Biochemical and Biophysical Research Communications 445 (2014) 717–723

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

PeptiSite: A structural database of peptide binding sites in 4D

Chayan Acharya [a], Irina Kufareva [a], Andrey V. Ilatovskiy [a,b,c], Ruben Abagyan [a,*]

[a] UCSD, Skaggs School of Pharmacy and Pharmaceutical Sciences, La Jolla, CA 92093, USA
[b] Division of Molecular and Radiation Biophysics, Petersburg Nuclear Physics Institute, Gatchina 188300, Russia
[c] Research and Education Center "Biophysics", PNPI and St. Petersburg State Polytechnical University, St. Petersburg 195251, Russia

ARTICLE INFO

Article history:
Received 28 November 2013
Available online 6 January 2014

Keywords:
PeptiSite
Peptide binding site
Pocketome
Induced-fit
SiteFinder

ABSTRACT

We developed PeptiSite, a comprehensive and reliable database of biologically and structurally characterized peptide-binding sites, in which each site is represented by an ensemble of its complexes with protein, peptide and small molecule partners. The unique features of the database include: (1) the ensemble site representation that provides a fourth dimension to the otherwise three dimensional data, (2) comprehensive characterization of the binding site architecture that may consist of a multimeric protein assembly with cofactors and metal ions and (3) analysis of consensus interaction motifs within the ensembles and identification of conserved determinants of these interactions. Currently the database contains 585 proteins with 650 peptide-binding sites. http://peptisite.ucsd.edu/ link allows searching for the sites of interest and interactive visualization of the ensembles using the ActiveICM web-browser plugin. This structural database for protein–peptide interactions enables understanding of structural principles of these interactions and may assist the development of an efficient peptide docking benchmark.

1. Introduction

Molecular recognition is the central element of biological communication. The endogenous ligand binding sites on receptors are often targeted by therapeutics [1–5]. Interactions between globular proteins and short peptides constitute 15–40% of all protein interactions and are vital for various cellular processes [6,7]. Peptides and peptidemimetics are increasingly attracting attention as therapeutics [7–9]. Currently around 60 peptide drugs are sold worldwide.

Incomplete understanding of the receptor–ligand non-covalent interaction is one of the major limitations for lead discovery [10]. Computational modeling is used to complement the experimental studies for explaining the structural features of the protein–ligand interactions [11–13]. Such computational studies greatly depend on the crystallographic static structure of the protein–peptide complexes [14–16]. However, proteins often undergo significant conformational changes, known as induced fit, to accommodate ligands of various shapes, sizes and chemical properties in the same binding site [17–23]. Such conformational changes may range from small-scale side-chain rearrangement to large-scale loop movement or domain shifts [20,24–28]. Hence, it is difficult to develop an unambiguous structural model of the binding pockets that can be used in computational docking studies [21,29–34]. The effect of induced fit becomes more crucial for binding sites interacting with peptides due to higher number of conformational degrees of freedom of the ligands.

Many efforts have been undertaken to account for the structural flexibility in predicting the active conformation of the binding sites [20,24,35–53], including building of databases of experimental structures of the protein complexes with small molecule and peptide ligands [7,54–60]. Comparison of multiple crystal structures of a given protein bound to various ligands can reveal the conformational space that needs to be considered to obtain a correct definition of the binding pocket. Our laboratory recently developed an database of small-molecule binding pockets, known as Pocketome [60]. Here we extended the Pocketome to build a database, called PeptiSite, with a comprehensive collection of crystal structures of various receptors interacting with peptide ligands that are reported in the Protein Data Bank (PDB) as of May 2013 [61]. A 3D structure of a single complex gives only a limited static view of this functionality; however, an ensemble of the multiple snapshots of each individual site adds the fourth dimension to the data [62], which not only allows separation of spurious or permanent complexes from the relevant transient interactions, but also provides insights into the effect of conformational variability on protein–peptide interactions.

