nodes with similar networking parameters are systematically prevented, or assortative, where these links are enhanced[52]. While a disassortative organization allows for rapid signal transmission between segregated modules but may produce more vulnerable to random attacks networks, assortative networks may sacrifice the efficiency of long-range communication to achieve a greater resilience against random perturbations. A dynamic balance between disassortative organization of spatially separated, highly connected hubs and assortative features that are responsible for the network integrity and functional redundancy is believed to be maintained in the protein environment within the cell[53].

In this work, we integrate MD simulations with the network modelling to characterize organization and dynamics of the residue interaction networks and allosteric communications pathways in Abl complexes with cancer drugs Imatinib and Dasatinib. We show that Imatinib binding to a small number of highly connected residues could allow for efficient allosteric communication mediated by a dominant single ensemble of short path length routes that are "addicted" to the unphosphorylated state of the A-loop. In contrast, Dasatinib binding may activate a broad ensemble of alternative allosteric pathways that are essentially independent on the phosphorylation status of Abl and may provide a better balance between the efficiency and resilience of signalling routes. We also demonstrate how differences in centrality and communication pathways of Abl kinases complexes with Imatinib and Dasatinib are linked with differences in drug resistant profiles of these inhibitors.

Materials and methods

MD Simulations

MD simulations of the protein kinase crystal structures (each of 500 ns duration) were performed for structures of the Abl complexes with Imatinib (pdb id 1IEP, 1OPJ) and Dasatinib (2GQG). The crystal structures of the Abl kinases were obtained from the Protein Data Bank[54]. The missing residues, unresolved structural segments and disordered loops were modelled with the ArchPRED server[55]. MD simulations were carried out using NAMD2.6[56] with the CHARMM27 force field[57,58] and the explicit TIP3P water model as implemented in NAMD2.6[59]. The employed MD protocol is consistent with the overall setup described in details in our earlier studies[60]. In a nutshell, the initial structures were solvated in a water box with the buffering distance of 10 Å. The system was heated from 100 K to 300 K in 30 ps and then cooled down again to 100 K in 30 ps in the NVT ensemble. In the following step, the system was heated in the NPT ensemble to 300 K over 30 ps keeping a restraint of 10 Kcal mol-1 Å-2 on protein alpha carbons (C_α). The system was then equilibrated for 300ps at 300K in the NVT ensemble without restraining forces on the atoms and then for further 300ps at 300K using the NPT ensemble to achieve uniform pressure. An NPT production simulation was run on the equilibrated structures for 500 ns keeping the temperature at 300 K and constant pressure (1 atm) using Langevin piston coupling algorithm. The van der Waals interactions were treated by using a switching function at 10Å and reaching zero at a distance of 12Å. Principal component analysis (PCA) of the MD conformational ensembles was based on the Cα atoms to determine the essential dynamics of the protein systems. The calculations were performed using the CARMA package[61]. For comparison, we also employed the elastic network model (ENM) and computed ENM-based lowest normal modes using the Anisotropic Network Model web server[62].

Protein Structure Network Modelling

In the protein structure network analysis, a graph-based representation of proteins was used in which amino acid residues were considered as nodes connected by edges corresponding to the nonbonding residue-residue interactions. The pair of residues with the interaction strength I_{ij} greater than a user-defined cut-off ($I_{\min}$) are connected by edges and produce a protein structure network graph for a given interaction strength $I_{\min}$. The strength of interaction between two amino acid side chains is evaluated as follows :

$$I_{ij} = \frac{n_{ij}}{\sqrt{(N_i \times N_j)}} \times 100 \quad (1)$$

where n_{ij} is number of distinct atom pairs between the side chains of amino acid residues i and j that lie within a distance of 4.5 Å. N_i and N_j are the normalization factors for residues i and j respectively[63,64]. The number of interaction pairs including main-chain and side-chain made by residue type i with all its surrounding residues in a protein k is also evaluated. The normalization factors take into account the differences in the sizes of the side chains of the different residue types and their propensity to make the maximum number of contacts with other amino acid residues in protein structures. The pair of residues with the interaction I_{ij} greater than a user-defined cut-off ($I_{\min}$) are connected by edges and produce a protein structure network graph for a given interaction cut-off $I_{\min}$. The pair of residues with the interaction strength I_{ij} greater than a user-defined cut-off ($I_{\min}$) are connected by edges and produce a protein structure network (PSN) graph for a given interaction strength $I_{\min}$. Similar to the arguments presented in our earlier studies[60], we considered any pair of residues to be connected if $I_{\min}$ was greater than 3.0%. We treat protein–ligand complexes as interaction networks in which the nodes of the network are formed by both amino acid residues and ligand atoms. In the network model, the binding of a ligand introduces new edges in the protein network and more closely links important protein nodes.

Network Parameters

Protein structure networks were constructed by incorporating the topology-based residue connectivity in combination with the contact maps of residues cross-correlations obtained from MD simulations. A weighted network representation of the protein structure is adopted that includes non-covalent connectivity of side chains and residue cross-correlation fluctuation matrix[50]. In this model of a protein network, the weight w_{ij} of an edge between nodes i and j is determined by the dynamic information flow through that edge as measured by the correlation between respective residues. The weight w_{ij} is defined as $w_{ij} = -\log(|C_{ij}|)$ where C_{ij} is the element of the covariance matrix measuring the cross-correlation between fluctuations of residues is i and j obtained from MD simulations. The shortest paths between two residues are determined using the Floyd–Warshall algorithm[65] that compares all possible paths through the graph between each pair of residue nodes. At the first step, the distance between connected residues was considered to be one, and the shortest path was identified as the path in which the two distant residues were connected by the smallest number of intermediate residues. To select the shortest paths that consist of dynamically correlated intermediate residues, we considered the short paths that included sufficiently correlated (C_{ij} = 0.5–1.0) intermediate residues. Using the constructed protein structure networks, we computed the residue-based betweenness parameter. The betweenness centrality $C_b(n)$ of a node n is typically computed as $C_b(n) = \sum_{s \neq n \neq t} (\delta_{st}(n) / \delta_{st})$ where s and t are nodes different from n in the network, δ_{st} is the shortest paths number from s to t, and $\delta_{st}(n)$ is the shortest paths number from s to t in which n lies on. For each node n the betweenness value is normalized by the number of node pairs excluding n given as $(N-1)(N-2)/2$, where N is the total number of nodes in the connected component that node n belongs to. The betweenness of residue i is defined to be the sum of the fraction of shortest paths between all pairs of residues that pass through residue i :

$$C_b(n_i) = \sum_{j<k}^{N} \frac{g_{jk}(i)}{g_{jk}} \quad (2)$$

where g_{jk} denotes the number of shortest geodesics paths connecting j and k, and $g_{jk}(i)$ is the number of shortest paths between residues j and k passing through the node n_i. Residues with high occurrence in the shortest paths connecting all residue pairs have a higher betweenness values. The normalized betweenness of residue i can be expressed as follows:

$$C_b(n_i) = \frac{1}{(N-1)(N-2)} \sum_{\substack{j<k \\ j \neq i \neq k}}^{N} \frac{g_{jk}(i)}{g_{jk}} \quad (3)$$

g_{jk} is the number of shortest paths between residues j and k; $g_{jk}(i)$ is the fraction of these shortest paths that pass through residue i .The clustering coefficient C_n of a node n is the ratio between the total number of links actually connecting its nearest neighbours and the total number of possible links between the nearest neighbours of node n . The clustering coefficient is a ratio N/M where N is the edges numbers between the neighbours of a node n ; M is the maximum edge numbers which could possibly exist between the neighbours of a node n . The clustering coefficient C value of a node is a number between [0,1]. In undirected networks, the clustering coefficient C_n of a node n is defined as:

$$C_n = 2e_n / (k_n(k_n - 1)) \quad (4)$$

k_n is the number of neighbours of a node n and e_n is the number of connected pairs between all neighbours of a node n . The clustering coefficient of a residue node measures the number of interactions among its interacting residues, normalized by the maximum number of possible interactions. The network clustering coefficient is the average of the clustering coefficients for all nodes in the network.

To study assortative mixing in the interaction networks and determine if the high-degree nodes tend to be connected with other high-degree nodes or with other low-degree nodes, we computed the average neighborhood connectivity distribution. The neighborhood of a given node n is the set of its neighbors. The connectivity of a node n , denoted by k_n , is the size of its neighborhood. The neighborhood connectivity of a node n is defined as the average connectivity of all neighbours of a node n . A normalized version of this parameter is the network density. The density is a value between 0 and 1. A network which contains no edges and solely isolated nodes has a density of 0. In contrast, the density of a clique is 1. The average degree of neighbors of a node with degree *k* is calculated as follows:

$$< k_{nn} >= \sum_{k'} k' P(k'/\mathrm{k}) \quad (5)$$

$P(k'/k)$ is the conditional probability that an edge of node degree k is connected to a node with degree k' . If this function is increasing, the network is assortative, since it shows that nodes of high degree on average connect to nodes of high degree[66,67]. Conversely, if the function is decreasing, the network is disassortative, since nodes of high degree tend to connect to nodes of lower degree. The neighborhood connectivity distribution gives the average of the neighborhood connectivities of all nodes n with k neighbours. We have computed this parameter to study the tendency for residue nodes in the interaction network to be connected to other nodes with similar degree. Another measure of assortative mixing is the assortativity coefficient which is the Pearson correlation coefficient of degree between pairs of linked nodes. The assortativity coefficient is expressed as follows:

$$r = \frac{1}{\sigma_q^2} \sum_{jk} jk(e_{jk} - q_j q_k) \quad (6)$$

where q_k is the distribution of the remaining degree, which corresponds to the number of edges leaving the node, other than the one that connects the pair. The distribution q_k is derived from the degree distribution p_k as $q_k = \frac{p_{k+1}}{\sum_{j \geq 1} p_j}$ and e_{jk} the joint probability distribution of the remaining degrees of the two nodes at either end of a randomly chosen link. It is a symmetric quantity in an undirected graph, and satisfies the following normalizing conditions: $\sum_{jk} e_{jk} = 1$ and $\sum_{j} e_{jk} = q_k$. σ_q is the variance of the distribution q_k . The value of the assortativity coefficient r gives a quantitative estimation of the mixing behavior of nodes in a network. Positive values of r indicate a correlation between nodes of similar degree, while negative r values indicate relationships between nodes of different degree. In general, r lies between −1 and 1. When r = 1, the network is said to have perfect assortative mixing patterns, when r = 0 the network is non-assortative, while at r = −1 the network is completely disassortative. All topological measures were computed using the python module NetworkX (http://networkx.github.io/). For comparison, we also computed topological network parameters (clustering coefficient, shared neighbor distribution, and average neighborhood connectivity) using the Cytoscape platform for network analysis[68] where the RING program[69] was used to generate the residue interaction networks, and the RINalyzer[70] and NetworkAnalyzer[71] plugins were used in the analysis.

