

PocketQuery: protein–protein interaction inhibitor starting points from protein–protein interaction structure

David Ryan Koes* and Carlos J. Camacho

Department of Computational and Systems Biology, University of Pittsburgh, 3501 Fifth Avenue, Pittsburgh, PA 15260, USA

Received January 16, 2012; Revised March 20, 2012; Accepted April 4, 2012

ABSTRACT

PocketQuery (<http://pocketquery.csb.pitt.edu>) is a web interface for exploring the properties of protein–protein interaction (PPI) interfaces with a focus on the discovery of promising starting points for small-molecule design. PocketQuery rapidly focuses attention on the key interacting residues of an interaction using a ‘druggability’ score that provides an estimate of how likely the chemical mimicry of a cluster of interface residues would result in a small-molecule inhibitor of an interaction. These residue clusters are chemical starting points that can be seamlessly exported to a pharmacophore-based drug discovery workflow. PocketQuery is updated on a weekly basis to contain all applicable PPI structures deposited in the Protein Data Bank and allows users to upload their own custom structures for analysis.

INTRODUCTION

Protein–protein interactions (PPIs) are essential for biological function and are an emerging class of therapeutic targets (1,2). PPIs have proven to be a challenging target for drug discovery (3) and have performed poorly in high-throughput screens (4), possibly due to the historical bias of existing chemical libraries (5). One successful approach for targeting PPIs is the rational design of small-molecules that mimic the interaction of a few key residues at the protein–protein interface (6–8). These residues are typically deeply buried ‘anchor’ residues (9) and/or ‘hot-spot’ residues (10,11). The prediction of such residues has received a great deal of attention, and approaches use structural features (9,12–14), sequence conservation (15–18), or, most successfully, a multi-feature consensus approach (19–24). Although most predictors focus on identifying individual residues, analyzing *clusters* of nearby residues has been found to

be more informative when identifying chemical starting points for the design of small-molecule PPI inhibitors (25). PocketQuery provides an interface for exploring high-level features of a PPI interface and rapidly focusing attention on the key clusters of residues that are likely small-molecule inhibitor starting points.

There are numerous online resources for analyzing and predicting the properties of PPIs (26). In particular, several web servers support the exploration and visualization of properties of PPI interface residues such as sequence conservation (27), surface area calculations (14,28,29), and predicted hot spots (14,22). PocketQuery provides a 3D interface to explore all of these properties, and interface residue properties are precomputed for all PPI structures in the Protein Data Bank (PDB) resulting in the immediate availability of most structures of interest.

PocketQuery is unique in its focus on identifying small-molecule starting points from PPI structure in the form of clusters of interface residues. A cluster of co-located residues provides greater specificity than a single residue, and, unlike the full collection of interface residues, the molecular interactions of a small cluster are accessible to small-molecules. Clusters are ranked according to a ‘druggability’ score where high scoring clusters likely delineate a potential binding site on a receptor surface. PocketQuery is complementary to methods that identify binding sites through an analysis of the receptor (30). Unlike receptor-only methods, PocketQuery requires the full PPI structure, but the residues identified by PocketQuery not only delineate a putative binding site on the protein, they also define a set of molecular interactions selected for by evolution. These interactions define a pharmacophore that can be seamlessly exported into a virtual screening workflow.

MATERIALS AND METHODS

PocketQuery contains an analysis of every PPI structure in the PDB and is updated on a weekly basis as new

*To whom correspondence should be addressed. Tel: +1 412 383 5745; Fax: +1 412 648 3163; Email: dkoes@pitt.edu

structures are made available. Structures are not filtered by experimental method or resolution; users must make their own determination as to whether a structure is of high enough quality to support a meaningful analysis. The first biological assembly deposited in the PDB is analyzed. If no biological assembly is available (e.g. for an NMR structure), then the first model of the asymmetric unit is analyzed. Additionally, users may submit their own structures for analysis. Very large oligomeric structures (such as viral capsids) are reduced to a single monomer and its neighbors to reduce the computational overhead of the analysis. Each structure is preprocessed with CHARMM version 31b1 (31) to add missing atoms, including hydrogens, and optimize hydrogen bonding.