* Corresponding author. Address: University of California, San Diego, Skaggs School of Pharmacy and Pharmaceutical Sciences, 9500 Gilman Drive, La Jolla, CA 92093, USA. Fax: +1 858 822 5591.
E-mail address: rabagyan@ucsd.edu (R. Abagyan).

http://dx.doi.org/10.1016/j.bbrc.2013.12.132
0006-291X/

Unlike PeptiSite, other databases such as DOMINO [63], MPID [64], PCIDB [54], MOAD [55], IBIS [56,65], or ReliBase [57] mainly collect, enrich, and make inferences from individual PDB structures of protein complexes with small chemicals or peptides. The approach employed in PeptiSite shares some similarity with those of PCDB [58], PepX [7], PepSite [6] or IEDB-3D [59], however, PeptiSite specifically focuses on the detailed analysis of the peptide binding sites. Besides identifying the constituents of the receptor–peptide binding sites, the subsequent processing of the PeptiSite ensembles creates accurate interaction maps between the binding site and the peptide ligand residues, quantifies cross-compatibility between pockets and ligands from different structures.

The PeptiSite can assist the detection of the conserved determinants of molecular interactions, understanding the effects of SNPs and single-point mutations, explanation of induced fit phenomena, and development of flexible docking algorithms. It may also be used for structure-based prediction of novel activities of existing molecules or for activity and binding mode prediction of the new compounds [66–68]. With its unique interface providing versatile interactive molecular visualization, the PeptiSite is a valuable resource in understanding molecular interactions between protein and peptides.

2. Methods

2.1. Construction of the database

Each PeptiSite database entry contains an ensemble of the co-crystallized and superimposed structures of a peptide-binding site, along with the associated information. The definitions of various terms such as *pocket* and *site* used in our database are consistent with the Pocketome encyclopedia. A *pocket* describes the sets of atoms that are in contact with the ligand in a single co-crystallized structure, and a *site* represents the set of residues that have been experimentally observed to participate in ligand binding in at least one of the complexes in an ensemble. The detailed description of these terms can be found in [60]. Each member of an ensemble retains its individual characteristics such as the set of pocket atoms and the bound ligand.

One binding site may either reside only on a single protein chain or on a multimeric assembly interface. Currently, the maximum allowed number of hetero-multimeric partners comprising the binding site in PeptiSite database is two; however, there is no limit on the number of chains for homo-multimeric interface sites. PeptiSite entries do not use PDB BioMT records. One binding site in PeptiSite may or may not correspond to the protein biological unit [60].

A typical binding site may also include *cofactor* molecules (NAD, ATP etc.) and coordinated *metal ions*. A *cofactor* is defined as a molecule that non-competitively binds to the binding pocket simultaneously with the ligand molecule and essential for the biological

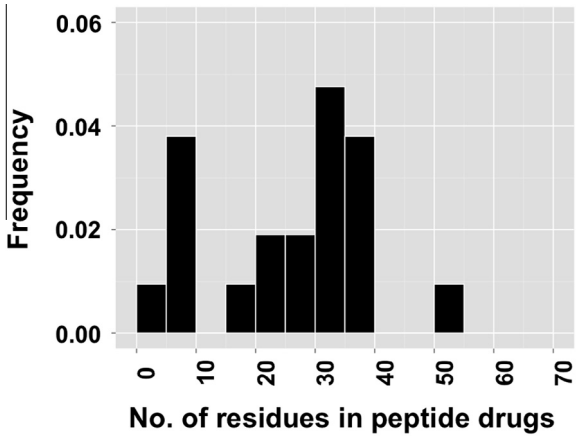

Fig. 2. Distribution of the number of residues of peptide drugs.