Results and discussion

MD simulations of the Abl complexes and dynamics of the inhibitor-kinase interactions

Using MD simulations, we analysed similarities and differences in conformational dynamics of the Abl complexes, particularly focusing on ligand-mediated variations in structural stability of the critical "triad" (DFG motif, αC-helix, and A-loop). To analyze functionally relevant motions we employed PCA of MD simulations[72,73]. Functional dynamics profiles in the space of the three lowest frequency modes were mapped onto respective c-Abl structures (Fig. 1A,B) showing that structural rigidity of the catalytic core may be coupled with conformational mobility of the αC-helix and the P-loop. In particular, conformational and positional variability of the αC-helix is linked with structural stability of the adjacent αC-β4 loop in the N-terminal lobe. We previously conferred that the "boundary" between the rigid αC-β4 loop

and a more flexible αC-helix can define a functional hinge connecting regions of high and low structural stabilities in the N-terminal lobe of the catalytic domain[60]. Functional dynamics maps of residue fluctuations also indicated that the Imatinib-bound c-Abl complex may be more rigid than the Dasatinib-Abl complex. An important distinct characteristic of the Imatinib-Abl complex was a markedly reduced mobility of the kinked P-loop in the DFG-out inactive conformation, owing to the correlated fluctuations between the P-loop and the A-loop (Fig. 1A). The main regions of conformational flexibility in the c-Abl kinase domain included the αC-helix (residues 279-292), and the A-loop (residues 379-407). The intramolecular networks that form a regulatory spine (R-spine) and a catalytic spine (C-spine) networks connecting the N-lobe and the C-lobe have been widely accepted as critical for kinase regulation and activation[74-77]. The R-spine in Abl kinase consists of M290 from the C-terminal end of the αC-helix, L301 from the β4-strand, F382 of the DFG motif in the beginning of the A-loop, H361 of the His-Arg-Asp (HRD) motif in the catalytic loop, and D421 of the αF-helix. Structural stability of the R-spine residues was seen in both complexes, even though the R-spine was partially disassembled in the Imatinib-bound inactive conformation (Fig. 1A) and fully assembled in the Dasatinib-Abl active form (Fig 1B). The differences in the binding modes of Imatinib and Dasatinib could become more apparent from the analysis of the inhibitor-kinase contacts (Fig. 1C). the distributions illustrated strong Imatinib contacts with the 361-HRD-363 motif and 381-DFG-383 motifs that include the key R-spine residues (M290, H361, and F382). Importantly, Imatinib maintains a large number of stable contacts with Y253 in the P-loop, catalytic residue E286 and gate-keeper T315.

The open binding mode of Dasatinib resulted in the reduced number of intermolecular contacts as compared to Imatinib binding, but stronger interactions in the binding site, particularly hinge residues (E316, F317, and M318) (Fig. 1B,C). Despite the fewer number of total contacts, Dasatinib interactions with the adenine binding pocket, that is lined up with the hydrophobic residues (L248 in β1, V256 in β2, A269 in β3, V299 in the loop between αC-helix and β4, and L370 in β6), were stronger than for Imatinib. During simulations Dasatinib maintained hydrogen bonds with M318, T315 and E316 that are observed in the crystal structure[14]. Due to moderate flexibility of the P-loop, Dasatinib also sustained a number of stable contacts with the P-loop residues G249 and Y253. At the same time, Dasatinib binding is less sensitive to the interactions with V289 (αC-helix), L354 (αE-helix), F359, I360, H361 (catalytic loop), A380 (in β8) D381, F382, G383 (DFG motif). Overall, we found that Imatinib binding may be strongly coupled with a number of stable functional residues from the P-loop, A-loop, HRD and DFG motifs that are critical in stabilization of the inactive c-Abl structure and may contribute to the binding specificity. At the same time, Dasatinib binding to the active kinase form may primarily rely on structural stability of the hinge residues in the adenine binding pocket and allow for a greater mobility and variations in the intermolecular contacts with other functional regions.

Small-world topology of the residue interaction networks in the Abl kinase complexes

Although functional dynamics and collective motions provided certain evidence of allosteric coupling in c-Abl complexes, the residue interaction networks and communication pathways between various regions cannot be fully quantified using only analysis of conformational dynamics. The network approach can describe the global structural organization of proteins and complement PCA and ENM-based normal mode analyses of global collective motions. We employed a graph-based representation of protein structures by combining MD simulations and structure-based network analysis to characterize the residue interaction networks in the c-Abl kinase complexes. In this model, we integrated the topology-based residue connectivity and dynamics-based contact maps of residues cross-correlations obtained from MD simulations, thus expanding our analysis beyond a static description of protein structure networks. By weighing the residue interaction networks on the basis of cross-correlations between residues, a more accurate picture of the network topology may be obtained in which the average strength of allosteric signalling in c-Abl complexes can be related to experimental observations.

Residue betweenness, which is a global centrality measure, was used to compute residue interaction networks in the MD-derived ensembles of kinase conformations. The betweenness of a node is defined as the number of shortest paths that pass through that node in the network, representing a global medial measure of the node contribution to the communication within the network. These network parameters can characterize highly connected residues that mediate stable interaction networks and allosteric communications in protein structures. Using the MD trajectories we computed distributions of residue centrality in the c-Abl kinase complexes. In this analysis, it was observed that the residue betweenness could be anti-correlated with local conformational fluctuations (Fig. 2A,B). According to this finding, highly connected central residues in the kinase complexes are mostly rigid and such hub residues tend to have lower flexibility than sparsely connected residues. Consistent with previous studies[40], these results suggested that the average shortest path lengths in the kinase complexes may be associated with the extent of residue fluctuations, revealing an important link between protein network topology and conformational dynamics. Despite the observed correlation pattern, structurally rigid residues with low B-factors may exhibit a range of medium and high betweenness values, as could be seen in the Dasatinib-Abl complex (Fig. 2B). Therefore, low residue flexibility and cross-correlations of with other protein kinase residues could ensure a certain level of residue connectivity in the global network, but by itself may not be sufficient to infer functional significance. In other words, local measures of residue flexibility and global centrality parameters could represent complementary features in relating structural stability of a given residue with

its global mediating role in the interaction network and functional significance.

We found that Imatinib binding could induce a significant network-bridging effect in the Abl kinase and result in the betweenness profile where a small number of highly connected residues have high centrality, exemplified by the long distribution tail (Fig. 3A). This finding is reminiscent of a similar effect seen in computational studies of ligand binding with dihydrofolate reductase (DHFR) where the actual term "network-bridging" ligand effect was originally coined[47]. In the network terms, it would imply that Imatinib binding could induce the residue interaction network, where the short path length routes proceed through a small number of critical central residues and the bound ligand. The number of hub nodes seemed to exponentially decay as the betweenness of a residue node increased. This is indicative of a small-world network topology, in which a new node in the network is more likely to form an edge with a node that has higher than average number of connections, or with a node that has the short path length to such a node.

A scale-free topology is considered as an extreme case of small world networks, where path lengths between any two nodes are significantly shorter than can be predicted by the small-world effect. Due to inherent limitations imposed on the number of residue edges (interactions) and topological constraints, protein structure networks do not follow a scale-free behavior[78,79] but rather tend to display a Poisson-like centrality distribution dictated by a small-world network organization. The centrality profile of the Dasatinib-Abl complex featured signs of a random graph, where the majority of residues in the kinase domain have a relatively moderate level of centrality (Fig. 3B). This observation is in line with the evidence that side-chain interaction networks of flexible proteins may be described by random graph models[80]. A common feature of the random graph model is that the centrality distribution could peak at an average value, remaining homogeneous as most of the nodes have similar number of connections. Overall, the residue interaction networks in the kinase complexes tend to fall in between scale-free and random graph patterns, displaying a certain bias towards a Poisson-like centrality distribution.

We specifically focused on differences in the centrality profiles of the inhibitor-interacting residues in the Imatinib and Dasatinib complexes with c-Abl (Fig. 3C,D). To highlight the ligand-induced bridging effect, we compared centrality profiles of the binding site residues in the absence of bound inhibitors (Fig. 3C) and in the inhibitor-bound complexes (Fig. 3D). The betweenness values of the Imatinib-interacting residues were significantly greater for Y253 (P-loop), T315 (gate-keeper), H361, R362 (HRD motif), D381, F382 (DFG motif) and Y393 (A-loop). These findings revealed a stronger inhibitor bridging effect in the Imatinib-Abl complex and pointed to a global mediating role of these residues in the specific Abl conformation. From structural perspective, this may reflect the integrating role of Y253, D363, and Y393 residues in stabilizing allosteric interaction networks between the kinked P-loop, catalytic core and unpohosphorylated A-loop of Abl (Fig. 1A). It has been well recognized that phosphorylation status of Y393 is crucial for regulation of Abl activity and is necessary to stabilize the active kinase activation[11-13]. Our findings also indicated that structural stability of the inactive Abl conformation may be dependent on precise coupling between high centrality residues Y253, D363, and Y393. This may be relevant in explaining the experimentally observed high sensitivity of Imatinib binding to the kinked P-loop conformation and obligatory unphosphorylated form of the A-loop.