The following energetic, structural and evolutionary properties are computed for each interface residue of a PPI:

ΔG^{FC} : an estimate of the change of free energy (kcal/mol) for a residue upon complexation. Computed using FastContact (12). More negative values indicate a stronger interaction.

$\Delta\Delta G^R$: an estimate of the change in free energy of an alanine mutation. Computed using Rosetta (13). More positive values indicate the mutation destabilizes the complex and thus the original residue has a stronger interaction.

Δ SASA: the change in solvent accessible surface area (SASA) of a residue. This is the difference between the SASA of the bound conformation of a chain in the complexed state and the bound conformation as an independent chain. Computed using naccess (<http://www.bioinf.manchester.ac.uk/naccess/>).

Δ SASA%: the relative Δ SASA as computed by naccess. Expressed as a percentage.

Cons: a conservation score computed using Scorecons (32). A higher score indicates a higher degree of conservation.

Rate: an evolutionary rate computed using Rate4Site (33). A higher score indicates a higher rate and lower degree of conservation.

The full protocol for computing these properties is reported elsewhere (25). Any residue with Δ SASA $>$ 0.05 \AA^2 is treated as an interface residue. All possible clusters of interface residues with a maximal span of 12 \AA are computed. The cluster properties include the aggregated residue properties (minimum, maximum, average and total values) as well as the types of residues in the cluster, the size of the cluster (number of residues) and the maximal distance between cluster residues.

Most significantly, PocketQuery provides a consensus score for each cluster that is derived from a structural analysis of protein–ligand and protein–protein structures (25). The score, which ranges from zero to one, is the output of a support vector machine trained using structures of small-molecules that bind at the PPI interface. When these ligand-bound structures are aligned with the PPI structure, the small-molecule overlaps a cluster of residues in the PPI, and these overlapping residues delineate a small-molecule binding site at the PPI interface.

Clusters with properties that are consistent with these overlapping clusters receive a higher score. That is, a high score suggests that the chemical mimicry of the cluster would be a good starting point for the design of a small-molecule inhibitor of the interaction, and the score provides a rough indication of the ‘druggability’ of the PPI. Since this druggability score is derived from a machine learning structural analysis, its quality and relevance should only increase as more protein–ligand and protein–protein structures are used in its computation. Consequently, the score for all PocketQuery clusters will be updated on a biennial basis to benefit from the constantly increasing amount of structural information.

The PocketQuery interface is implemented using JavaScript and the Java-based Jmol (<http://www.jmol.org/>) molecular viewer. A modern, standards compliant web browser with a recent Java plugin is required.

POCKETQUERY INTERFACE

PocketQuery provides an easy to use graphical interface for searching for clusters, displaying and browsing the clusters that are the results of a search, and visualizing the molecular structure and properties of clusters within the 3D PPI structure. The full results of a search as well as the structures of specific clusters can be exported into common, analyzable and file formats.

Search

The PocketQuery search interface is shown in Figure 1. The interface is divided into a Search panel on the left, where the criteria for the search can be entered, and a table of result clusters on the right. The search criteria includes structure-level information, such as the PDB id and chain, and all the computed cluster properties. An arbitrary number of search criteria may be conjunctively joined to precisely filter the results. For instance, searching for all clusters with a cluster size equal to one and a total ΔG^{FC} (FastContact energy) $<$ -3 would identify all predicted individual hot-spot residues. If a PDB id is not provided, a PDB-wide search is performed. A PDB-wide search can be narrowed by specifying keywords that must match within the PDB title or keywords fields.

As shown in the lower left of Figure 1, a number of preset settings for search criteria are available. Additionally, the search settings may be saved to or restored from a file to enable collaborative sharing of search results.