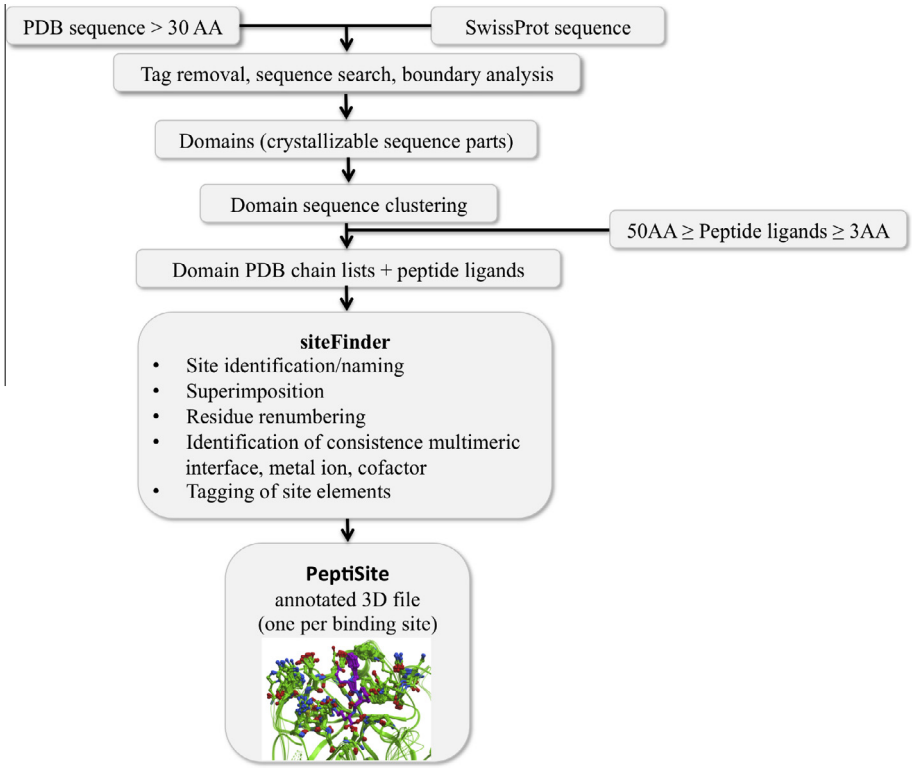

Fig. 1. PeptiSite data flow with the siteFinder algorithm. The output of the siteFinder utility in the form of tagged 3D ensembles forms the core of the PeptiSite database.

Table 1
Basic statistics of PeptiSite database.

Features	Median	Mean	Min:max
Number of structures	6	13	1:146
Number of PDB	4	10	1:80
Number of PDB/site with ligand	2	3	1:75
Number of receptor residues	423	609	45:7073
Number of peptide residues	11	15	3:50
Number of binding pocket residues	20	22	2:119
% Of hydrophobic residues in binding pocket	46	45	0:89
% Of polar residues in binding pocket	25	27	0:85
% Of acidic residues in binding pocket	11	12	0:60
% Of basic residues in binding pocket	15	16	0:75
ΔSASA (Å^2)	1103	1247	309:4170
$RMSD_{BB}$ (Å)	0.57	1.2	0.032:16
$RMSD_{sidechain}$ (Å)	1.3	1.8	0.032:17

activation of the receptor. Similarly *metal ions* consistently present in multiple crystal structures of a binding site are also assumed to be essential for the activation of the receptor. Each PeptiSite entry contains a ligand ensemble that may include other entities that bind competitively with these peptides, such as drug-like small molecules or nucleic acids. In some cases two or more ligands may concurrently occupy the space that contains a single ligand molecule in other pockets within the same ensemble.

The online version of the database includes only the non-redundant set of pocket compositions (i.e. distinct ligands in combinations with distinct point mutants only). However, complete sets of the related structures of any PeptiSite entry are available on request.

2.2. PeptiSite data flow and filtering criteria

The PeptiSite database is built based on the siteFinder algorithm which is explained in detail in the Pocketome (http://www.pocketome.org) [60]. In brief, siteFinder automatically collects, clusters, analyzes and validates the binding pocket structures based on consistency of their composition and spatial configuration between the multiple members of the structural ensemble. The siteFinder also clusters the highly homologous ($\geqslant$94% sequence identity) proteins originating from different organisms into one entry to minimize unnecessary fragmentation and enhance the representative