To directly evaluate the "small-worldness" of the residue interaction networks in the kinase complexes, we used MD simulations to compute distributions of several topological network parameters. The employed parameters included the short path length (Fig. 4A,B), the clustering coefficient distribution (Fig. 4C), and the neighbourhood connectivity (Fig. 4D). High overall clustering and a relatively short path length between any pair of nodes are key defining features of a small-world network[40-48]. In the Abl complexes the majority of short paths require ~3-5 nodes to connect any pair of residues in the interaction network. The results are consistent with previous studies suggesting that short path lengths in protein structures are typically distributed by small-world topology[47,48]. Despite an overall similarity, the important feature of the Imatinib-bound complex was a noticeable distribution shift towards the dominant peak corresponding to the shorter path length (~3 nodes). Accordingly, Imatinib binding may lead to a small-world network topology that can minimize transition times and optimize allosteric communications in the specific Abl complex. Strikingly, the network-bridging effect produced by Imatinib may amplify connectivity of some central residues and render a more "scale-free like" topology with efficient communications between rigid functional residues.

A network cluster (or module) defines a highly interconnected group of nodes that can be determined by the clustering coefficient, which is a signature of the network modularity. The clustering coefficient quantifies the number of connected pairs between a node and its neighbors and can be measured both local and globally as the average of the clustering coefficients for all nodes with k neighbours. The clustering coefficient distribution C (k) (Fig. 4C) is defined by the average of the clustering coefficient of the nodes with degree equal to k. In a random network, C(k) is independent of the node degree k since the vast majority of nodes have about the same number of connections, and their tendency to form clusters does not depend on the node degree[78,79]. Similar to random networks, C (k) is independent of k in scale-free networks[81,82]. Hierarchical networks are considered as a special type of scale-free networks with a large clustering coefficient that can be approximated by a power law[83]. For the kinase complexes, the clustering coefficient distribution follows as C (k) ~ $k^{-0.6}$ indicating an intermediate small-world regime between random and scale-free topology (Fig. 4C). In

such hierarchical organization, locally connected residue nodes typically belong to large clusters, while long-range communication between non-overlapping modules can be maintained by a relatively small number of central residues. Clustering coefficient corresponds to the density of the neighborhood of a residue in the network, and thus highly connected residue nodes have less densely clustered neighborhoods. High centrality (low clustering coefficient) residues are typically associated with the intermodular interactions and mutations of these residues may affect global connectivity of the network and the average path length of the network. In particular, we found that the high betweenness residues from the regulatory HRD and DFG motifs could be involved in mediating the intermodular connections.

A network measure that has recently caught substantial attention is the assortativity coefficient, which quantifies the preference of a node to attach to another one with similar (assortative mixing) or dissimilar (disassortative mixing) number of connections[84,85]. In assortative mixing the high-degree nodes tend to be connected with other high-degree nodes, while in disassortative networks the tendency of high degree nodes to connect with low degree ones results in a topology that favors speedy information processing across the network. These network characteristics can be assessed by computing the average neighborhood connectivity of a node which is defined as the average connectivity of all neighbors of this node. We computed the average neighbourhood connectivity distribution as the average of the neighborhood connectivities of all nodes with k neighbours (Fig. 4D). In the presence of assortative mixing, this parameter tends to increase with increasing k, while it decreases with k for a disassortative network. We observed that the average neighboring connectivity of a residue node generally increased with the number of neighbours k. After reaching the plateau level, this distribution begins to gradually falloff for highly connected residues (Fig. 4D). A plateau level approached in the distributions at larger values of k may arise from steric constraints of the structural fold. The character of the distribution profile remained largely the same for the kinase complexes, signalling the presence of assortative organization shared by the residue interaction networks. Interestingly, in the Imatinib-bound complex, the average connectivity for neighbors of the highest connected residues begins to gradually decline, pointing to the emergence of unique "mega-hubs" in the interaction network of the specific kinase complex.

It has been observed that assortative networks may exhibit a modular organization and display an efficient dynamics that is stable to noise and resilient to node deletion[78,79]. Unlike biological and protein-protein interaction networks that are known to be disassortative with the prevalent connectivity between high-degree and low-degree nodes[86], protein structure networks tend to shape a topology that is distinctively assortative. Our results supported this notion, demonstrating that hierarchical modularity (high clustering coefficient) and positive assortative mixing (increasing average neighborhood connectivity) may be general properties of kinase catalytic domains and their complexes with inhibitors. It has been also noted that increasing the assortativity may eventually decrease the network stability and that disassortative networks can be more resistant to the effect of dynamical fluctuations than assortative networks[87]. Hence, the observed positive assortativity and intermodular interactions in the inactive and active kinase structures corroborates with the marginal thermodynamic stability of structurally distinct kinase states that enables conformational transitions required for regulation.

Communication pathways in the Abl complexes: high centrality residues mediate allosteric interactions

Our results were particularly intriguing in light of recent experimental data demonstrating that selective Abl inhibitors, such as Imatinib, may be susceptible to the allosteric coupling between the kinked P-loop conformation and unphosphorylated state of the A-loop in Abl kinase[18]. These experiments posted an important question concerning a mechanism of allosteric communications between these regions that may be linked with binding preferences of specific inhibitors. We employed the network analysis to model how allosteric signals may be transmitted in the catalytic core. Modelling of allosteric communication pathways was based on the centrality analysis which generated the ensemble of short paths between any pair of residues in the Abl and c-Src complexes. In this analysis, we focused on the following objectives : (a) to map the ensemble of short path length between high centrality residues in the P-loop and phosphorylation sites Y393 and Y413 in the A-loop; (b) to determine the contribution of functional regions to the optimal communication pathways; (c) to determine differences in allosteric coupling of Imatinib and Dasatinib-bound complexes. We specifically analyzed the ensembles of communication pathways connecting the conserved high centrality Y253 residue (P-loop) with the primary phosphorylation site in the A-loop (Y393) and another phosphorylation site Y413.

The important finding of this analysis is the emergence of a dominant optimal pathway in the Imatinib-Abl complex that accounts for more than 30% of the total ensemble of short length pathways (Fig. 5A,B). This path linked Y253 from the P-loop with the nearby F382 and D381 residues (DFG motif), via D363 (HRD motif) that hydrogen bonded with Y393. This route could further link Y393 via H396 and N414 with another phosphorylation site Y413 in the A-loop (Fig. 5B). Consequently, allosteric communication pathways between the P-loop and A-loop in the Imatinib-Abl complex may be primarily mediated by high centrality residues from the HRD and DFG motifs and can be highly dependent on the hydrogen bonding between D363 and unphosphorylated Y393. Additionally, the optimal pathway seemed to exploit support of neighbouring hydrophobic residues of the A-loop (L384, L387, and H396) that may serve as "guardians" protecting the optimal route. Of special notice was a sharp decline in

the population of optimal pathways in the Imatinib-Abl specific complex and strong dependence of the single dominant path on a small number of highly connected residues (Fig. 5A,B). Strikingly, functional residues that constitute the dominant allosteric pathway are known to be critical for kinase regulation. For instance, mutation of a highly conserved R367 (R367A) showed a reduction in catalytic efficiency of about 5000-fold compared with the wild-type enzyme[88]. The strategic position of Y393 and its integrating central role in the optimal pathways is consistent with the high sensitivity of Imatinib binding to the structural arrangement and obligatory unphosphorylation status at this position[11-13,18]. Indeed, phosphorylation of Y393 can stabilize the open conformation of c-Abl kinase, thus blocking the access of Imatinib to the catalytic region and severely compromising its inhibitory function[89,90]. The importance of unphosphorylated Y393 was further demonstrated from mutagenesis analysis as Y393F variant formed a constitutively dephosphorylated state of Abl that resulted in increased kinase sensitivity towards Imatinib[89].

Structural and biochemical studies of Dasatinib binding[24,91] have convincingly shown that Abl can adopt a stable active conformation independent of the A-loop phosphorylation at Y393. Modelling of allosteric communications in the Dasatinib-Abl complex revealed a number of equally probable alternative routes connecting the P-loop and the A-loop regions (Fig. 5C, D). One of the optimal pathways connected Y253 via a catalytic pair (K271, E286) with the regulatory αC-helix (M290 from the R-spine) and then via V289, F359, L387 and M388 residues to the unphosphorylated Y393 (Fig. 5D). This pathway utilized the catalytic salt bridge K271-E286 coupled to the active position of the αC-helix to transmit allosteric signal via a set of rigid and primarily hydrophobic residues. Notably, most of these residues (K271, E286, V289, M290, F359, L387) are known to be functionally important for modulating conformational transitions and stabilization of the active kinase form[11-13]. We also found a number of equally probable pathways that explore alternative routes. For example, a potential route connected Y253 via a catalytic pair (K271, E286) with D381 (HRD motif), F382 (DFG motif), L384, S385 and R386 to the unphosphorylated Y393. This pathway navigated via high centrality residues and utilized hydrogen bonding between backbone carbonyl of S385 and guanidinium group of R386 as well as stacking interactions of Y393 with the hydrophobic part of R386. In the presence of phosphotyrosine, this route becomes somewhat more dominant as it is stabilized by direct hydrogen bonding between p-Y393 and the side-chain of R386.

To summarize, we determined that Dasatinib binding may activate a broad ensemble of allosteric pathways that could exploit structurally alternative routes between the P-loop and A-loop, exhibiting only marginal sensitivity to the phosphorylation status of Y393. The observed structural diversity of communication paths may be associated with the broader distribution of central residues. This may provide a better balance between the efficiency and resilience of the interaction networks in the Dasatinib complex.