Clusters

The cluster results table is shown in the right of Figures 1, 2 and 3 and is always visible. Each row of the table corresponds to a single cluster of residues and each column to a cluster property. Clusters may be sorted by any available property by clicking the column heading. As each cluster has dozens of associated properties, it is unwieldy to view them all at once. Instead, as shown in the lower right of Figure 2, different views of the data may be selected to focus in on different aspects of the interface. For example, the Residue Centric view, which is shown in

Search

Average Δ SASA

The average absolute change in solvent accessible surface area of a cluster residue

Between: 44.6233 and

Average Δ SASA%

The average relative change in solvent accessible surface area of a cluster residue

Between: 39.6 and

PDB ID

Search only within structure specified by 4 character PDB identifier

ID: 1YCR

Chain

Cluster must reside on specified chain (case-sensitive)

ID: B

Cluster Size

The number of residues contained within a cluster. Set max to 1 to screen for hotspot residues.

Between: 3 and

Empty
Hot Spot Residues
Anchor Residues
Druggable Site

Add Search Criteria

Preset: Choose... Load... Save... Search

Clusters

PDB	Ch	Sz	Dist	Av ΔG^{FC}	Av $\Delta \Delta G^R$	Av Δ SASA	Av Δ SASA%	Score
1YCR	B	3	10.2463	-5.30333	3.47917	113.723	65.5333	0.954928
1YCR	B	4	10.2463	-4.425	2.90053	98.2475	58.325	0.941059
1YCR	B	3	6.7549	-4.80333	3.8557	106.54	60.4333	0.939554
1YCR	B	5	10.3861	-3.624	2.32342	83.718	53.22	0.925044
1YCR	B	4	10.3861	-4.0825	2.61313	91.6925	57.35	0.920452
1YCR	B	4	6.7549	-3.7075	2.89552	86.305	53.525	0.911256
1YCR	B	3	6.7265	-4.34667	3.4725	97.8	59.1333	0.910822
1YCR	B	5	10.2463	-3.478	2.30814	80.908	48.3	0.900284
1YCR	B	6	10.3861	-2.96833	1.92595	71.69	45.7167	0.895575
1YCR	B	4	10.2463	-3.9	2.59402	88.18	51.2	0.893227
1YCR	B	4	10.3831	-3.8	2.89485	84.005	49.375	0.893056
1YCR	B	3	10.3831	-4.47	3.4716	94.7333	53.6	0.89094
1YCR	B	3	10.2463	-3.89667	2.27077	85.43	56.2	0.889613
1YCR	B	5	10.2463	-3.518	2.30054	79.242	47.06	0.888486
1YCR	B	5	10.3831	-3.124	2.31888	72.324	46.06	0.883666
1YCR	B	6	10.3861	-3.00167	1.91962	70.3017	44.6833	0.883542
1YCR	B	5	10.3861	-3.204	2.07822	75.664	47.52	0.88299
1YCR	B	4	9.9993	-3.525	2.87643	82.7925	47.375	0.881381
1YCR	B	4	10.2463	-3.95	2.58453	86.0975	49.65	0.880317
1YCR	B	3	9.9993	-4.10333	3.44703	93.1167	50.9333	0.879577

<< < 1 2 3 4 5 > >>

85 clusters

Save Results... View: Cluster Averages

Figure 1. The PocketQuery search interface. The Search panel on the left allows the user to set an arbitrary number of conjunctively joined filters, such as specifying a minimum average Δ SASA for each cluster. Preset filtering criteria are available (bottom left) or custom searches may be created and shared using the Load/Save buttons. The matching clusters from the PPI structures deposited in the PDB are displayed in the right panel.