The Network Effect of Ligand Binding and Drug Resistance: Correlating Mutational Effects with Residue Centrality

A large number of studies have investigated structural mechanisms of Imatinib-resistant mutations[92-94]. We hypothesized that network signatures of the Abl complexes may be relevant in quantifying drug resistance effects, particularly explaining the greater tolerance of Dasatinib to a number of mutations known to confer resistance to Imatinib. Although many mutations in the Abl kinase domain can be associated with Imatinib resistance, there are ~ 15 major point mutations that account for 85% of observed alterations in Abl kinase. We considered a panel of these major mutational forms associated with Imatinib resistance[94] that cover a range of residues in different kinase regions: L248, G250, Q252, Y253, E255 (P-loop), D276, E279 (C-helix), ATP binding site (V299, T315, F317), SH2 domain contact (M351), substrate binding region (F359), A-loop (L384, H3396, G398), and C-terminal lobe (F486). In the context of the current study, we elected to simplify analysis and perform a direct mapping of targeted residues onto the centrality profiles of the Imatinib and Dasatinib complexes (Fig. 6).

In this analysis, we correlated residue centralities against a consistent set of experimental data[92-95] on the following mutations: L248V, G250E, Q252H, Y253F, E255V, D276G, E279K, V299L, T315I/A, F317L/V/C, F359V, L384M, H396R, G398R, and F486S. We assembled all available IC50 values for Imatinib and Dasatinib binding with different mutant forms, including information on multiple mutations of the same binding site residues such as T315I/A and F317L/V/C. The utilized IC50 values against mutated forms of Abl were obtained under similar experimental conditions and corresponded to the average of at least three independent proliferation assay experiments[92-95]. We analyzed correlation patterns between residue centrality and the fold change in IC50 values of drug binding to the mutants. Given a significant chemical diversity of mutations and inclusion of multiple modifications of the same residues (T315 and F317), we tested whether the network signature of kinase residues could be a robust indicator of drug resistance effects.

The central finding of this analysis is that residue centrality in the kinase complexes may be strongly associated with the susceptibility to drug resistant mutations. We detected a fairly strong signal pointing to a relationship between residue centrality and mutation-induced fold increase in IC50 values for the Abl complexes (Fig. 6A, B). For the Imatinib-Abl binding, the correlation trend (R=0.6258) captured fairly well Imatinib resistance to Y253H, E255V, and T315I mutational variants (Fig. 6A) Dasatinib is active against many of the Abl mutations known to confer resistance to Imatinib except for the T315I mutation and some other mutations (V299L, F317L/V/C, and Q252H) that could be resistant or intermediately sensitive to Dasatinib[92-95]. These residues corresponded to the highest centrality sites and the correlation between the residue betweenness values and the IC50 fold increase upon mutations for Dasatinib-Abl binding approached R=0.766 (Fig. 6B). Importantly, this analysis provided an interesting interpretation of allosteric drug

resistance effects, whereby mutations of the kinase residues not directly interacting with the drug could still have significant impact on binding. In particular, some of the A-loop residues that contribute to the dominant communication route in the Imatinib-Abl complex include F382, L387, and H396 residues (Fig. 6C). While Imatinib maintains a significant number of contacts with F382, the drug does not form any interactions with L387 and H396 residues. At the same time, mutations of these residues were reported in patients including F382L, L387M and H396P/R, among which H396R was observed most frequently[96]. In the inactive Abl conformation L387 is orientated toward the P-loop and involved in communication pathways connecting P-loop and A-loop, so its mutation to methionine may affect the packing with the P-loop and affect allosteric signalling in the Imatinib-bound Abl. According to our observations, a fairly exposed A-loop residue H396 appeared to contribute to the optimal pathways linking the P-loop and A-loop. The crystal structure of the Abl kinase domain carrying the H396P mutation displays the active conformation of the activation loop to which Imatinib binds less readily[97]. In line with this observation, the H396P mutation was shown to strongly activate c-Abl activity[98] providing a possible explanation for the observed Imatinib resistance.

According to this analysis, mutations of these A-loop mutations, although not contacting Imatinib directly, may compromise the integrity and efficiency of short communication pathways mediating inhibition signal in the Imatinib-Abl complex, which may trigger shift in the conformational equilibrium towards the active conformation. Some of the key residues involved in multiple pathways connecting P-loop and A-loop in the Dasatinib-Abl complex include Y253 and F359 residues (Fig. 5D). Indeed, targeted mutations of these residues could cause some level of resistance to Dasatinib, which is tolerant to majority of the Imatinib-resistant mutations. However, patients with Y253H or F359V mutation have a high likelihood of developing new mutations in the setting of Dasatinib resistance[99].

The important conclusion from this analysis is that centrally positioned stable residues that preserve the short path length routes and ensure the efficiency of allosteric networks in the kinase structures are important for kinase regulation and can be also associated with drug resistance effects.

Conclusions

In this work, MD simulations of the Abl kinase complexes with cancer drugs Imatinib and Dasatinib were combined with structure-based network modelling to characterize dynamics of the residue interaction networks. The results have demonstrated that structural architecture of Abl kinase complexes can produce small-world topology conducive to a modular organization and assortative mixing of the residue interaction networks. An intriguing finding of this analysis revealed a significant network-bridging effect of Imatinib in Abl and pointed to a small-world topology of the residue interaction network, in which a small number of central residues could mediate binding preferences and coordinate allosteric communications. According to our findings, rapid communication in the Imatinib-Abl complex may come at the expense of high dependency on a small number of coordinating modes and sensitivity to the unphosphorylated state of the A-loop, making specific drug binding vulnerable to targeted mutations. A different organization of optimal pathways in the Dasatinib complex may provide a better balance between the efficiency and resilience of the interaction networks in the Dasatinib complex. Our results supported this notion, demonstrating that hierarchical modularity (high clustering coefficient) and positive assortative mixing (increasing average neighborhood connectivity) may be general properties of protein domains, including kinase catalytic domains and their complexes with inhibitors. The emergence of these topological traits suggests a similar organizing topology that enhances resilience of the interaction networks. Interestingly, the short path length distribution emerged as a topological indicator of differential kinase sensitivity to Imatinib binding, suggesting that variations in allosteric communication pathways may be linked with protein responses to ligand binding.

Another important implication of this study is a potential utility of network parameters in predicting drug resistance effects. Our results have related differences in the residue interaction networks and allosteric communications of the Abl kinase complexes with their binding preferences and drug resistance profiles. In particular, we have found that the severity of Imatinib-resistant effects may be differentiated based on network centrality properties. This finding may be particularly interesting since a number of Imatinib-resistant mutations may arguably exist before treatment and may contribute to tumorigenesis.

The structure-based network approach could provide a useful conceptual perspective on structural stability and function, also addressing the efficiency and robustness of proteins against failure of function due to mutations. The network approach is also attractive from a protein modelling perspective, enabling both local and global interaction effects to be considered from a unified view. The simplicity and transparency of the network analysis coupled with the rigor of biophysical simulations and power of structural modelling may prove to be a useful approach complementing existing computational and structural biology studies of protein kinase function and regulation.

Acknowledgements

This work is partly supported by funding from Chapman University.

Notes and references

1 M. Huse and J. Kuriyan, *Cell*, 2002, **109**, 275.
2 M.E. Noble, J.A. Endicott and L.N. Johnson, *Science*, 2004, **303**, 1800.
3 S.S.Taylor and A.P. Kornev, *Trends Biochem Sci*, 2011, **36**, 65.