1YCR: MDM2 BOUND TO THE TRANSACTIVATION DOMAIN OF P53

Clusters

PDB	Ch	Sz	Dist	Av ΔG^{FC}	Av $\Delta \Delta G^R$	Av Δ SASA	Av Δ SASA%	Score
1YCR	B	3	10.2463	-5.30333	3.47917	113.723	65.5333	0.954928
1YCR	B	4	10.2463	-4.425	2.90053	98.2475	58.325	0.941059
1YCR	B	3	6.7549	-4.80333	3.8557	106.54	60.4333	0.939554
1YCR	B	5	10.3861	-3.624	2.32342	83.718	53.22	0.925044
1YCR	B	4	10.3861	-4.0825	2.61313	91.6925	57.35	0.920452
1YCR	B	4	6.7549	-3.7075	2.89552	86.305	53.525	0.911256
1YCR	B	3	6.7265	-4.34667	3.4725	97.8	59.1333	0.910822
1YCR	B	5	10.2463	-3.478	2.30814	80.908	48.3	0.900284
1YCR	B	6	10.3861	-2.96833	1.92595	71.69	45.7167	0.895575
1YCR	B	4	10.2463	-3.9	2.59402	88.18	51.2	0.893227
1YCR	B	4	10.3831	-3.8	2.89485	84.005	49.375	0.893056
1YCR	B	3	10.3831	-4.47	3.4716	94.7333	53.6	0.89094
1YCR	B	3	10.2463	-3.89667	2.27077	85.43	56.2	0.889613

<< < 1 2 3 4 5 6 7 > >>

83 clusters

Save Results... View: Cluster Averages

Residue Centric
Cluster Averages
Cluster Totals
Energy Centric (FastContact)
Energy Centric (Rosetta)
Absolute Δ SASA Centric
Relative Δ SASA Centric
Conservation Centric
Everything

Residues

Res	#	ΔG^{FC}	$\Delta \Delta G^R$	Δ SASA	Δ SASA%	SASA	Cons	Rate
PHE	19	-6.61	5.61	131.10	79.90	2.61	0.00	0.00
TRP	23	-6.01	4.79	136.70	64.70	9.06	0.00	0.00
LEU	26	-3.29	0.04	73.37	52.00	1.86	0.00	0.00

Figure 2. Browsing the PocketQuery search results. The Clusters panel on the right displays the table of clusters matching the specified search criteria. Different consolidated views of the many available cluster properties can be selected, as shown in the bottom right. Results may be sorted by any numerical property. Here they are shown sorted by a 'druggability' score, where higher score indicates the similarity of a cluster to residue clusters known to delineate small-molecule binding sites at a PPI interface. Selecting a result cluster in the results table brings up a molecular viewer panel featuring a 3D Jmol interface of the PPI structure. The residues of the selected cluster(s) and their properties are shown in the bottom of the molecular viewer panel.

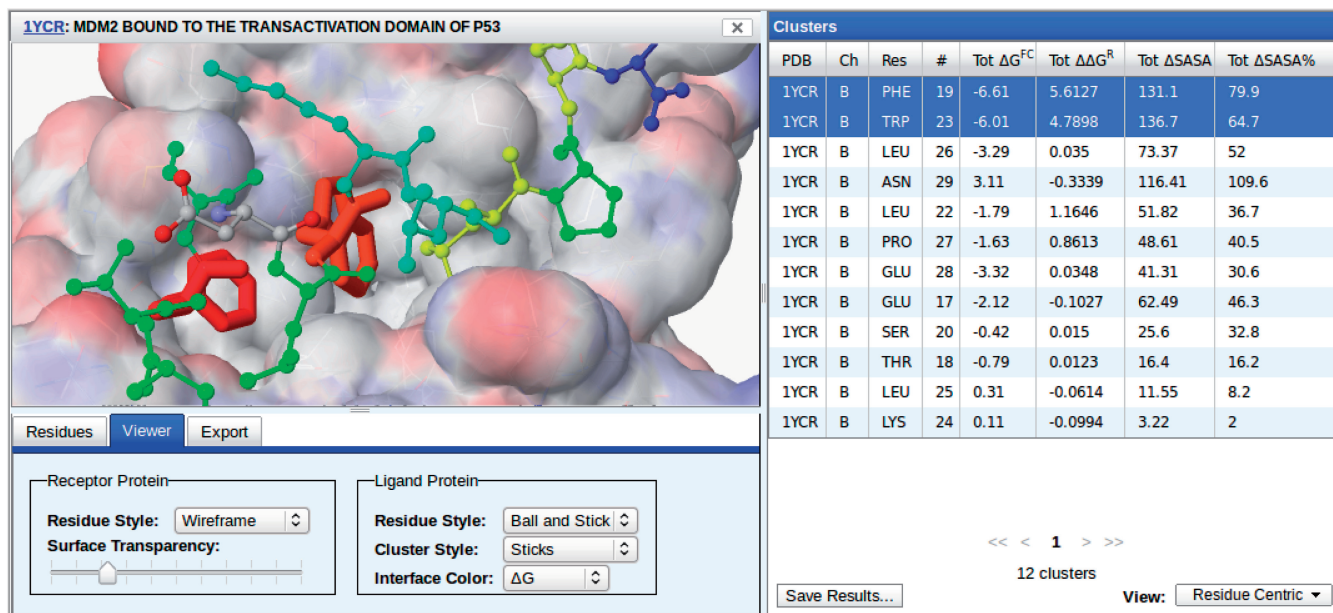


Figure 3. Visualizing the properties of a protein–protein interface. A variety of display styles are available for the receptor protein, ligand protein and selected cluster(s). The receptor surface is shown color mapped by partial charge. The color of the interface residues of the ligand protein (including the cluster residues) may be color mapped to any of the computed properties, such as energy estimates, sequence conservation or $\Delta SASA$.

Figure 3 and is most appropriate when searching for only clusters of size one, displays the residue names and sequence identifiers, but omits the cluster size and maximum distance (which are always one and zero for single-residue clusters). The full search results, consisting of all the computed properties for every matching cluster, can be downloaded as a comma separated text file through the Save Results button.

Clusters may be visualized in the context of the PPI by clicking on the corresponding row. As shown in Figure 2, this launches a molecular viewer in the left panel of the interface, replacing the search panel. If multiple clusters from the same chain are selected (i.e. using the Ctrl and Shift keys) as shown in Figure 3, their residues are merged into a single cluster in the molecular viewer panel. The residue properties of the displayed cluster are shown in the bottom of the molecular viewer panel, as shown in Figure 2.

Visualization

The molecular viewer may be manipulated using the standard Jmol controls. Additionally, the viewer panel, shown in Figure 3, can be used to customize the visual styles of the receptor protein, ligand protein and cluster residues. The ligand protein is the single chain that contains the selected cluster whereas the receptor protein consists of all other chains in the PPI complex. The default display style is for the receptor and ligand protein residues to be shown as wireframes with the cluster residues shown as sticks, as shown in Figure 2. The receptor protein is displayed with a rendered surface that is color mapped by the partial charge of the residues (the surface may be omitted by setting the transparency to 100%). The interface residues of the ligand protein may be colored

by residue property, as shown in Figure 3. Property values are mapped to a rainbow (ROYBG) spectrum where the smallest values are always red and the largest blue.

Export

Once a cluster has been identified as a promising starting point for small-molecule design, it can be saved as a PDB file or exported directly into a pharmacophore-based virtual screening workflow. This functionality is provided under the Export tab of the molecular viewer panel. Export is available for two online virtual screening search engines: AnchorQuery (<http://anchorquery.csb.pitt.edu>) and ZINCPharmer (<http://zincpharmer.csb.pitt.edu>). Both AnchorQuery and ZINCPharmer screen for compounds using a 3D pharmacophore, the spatial arrangement of the essential features (such as hydrophobic regions or hydrogen bonds) of an interaction. A candidate pharmacophore is derived directly from the cluster–receptor structure exported by PocketQuery and can be easily refined by the user.

AnchorQuery is a specialized interactive pharmacophore search technology. AnchorQuery includes a library of > 1.5 billion conformers of > 21 million novel chemical compounds. These compounds are accessible through one-step multi-component reaction chemistry. The compounds are designed to be biased to target PPIs by virtue of always containing a functional mimic of a specific amino acid (the ‘anchor’ feature of the pharmacophore). AnchorQuery provides rapid access to a large chemical space specialized for targeting PPIs, but synthesis is required for experimental validation.