4 J.A. Endicott, M.E. Noble and L.N. Johnson, *Annu. Rev. Biochem.* , 2012, **81**, 587.
5 S.S. Taylor, M.M. Keshwani, J.M. Steichen and A.P. Kornev, *Philos. Trans. R Soc. Lond. B Biol. Sci.*, 2012, **367**, 2517.
6 S.S. Taylor, R. Ilouz, P. Zhang and A.P. Kornev, *Nat. Rev. Mol. Cell. Biol.*, 2012, **13**, 646.
7 S.C. Artim, J.M. Mendrola and M.A. Lemmon, *Biochem. J.,* 2012, **448**, 213.
8 K. Oruganty and N. Kannan, *Philos. Trans. R Soc. Lond. B Biol. Sci.*, 2012, **367,** 2529.
9 K. Oruganty and N. Kannan, *Biochim. Biophys. Acta*, 2013, **1834**, 1322.
10 H.S. Meharena, P. Chang, M.M. Keshwani, K. Oruganty, A.K. Nene, N. Kannan N, S.S. Taylor and A.P. Kornev, *PLoS Biol.*, 2013, **11**, e1001680.
11 O. Hantschel, *Genes Cancer*, 2012, **3**, 436.
12 S.Panjarian, R.E. Iacob, S. Chen, J.R. Engen and T.E. Smithgall, *J. Biol. Chem.*, 2013, **288**, 5443.
13 E.P. Reddy and A.K. Aggarwal, *Genes Cancer*, 2012, **3**, 447.
14 J.S. Tokarski, J.A. Newitt, C.Y. Chang, J.D. Cheng, M. Wittekind, S.E. Kiefe, K. Kish, F.Y. Lee, R. Borzillerri, L.J. Lombardo, D. Xie, Y. Zhang and H.E. Klei, *Cancer Res.*, 2006, **66**, 5790.
15 T. Schindler, W. Bornmann, P. Pellicena, W.T. Miller, B. Clarkson and J. Kuriyan, *Science*, 2000, **289**, 1938.
16 B. Nagar, W.G. Bornmann, P. Pellicena, T. Schindler, D.R. Veach, W.T. Miller, B. Clarkson and J. Kuriyan, *Cancer Res.*, 2002, **62**, 4236.
17 N.M. Levinson, O. Kuchment, K. Shen, M.A. Young, M. Koldobskiy, M. Karplus, P.A. Cole and J. Kuriyan, *PLoS Biol.* 2006, **4**, e144.
18 S.B. Hari, B.G. Perera, P. Ranjitkar, M.A. Seeliger and D.J. Maly, *ACS Chem. Biol*., 2013, **8**, 2734.
19 T. A. Carter, L. M. Wodicka, N. P. Shah, A. M. Velasco, M. A. Fabian, D. K. Treiber, Z. V. Milanov, C. E. Atteridge, W. H. Biggs, P. T. Edeen, M. Floyd, J. M. Ford, R. M. Grotzfeld, S. Herrgard, D. E. Insko, S. A. Mehta, H. K. Patel, W. Pao, C. L. Sawyers, H. Varmus, P. P. Zarrinkar and D. J. Lockhart, *Proc. Natl. Acad. Sci. U. S. A.,* 2005, **102**, 11011.
20 M. A. Fabian, W.H. Biggs, D.K. Treiber, C.E. Atteridge, M. D. Azimioara, M. G. Benedetti, T. A. Carter, P. Ciceri, P. T. Edeen, M. Floyd, J. M. Ford, M. Galvin, J.L. Gerlach, R. M. Grotzfeld, S. Herrgard, D. E. Insko, M. A. Insko, A. G. Lai, J. M. Lélias, S. A. Mehta, Z. V. Milanov, A. M. Velasco, L. M. Wodicka, H. K. Patel, P. P. Zarrinkar and D. J. Lockhart, *Nat. Biotechnol.*, 2005, **23**, 329.
21 M. Bantscheff, D. Eberhard, Y. Abraham, S. Bastuck, M. Boesche, S. Hobson, T. Mathieson, J. Perrin, M. Raida, C. Rau, V. Reader, G. Sweetman, A. Bauer, T. Bouwmeester, C. Hopf, U. Kruse, G. Neubauer, N. Ramsden, J. Rick, B. Kuster and G. Drewes, *Nat. Biotechnol*., 2007, **25**, 1035.
22 U. Rix, O. Hantschel, G. Durnberger, L.L. Remsing Rix, M. Planyavsky, N. V. Fernbach, I. Kaupe, K. L. Bennett, P. Valent, J. Colinge, T. Kocher and G. Superti-Furga, *Blood*, 2007, **110**, 4055.
23 O. Hantschel, U. Rix, U. Schmidt, T. Bürckstümmer, M. Kneidinger, G. Schütze, J. Colinge, K.L. Bennett, W. Ellmeier, P. Valent and G. Superti-Furga, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 13283.
24 L.M. Wodicka, P. Ciceri, M.I. Davis, J.P. Hunt, M. Floyd, S. Salerno, X.H. Hua, J.M. Ford, R.C. Armstrong, P.P. Zarrinkar and D.K. Treiber, *Chem. Biol*., 2010, **17**, 1241.
25 G.M. Verkhivker, *Biopolymers*, 2007, **85**, 333.
26 T.S. Lee, S.J. Potts, H. Kantarjian, J. Cortes, F. Giles and M. Albitar, *Cancer*, 2008, **112**, 1744.
27 Y. Shan, M.A. Seeliger, M.P. Eastwood, F. Frank, H. Xu, M.Ø. Jensen, R. O. Dror, J. Kuriyan and D. E. Shaw, *Proc. Natl. Acad. Sci. U. S.A.,* 2009, **106**, 139.
28 A. Dixit and G.M. Verkhivker, *PLoS Comput. Biol.*, 2009, **5**, e1000487.
29 A. Dixit and G.M. Verkhivker, *PLoS Comput Biol.,* 2011, **7**, e1002179.
30 A. Aleksandrov and T. Simonson, *J. Biol. Chem*., 2010, **285**, 13807.
31 Y.L. Lin, Y. Meng, W. Jiang and B. Roux, *Proc. Natl. Acad. Sci. U. S. A.,* 2013, **110**, 1664.
32 S. Lovera, L. Sutto, R. Boubeva, L. Scapozza, N. Dölker and F.L. Gervasio, *J. Am. Chem. Soc*., 2012, **134**, 2496.
33 Y.L. Lin and B. Roux B., *J. Am. Chem. Soc*., 2013, **135**, 14741.
34 Y.L. Lin, Y. Meng, L. Huang and B. Roux, *J. Am. Chem. Soc*., 2014, **136**, 14753.
35 E. Laurini, P. Posocco, M. Fermeglia, D.L. Gibbons, A. Quintás-Cardama and S. Pricl, *Mol. Oncol*., 2013, **7**, 968.
36 D.J. Watts and S.H. Strogatz, *Nature*, 1998, 393, 440.
37 M. Vendruscolo, N.V. Dokholyan, E. Paci and M. Karplus, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 2002, **65**, 061910.
38 N.V. Dokholyan, L. Li, F. Ding and E.I. Shakhnovich, *Proc. Natl. Acad. Sci. U.S.A*., 2002, **99**, 8637.
39 L.H. Greene and V.A. Higman, *J. Mol. Biol.,* 2003*,* **334**, 781.
40 A.R. Atilgan, P. Akan and C. Baysal, *Biophys J.*, 2004, **8**6, 85.
41 A. del Sol and P. O'Meara, *Proteins*, 2005, **58**, 672.
42 A. del Sol, H. Fujihashi and P. O'Meara, *Bioinformatics*, 2005, 21, 1311.
43 R. Sathyapriya, M.S. Vijayabaskar and S. Vishveshwara, *PLoS Comput. Biol.,* 2008*,* **4**, e1000170.
44 G. Amitai, A. Shemesh, E. Sitbon, M. Shklar, D. Netanely, I. Venger and S. Pietrokovski, *J. Mol. Biol.,* 2004*,* **344**, 1135.
45 Z. Hu, D. Bowen, W.M. Southerland, A. del Sol, Y. Pan, R. Nussinov and B.Ma, *PLoS Comput. Biol.,* 2007*,* **3**, e117.
46 N.R. Taylor, *Comput. Struct. Biotechnol. J.*, 2013, **5**, e201302006.
47 A. del Sol, H. Fujihashi, D. Amoros and R. Nussinov, *Protein Sci.,* 2006, **15**, 2120.
48 A. del Sol, H. Fujihashi, D. Amoros and R. Nussinov, *Mol. Sys. Biol*., 2006, **2**, 2006.0019.
49 M. Bhattacharyya and S. Vishveshwara, *Biochemistry*, 2011, **50**, 6225.
50 A. Sethi, J. Eargle, A.A. Black and Z. Luthey-Schulten, *Proc. Natl. Acad. Sci. U.S.A.,* 2009*,* **106**, 6620.
51 A. Ghosh, R. Sakaguchi, C. Liu, S. Vishveshwara and Y.M. Hou, *J. Biol. Chem*., 2011, **286**, 37721.
52 G. Bagler and S. Sinha, *Bioinformatics*, 2007, **23**, 1760.
53 D. Hao and C. Li, *PLoS One*, 2011, **6**, e28322.
54 H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov and P.E. Bourne, *Nucleic Acids Res*., 2000, **28**, 235.
55 N. Fernandez-Fuentes, J. Zhai and A. Fiser, *Nucleic Acids Res*., 2006, **34**, W173.
56 J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, C. Chipot, R.D. Skeel, L. Kalé and K. Schulten, *J. Comput. Chem*., 2005, **26**, 1781.
57 A.D. MacKerell, D. Bashford, M. Bellott, R.L. Dunbrack, J.D. Evanseck, M.J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. Lau, C. Mattos, S. Michnick, T. Ngo, D. Nguyen, B. Prodhom, W. Reiher, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B.,* 1998, **102**, 3586.
58 A.D. MacKerell, N. Banavali, and N. Foloppe, *Biopolymers,* 2001*,* **56***, 257.*
59 W.L. Jorgensen, J. Chandrasekhar, J.D. Madura, R.W. Impey and M.L. Klein, *J. Chem. Phys*., 1983, **79**, 926.
60 K.A. James and G.M. Verkhivker, *PLoS One*, 2014, **9**, e113488.

61 P.I. Koukos and N.M. Glykos, *J. Comput. Chem*., 2013, **34**, 2310.
62 E. Eyal, L.W. Yang and I. Bahar, *Bioinformatics*, 2006, **22**, 2619.
63 N. Kannan and S. Vishveshwara, *J. Mol. Biol*., 1999, **292**, 441.
64 K.V. Brinda and S. Vishveshwara, *Biophys. J.,* 2005*,* **89**, 4159.
65 R.W. Floyd, *Commun. A.C.M*., 1962, **5**, 345.
66 M.E. Newman, *Phys. Rev. Lett.*, 2002, **89**, 208701.
67 M.E. Newman, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.,* 2003, **67**, 026126.
68 P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski and T. Ideker, *Genome Res.,* 2003, **13**, 2498.
69 A.J. Martin, M. Vidotto, F. Boscariol, T. Di Domenico, I. Walsh and S.C. Tosatto, *Bioinformatic*s, 2011, **27**, 2003.
70 N.T. Doncheva, Y. Assenov, F.S. Domingues and M. Albrecht, *Nat. Protoc*., 2012, **7**, 670.
71 Y. Assenov, F. Ramírez, S.E. Schelhorn, T. Lengauer and M. Albrecht, *Bioinformatics*, 2008, **24**, 282.
72 A. Amadei, A.B. Linssen and H.J. Berendsen, *Proteins*, 1993, **17**, 412.
73 A. Amadei, M.A. Ceruso and A. Di Nola, *Proteins*, 1999, **36**, 419.
74 A.P. Kornev, N.M. Haste, S.S. Taylor and L.F. Ten Eyck, *Proc. Natl. Acad. Sci. U. S. A.,* 2006, **103**, 17783.
75 A.P. Kornev, S.S. Taylor and L.F. Ten Eyck, *Proc. Natl. Acad. Sci. U. S. A*., 2008, **105**, 14377.
76 L.F. Ten Eyck, S.S. Taylor and A.P. Kornev, *Biochim. Biophys. Acta*, 2008, **1784**, 238.
77 M. Azam, M. A. Seeliger, N. S. Gray, J. Kuriyan and G.Q. Daley, *Nat. Struct. Mol. Biol.,* 2008, **15**, 1109.
78 A.L. Barabási and Z.N. Oltvai, *Nat. Rev. Genet*., 2004, **5**, 101.
79 C. Böde, I.A. Kovács, M.S. Szalay, R. Palotai, T. Korcsmáros and P. Csermely, *FEBS Lett*., 2007, **58**, 2776.
80 L. Bartoli, P. Fariselli and R. Casadio, *Phys. Biol*., 2008, **4**, L1.
81 A.L. Barabási, *Science*, 2009, **325**, 412.
82 A.L. Barabási, *Philos. Trans. A Math. Phys. Eng. Sci*., 2013, **371**, 20120375.
83 E. Ravasz, A.L. Somera, D.A. Mongru, Z.N. Oltvai and A.L. Barabási, *Science*, 2002, **297**, 1551.
84 J. Park and M.E. Newman, *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.*, 2004, **70**, 066117.
85 G. Bagler and S. Sinha, *Bioinformatics*, 2007, **23**, 1760.
86 S. Maslov and K. Sneppen, *Science*, 2002, **296**, 910.
87 N.A. Alves and A.S. Martinez, *Physica A*., 2007, **375**,336.
88 K.E. Muratore, M.A. Seeliger, Z. Wang, D. Fomina, J. Neiswinger, J. J. Havranek, D. Baker, J. Kuriyan and P. A. Cole, *Biochemistry*, 2009, **48**, 3378.
89 B.B. Brasher and R. A. Van Etten, *J. Biol. Chem*., 2000, **275**, 35631.
90 K. Dorey, J.R. Engen, J. Kretzschmar, M. Wilm, G. Neubauer, T. Schindler and G. Superti-Furga, *Oncogene*, 2001, **20**, 8075.
91 A.J. Lamontanara, S. Georgeon, G. Tria, D.I. Svergun and O. Hantschel, *Nat. Commun.,* 2014, **5**, 5470.
92 S. Soverini, A. Hochhaus, F. E. Nicolini, F. Gruber, T. Lange, G. Saglio, F. Pane, M.C. Muller, T. Ernst, G. Rosti, K. Porkka, M. Baccarani, N. Cross and G. Martinelli, *Blood*, 2011, **118**,1208.
93 S. Soverini, C. De Benedittis, C. Papayannidis, S. Paolini, C. Venturi, I. Iacobucci, M. Luppi, P. Bresciani, M. Salvucci, D. Russo, S. Sica, E. Orlandi, T. Intermesoli, A. Gozzini, M. Bonifacio, G.M. Rigolin, F. Pane, M. Baccarani, M. Cavo and G. Martinelli, *Cancer*, 2014, **120**, 1002.
94 S. Redaelli, R. Piazza, R. Rostagno, V. Magistroni, P. Perini, M. Marega, C. Gambacorti-Passerini and F. Boschelli, *J. Clin. Oncol.,* 2009, **27**, 469.
95 H.A. Bradeen, C.A. Eide, T. O'Hare, K.J. Johnson, S.G. Willis, F.Y. Lee, B.J. Druker and M.W. Deininger, *Blood*, 2006 **108**, 2332.
96 A.J. Lamontanara, E.B. Gencer, O. Kuzyk and O. Hantschel, *Biochim. Biophys. Acta*, 2013, **1834**, 1449.
97 M.A. Young, N.P. Shah, L.H. Chao, M. Seeliger, Z.V. Milanov, W.H. Biggs, D.K. Treiber, H.K. Patel, P.P. Zarrinkar, D.J. Lockhart, C.L. Sawyers and J. Kuriyan, *Cancer Res*., 2006, **66**, 1007.
98 D. W. Sherbenou, O. Hantschel, I. Kaupe, S. Willis, T. Bumm, L. P. Turaga, T. Lange, K. H. Dao, R.D. Press, B. J. Druker, G. Superti-Furga and M. W. Deininger, *Blood*, 2010, **116**, 3278.
99 Q. Jiang, Y. Qin, Y. Lai, H. Jiang, J. Wang and X. Huang, *Leuk. Lymphoma*, 2015, **14**, 1.

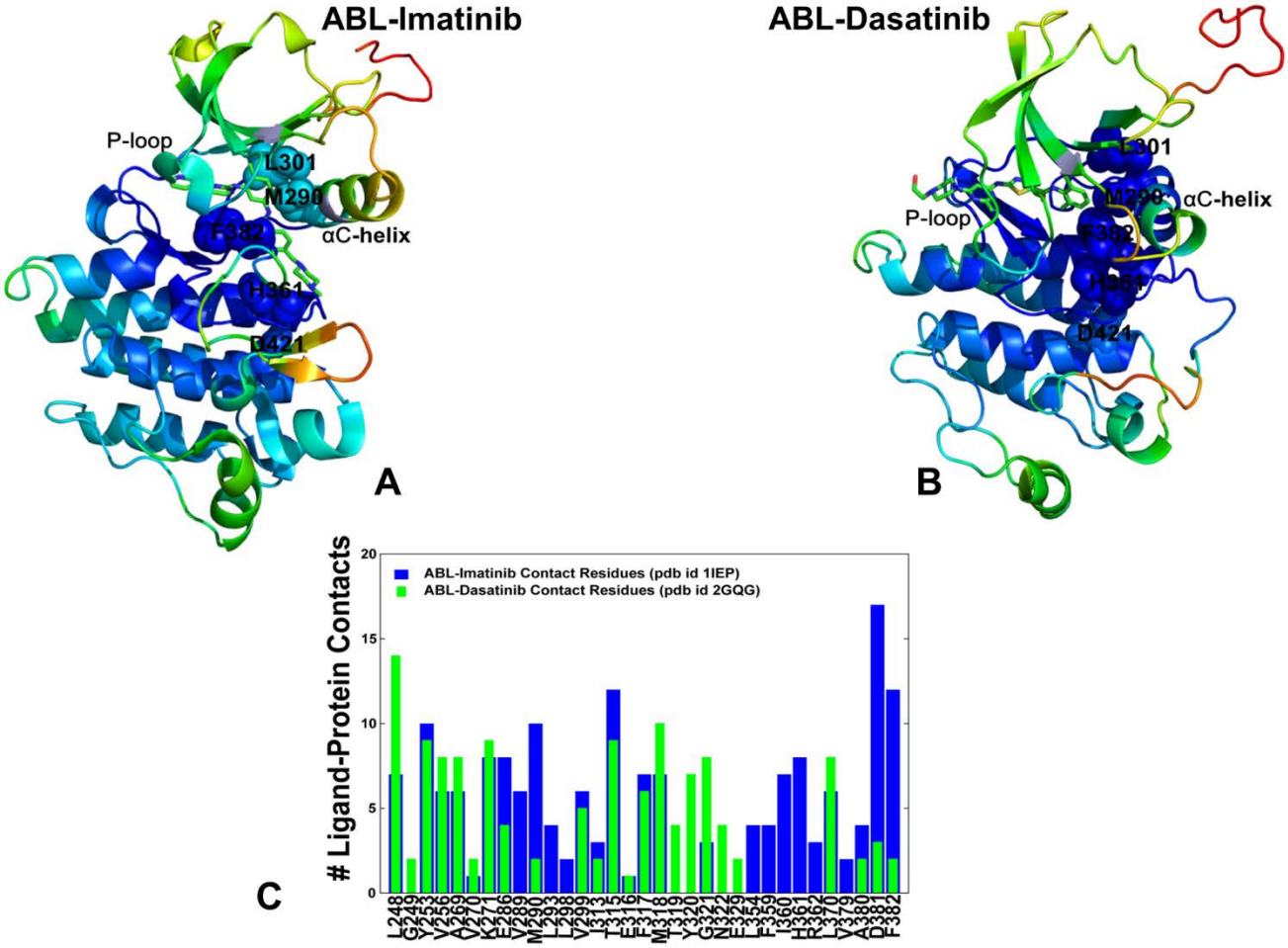


Fig. 1. Functional dynamics of the Abl-inhibitor complexes. Conformational dynamics profiles of the Abl structures in the essential space of three low frequency modes are shown for the inactive Imatinib-bound form (A) and the Dasatinib-bound active kinase form (B). The colour gradient from blue to red indicates the decreasing structural rigidity (or increasing conformational mobility) of the protein residues and refers to an average value over the backbone atoms in each residue. The R-spine residues are annotated in spheres and colored according to their degree of structural stability. (C) The ensemble-based distribution of the intermolecular contacts in the Imatinib-Abl and Dasatinib-Abl complexes. The number of the intermolecular contacts formed by the inhibitors with the interacting residues is shown for Imatinib in blue filled bars and for Dasatinib in green filled bars.