ZINCPharmer is a general interactive pharmacophore search technology for the ZINC database (34).

ZINCPharmer is routinely updated to search the most recent set of purchasable compounds from ZINC. The size of this library is on the order of 100–200 million conformations of 10–20 million compounds. Although historical collections of compounds are thought to be poorly suited for targeting PPIs (5), compounds found through ZINCPharmer have the advantage that they may be immediately purchased for experimental validation.

DISCUSSION

PocketQuery includes precomputed properties for all the PPI structures in the PDB resulting in the immediate availability of almost all structures of interest. Not only are multiple properties computed, such as energy estimates and sequence conservation, but these properties are also computed with multiple methods. Property values can be displayed both numerically and visually color-mapped directly onto the PPI structure. This allows for a critical assessment of the potential contribution of an interface residue in the formation and stabilization of the PPI complex.

PocketQuery is a valuable tool for the interactive, user-guided exploration of PPI interfaces, and the analysis of clusters of residues combined with the consensus ‘druggability’ score results in the rapid focusing of attention on the key residues of the interaction. Once an appropriate set of residues is identified, PocketQuery provides an entry point into an interactive online drug discovery workflow for the development of PPI inhibitors. PocketQuery is freely available at <http://pocketquery.csb.pitt.edu>

FUNDING

National Institutes of Health (NIH) [R01GM097082]. Funding for open access charge: NIH [1R21GM087617].

Conflict of interest statement. None declared.