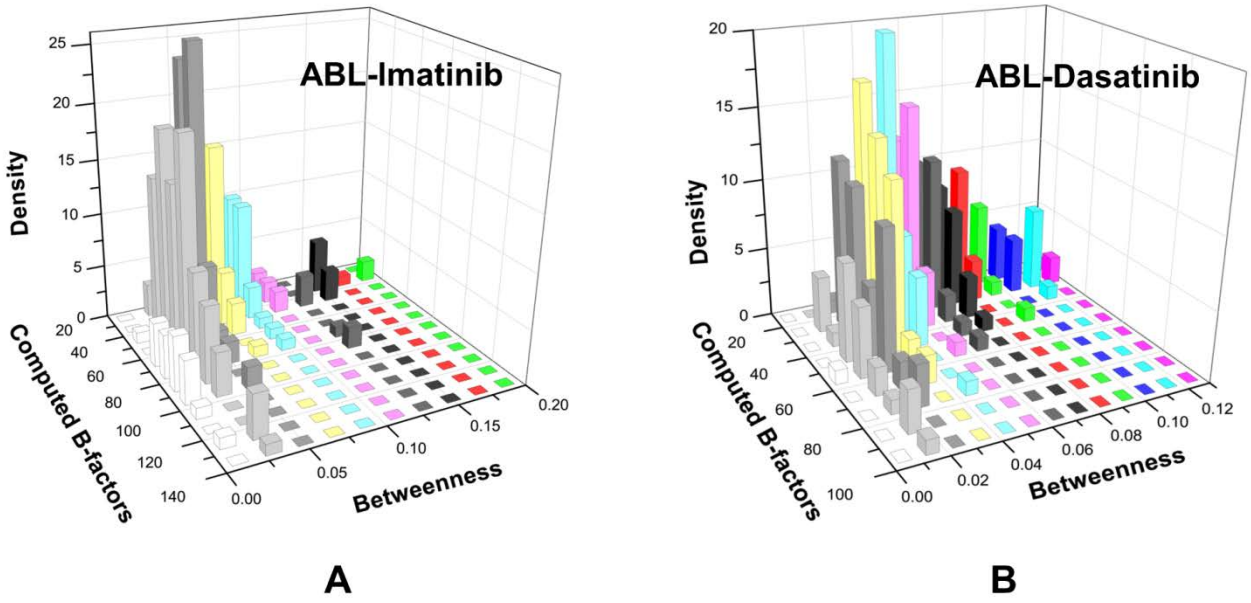


Fig. 2. Structure-based network analysis of the Abl-inhibitor complexes. The joint probability distribution of computed B-factors and residue betweenness values in the Imatinib-Abl (A) and Dasatinib-Abl complexes (B). The depicted profile combines local estimate of conformational fluctuations with the global measure of residue centrality in the interaction networks.

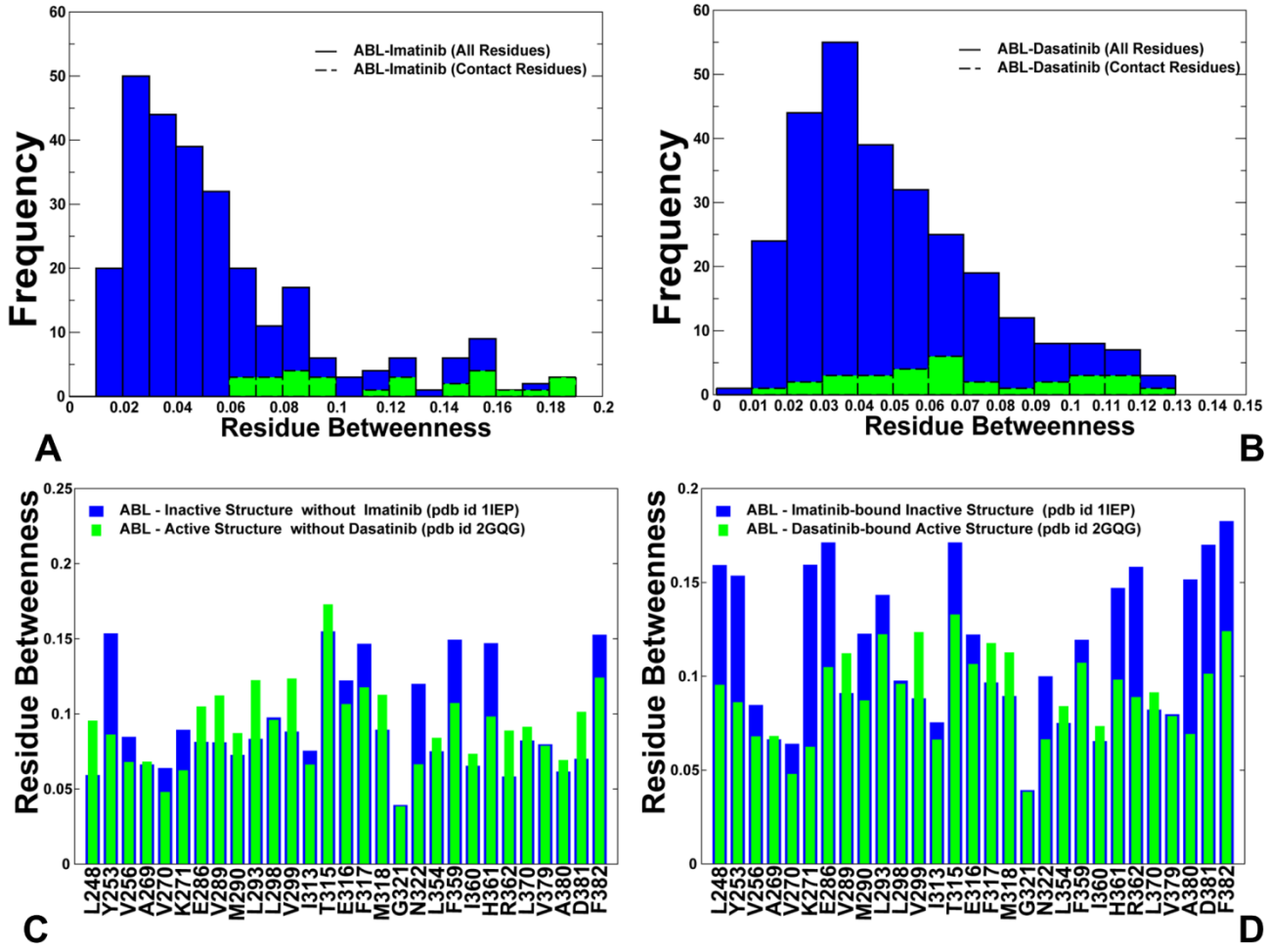


Fig. 3. Structure-based network analysis of the Abl-inhibitor complexes. The residue-based betweenness distribution in the Imatinib-Abl (A) and Dasatinib-Abl complexes (B). The distributions are shown in blue bars and the population of the inhibitor-interacting residues is depicted in green. (C, D) A comparative analysis of centrality for the binding site residues in the inactive (blue bars) and active Abl structures (green bars) in the absence of bound inhibitors (C) and in the Imatinib-bound (blue bars) and Dasatinib-bound (blue bars) Abl complexes (D).

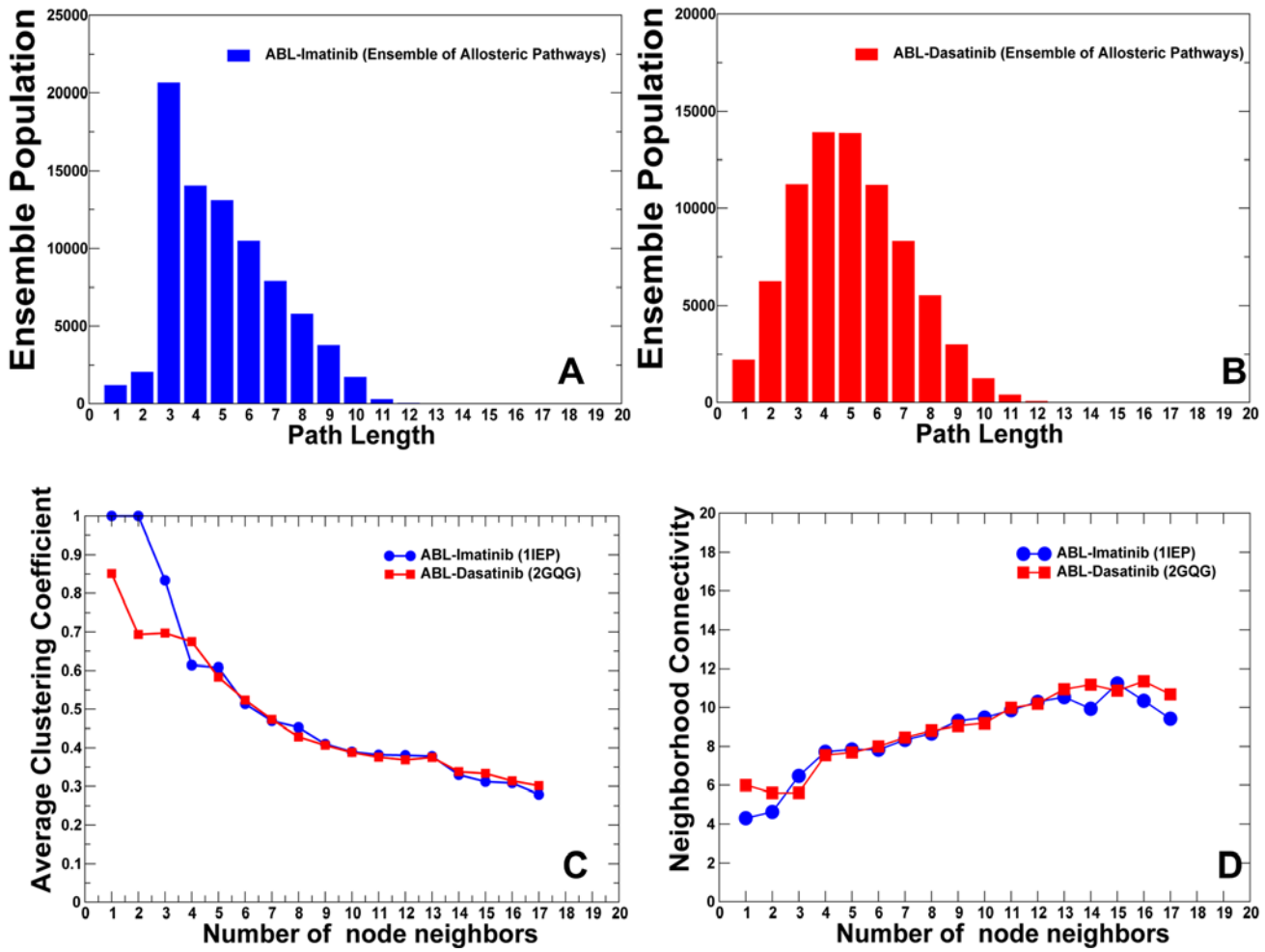


Fig. 4. The distribution of topological network parameters in the Abl-inhibitor complexes. (A,B) The distribution of the short path length between residue pairs in the Imatinib-Abl complex (in blue bars) and Dasatinib-Abl complex (in green bars). (C) The distribution of the average clustering coefficient and (D) the distribution of the average neighbourhood connectivity. The average clustering coefficient and the average neighbourhood connectivity distributions are shown in colored lines, with Imatinib-Abl (in blue circles), Dasatinib-Abl (in red squares).