REFERENCES

- Wells,J. and McClendon,C. (2007) Reaching for high-hanging fruit in drug discovery at protein–protein interfaces. *Nature*, **450**, 1001–1009.
- Dömling,A. (2008) Small molecular weight protein-protein interaction antagonists—an insurmountable challenge? *Curr. Opin. Chem. Biol.*, **12**, 281–291.
- Whitty,A. and Kumaravel,G. (2006) Between a rock and a hard place? *Nat. Chem. Biol.*, **2**, 112–118.
- Macarron,R. (2006) Critical review of the role of HTS in drug discovery. *Drug Discov. Today*, **11**, 277–279.
- Sperandio,O., Reynès,C.H., Camproux,A.C. and Villoutreix,B.O. (2010) Rationalizing the chemical space of protein-protein interaction inhibitors. *Drug Discov. Today*, **15**, 220–229.
- Popowicz,G., Dömling,A. and Holak,T. (2011) The structure-based design of Mdm2/Mdmx–p53 inhibitors gets serious. *Angew. Chem. Int. Ed.*, **50**, 2680–2688.
- Liu,S., Wu,S. and Jiang,S. (2007) HIV entry inhibitors targeting gp41: from polypeptides to small-molecule compounds. *Curr. Pharm. Des.*, **13**, 143–162.
- Christ,F., Voet,A., Marchand,A., Nicolet,S., Desimie,B.A., Marchand,D., Bardiot,D., Van derVeken,N.J., Van Remoortel,B. and Strelkov,S.V. (2010) Rational design of small-molecule inhibitors of the LEDGF/p75-integrase interaction and HIV replication. *Nat. Chem. Biol.*, **6**, 442–448.
- Rajamani,D., Thiel,S., Vajda,S. and Camacho,C.J. (2004) Anchor residues in protein-protein interactions. *PNAS*, **101**, 11287.
- Clackson,T. and Wells,J.A. (1995) A hot spot of binding energy in a hormone-receptor interface. *Science*, **267**, 383–386.
- Moreira,I., Fernandes,P. and Ramos,M. (2007) Hot spots: a review of the protein–protein interface determinant amino-acid residues. *Proteins: Struct. Funct. Bioinf.*, **68**, 803–812.
- Camacho,C. and Zhang,C. (2005) FastContact: rapid estimate of contact and binding free energies. *Bioinformatics*, **21**, 2534.
- Kortemme,T., Kim,D. and Baker,D. (2004) Computational alanine scanning of protein-protein interfaces. *Sci. STKE*, **2004**, 12.
- Meireles,L., Dömling,A.S. and Camacho,C.J. (2010) ANCHOR: a web server and database for analysis of protein-protein interaction binding pockets for drug discovery. *Nucleic Acids Res.*, **38**, W407.
- Bromberg,Y. and Rost,B. (2008) Comprehensive in silico mutagenesis highlights functionally important residues in proteins. *Bioinformatics*, **24**, i207.
- Keskin,O., Ma,B. and Nussinov,R. (2005) Hot regions in protein-protein interactions: the organization and contribution of structurally conserved hot spot residues. *J. Mol. Biol.*, **345**, 1281–1294.
- Ofran,Y. and Rost,B. (2007) Protein-protein interaction hotspots carved into sequences. *PLoS Comput. Biol.*, **3**, e119.
- Lichtarge,O., Bourne,H.R. and Cohen,F.E. (1996) An evolutionary trace method defines binding surfaces common to protein families. *J. Mol. Biol.*, **257**, 342–358.
- Cho,K., Kim,D. and Lee,D. (2009) A feature-based approach to modeling protein–protein interaction hot spots. *Nucleic Acids Res.*, **37**, 2672.
- Lise,S., Archambeau,C., Pontil,M. and Jones,D. (2009) Prediction of hot spot residues at protein-protein interfaces by combining machine learning and energy-based methods. *BMC Bioinformatics*, **10**, 365.
- Darnell,S., Page,D. and Mitchell,J. (2007) An automated decision-tree approach to predicting protein interaction hot spots. *Proteins Struct. Funct. Bioinf.*, **68**, 813–823.
- Zhu,X. and Mitchell,J.C. (2011) KFC2: a knowledge-based hot spot prediction method based on interface solvation, atomic density, and plasticity features. *Proteins Struct. Funct. Bioinf.*, **79**, 2671–2683.
- Tuncbag,N., Gursoy,A. and Keskin,O. (2009) Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics*, **25**, 1513.
- Guney,E., Tuncbag,N., Keskin,O. and Gursoy,A. (2008) HotSprint: database of computational hot spots in protein interfaces. *Nucleic Acids Res.*, **36**, D662.
- Koes,D.R. and Camacho,C.J. (2012) Small-molecule inhibitor starting points learned from protein-protein interaction inhibitor structure. *Bioinformatics*, **28**, 784–791.
- Tuncbag,N., Kar,G., Keskin,O., Gursoy,A. and Nussinov,R. (2009) A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. *Brief. Bioinform.*, **10**, 217–232.
- Glaser,F., Pupko,T., Paz,I., Bell,R.E., Bechor-Shental,D., Martz,E. and Ben-Tal,N. (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics*, **19**, 163.
- Gabdoulline,R.R., Wade,R.C. and Walther,D. (2003) MolSurfer: a macromolecular interface navigator. *Nucleic Acids Res.*, **31**, 3349–3351.
- Dundas,J., Ouyang,Z., Tseng,J., Binkowski,A., Turpaz,Y. and Liang,J. (2006) CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.*, **34**, W116–W118.

30. Henrich,S., Salo-Ahen,O., Huang,B., Rippmann,F., Cruciani,G. and Wade,R. (2010) Computational approaches to identifying and characterizing protein binding sites for ligand design. *J. Mol. Recogn.*, **23**, 209–219.
31. Brooks,B.R., Bruccoleri,R.E. and Olafson,B.D. (1983) CHARMM: a program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **4**, 187–217.
32. Valdar,W.S.J. (2002) Scoring residue conservation. *Proteins Struct. Funct. Bioinf.*, **48**, 227–241.
33. Mayrose,I., Graur,D., Ben-Tal,N. and Pupko,T. (2004) Comparison of site-specific rate-inference methods for protein sequences: empirical Bayesian methods are superior. *Mol. Biol. Evol.*, **21**, 1781.
34. Irwin,J.J. and Shoichet,B.K. (2005) ZINC- a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model*, **45**, 177–182.