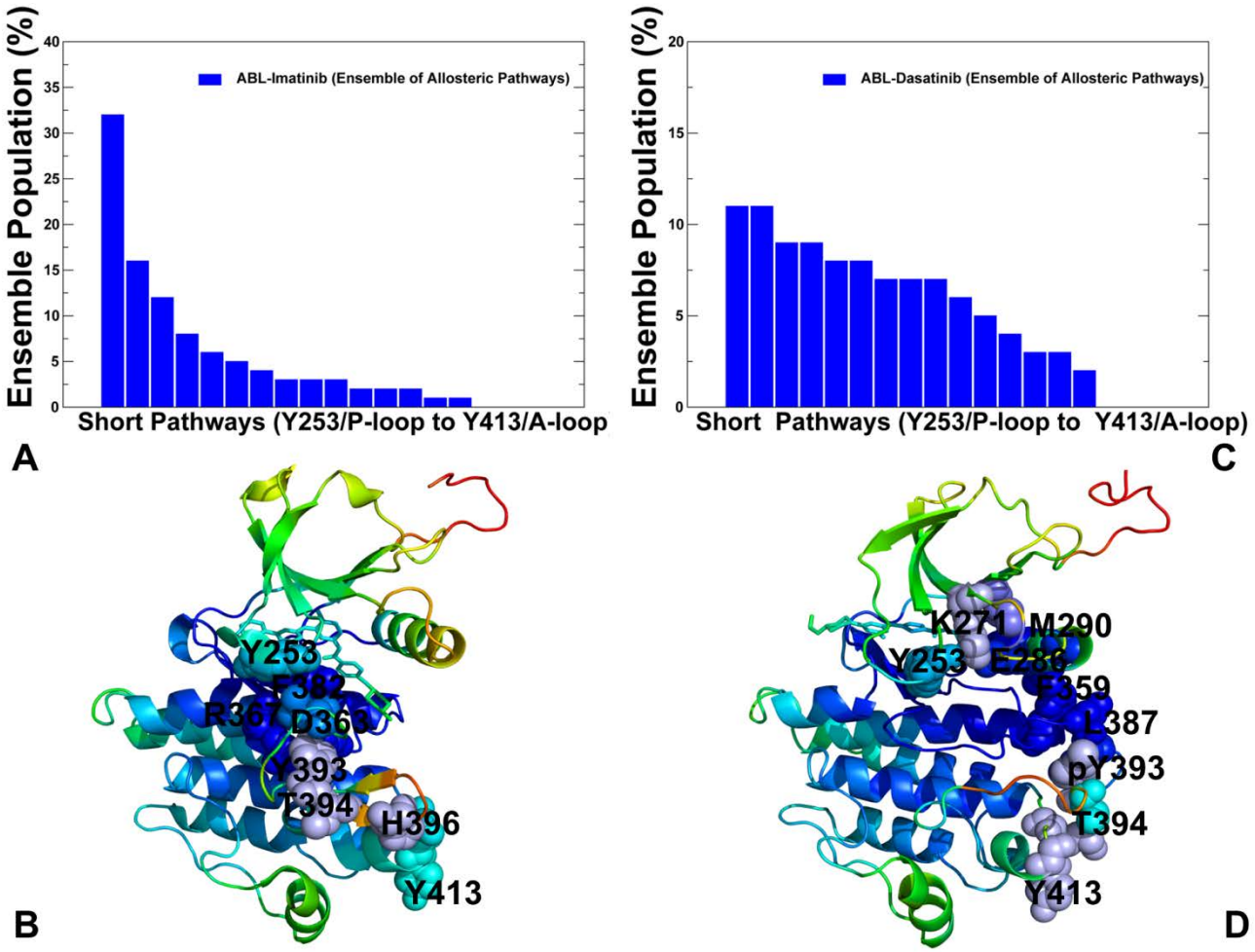


Fig. 5. Analysis of conformational allosteric pathways in the Abl-inhibitor complexes. (A,C) The population distribution of optimal communication pathways connecting the Y253 (P-loop) and phosphorylation site Y413 (A-loop) in the Imatinib-Abl and Dasatinib-Abl complexes. (B) The dominant communication pathway connecting Y253 of the P-loop with the phosphorylation sites in the A-loop (Y393 and Y413) in the Imatinib-Abl complex. (D) A semi-optimal communication pathway connecting the same residues in the Dasatinib-Abl complex. The residue numbering corresponds to the original annotation from the crystal structures (pdb id 1IEP and 2OIQ). The contributing residues are shown in filled spheres and colored according to their conformational mobility in the principal space of low frequency modes as in Fig. 1.

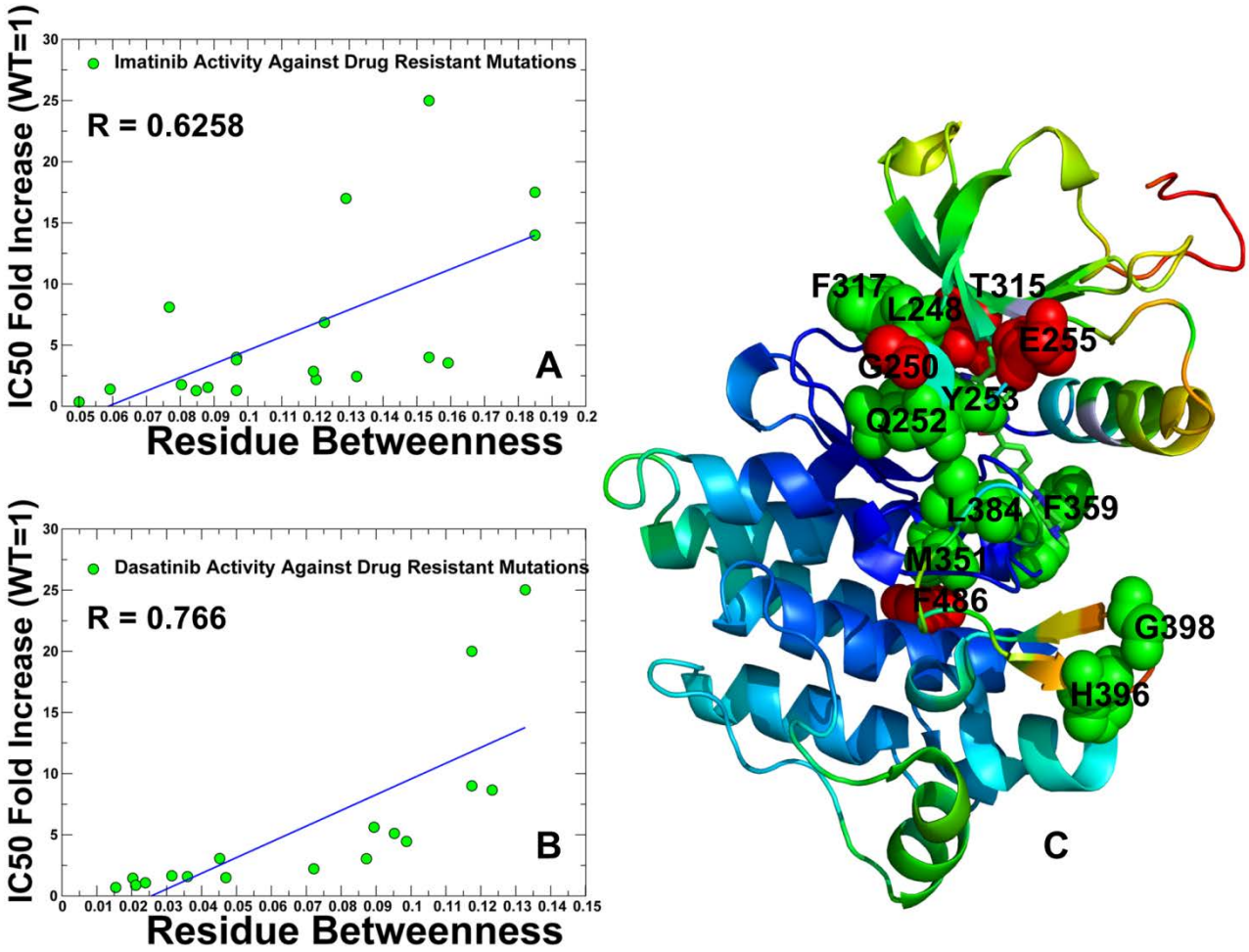
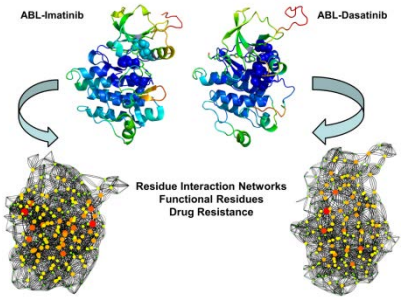


Fig. 6. The network analysis of drug resistance in the Abl-inhibitor complexes. (A,B) A correlation between residue betweenness and experimentally measured fold increases in IC50 values upon resistant mutations in the Imatinib-Abl and Dasatinib-Abl complexes respectively. (C) Structural mapping of Imatinib-resistant mutations onto Imatinib-bound inactive Abl structure. The sites of highly resistant Imatinib mutations (G250E, Y253F, E255V, T315I, F486S) are shown in red spheres, while residues targeted by moderately resistant Imatinib mutations (L248V, Q252H, V299L, F317L, F359V, H396, L384M, G398R) are shown as green spheres.

Table of contents entry



Computational modelling of efficiency and robustness of the residue interaction networks and allosteric pathways in kinase structures can characterize protein kinase sensitivity to drug binding and drug resistance effects