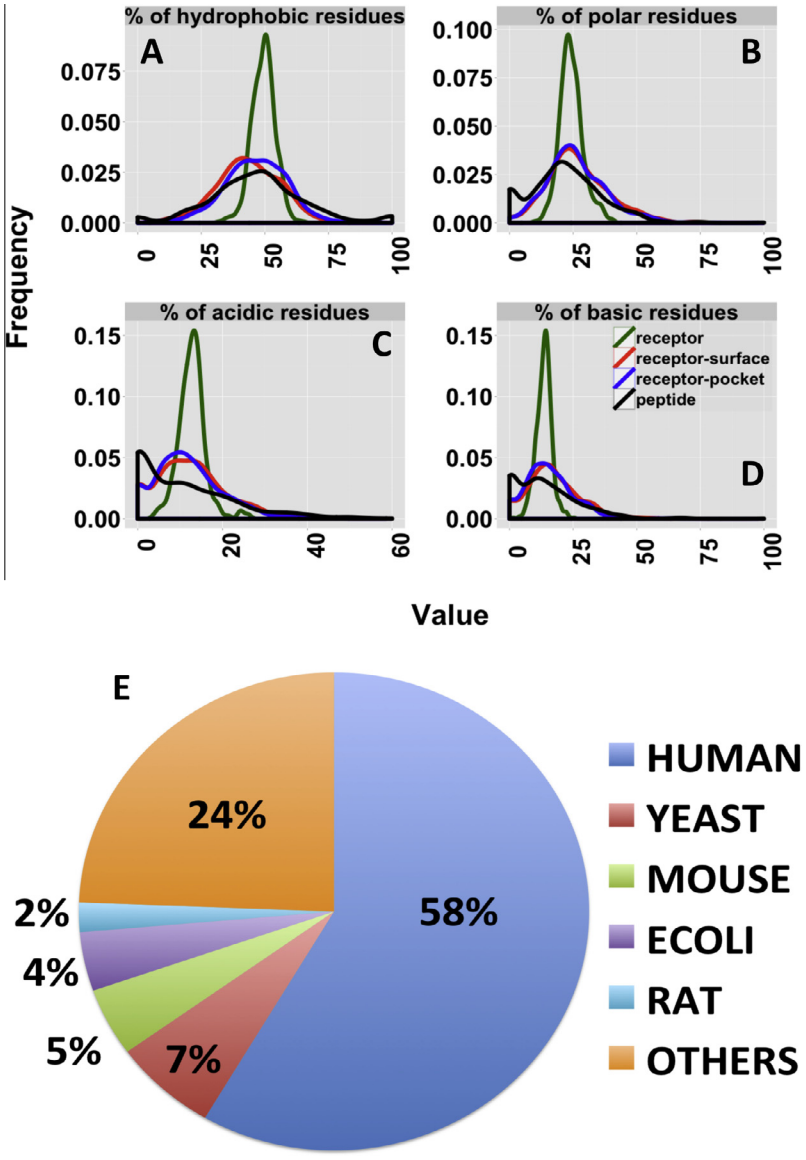

Fig. 3. (A) Distribution of the hydrophobic (G, A, V, L, I, F, W, C, M, P); (B) polar uncharged (S, T, Y, N, Q); (C) acidic (D and E) and (D) basic (H, R, K) residues on the receptor (green), receptor surface (red), binding site (blue) and peptide ligand (black). (E) Distribution of the organisms in PeptiSite.

structures of the corresponding ensembles. Fig. 1 depicts the schema of the PeptiSite.

Each peptide-binding site in PeptiSite satisfies the following criteria:

(i) Receptor must belong to SwissProt.
(ii) Protein is co-crystallized with at least one peptide ligand.

The allowed length of the peptide ligand in PeptiSite ranges from 3 to 50 amino acid residues. The choice of this criterion is based on the distribution of the number of residues of peptide drugs (Fig. 2) available in DrugBank [69,70]. Fig. 2 showed that more than 95% of the peptide drugs satisfy this condition. The siteFinder is launched with two input files. The first file includes the swissProt ID of the receptor along with their peptide-binding domain boundaries, permanent protein partners and associated PDB IDs. The second file includes a list of one *prototypical* peptide ligand, called seed ligands, for each receptor along with the PDB ID of the corresponding receptor–peptide complex and chain ID of the peptide ligand inside the PDB entry. The selection of the seed ligand is performed based on the following criteria:

(i) $3 \leqslant$ Peptide length $\leqslant 50$.
(ii) Bound within the domain boundaries.
(iii) Contains only natural amino acids.
(iv) At least one of the two conditions is met upon ligand binding to the receptor:
  (a) Loss of ligand solvent accessible surface area (fracBuried) $\geqslant 25\%$.

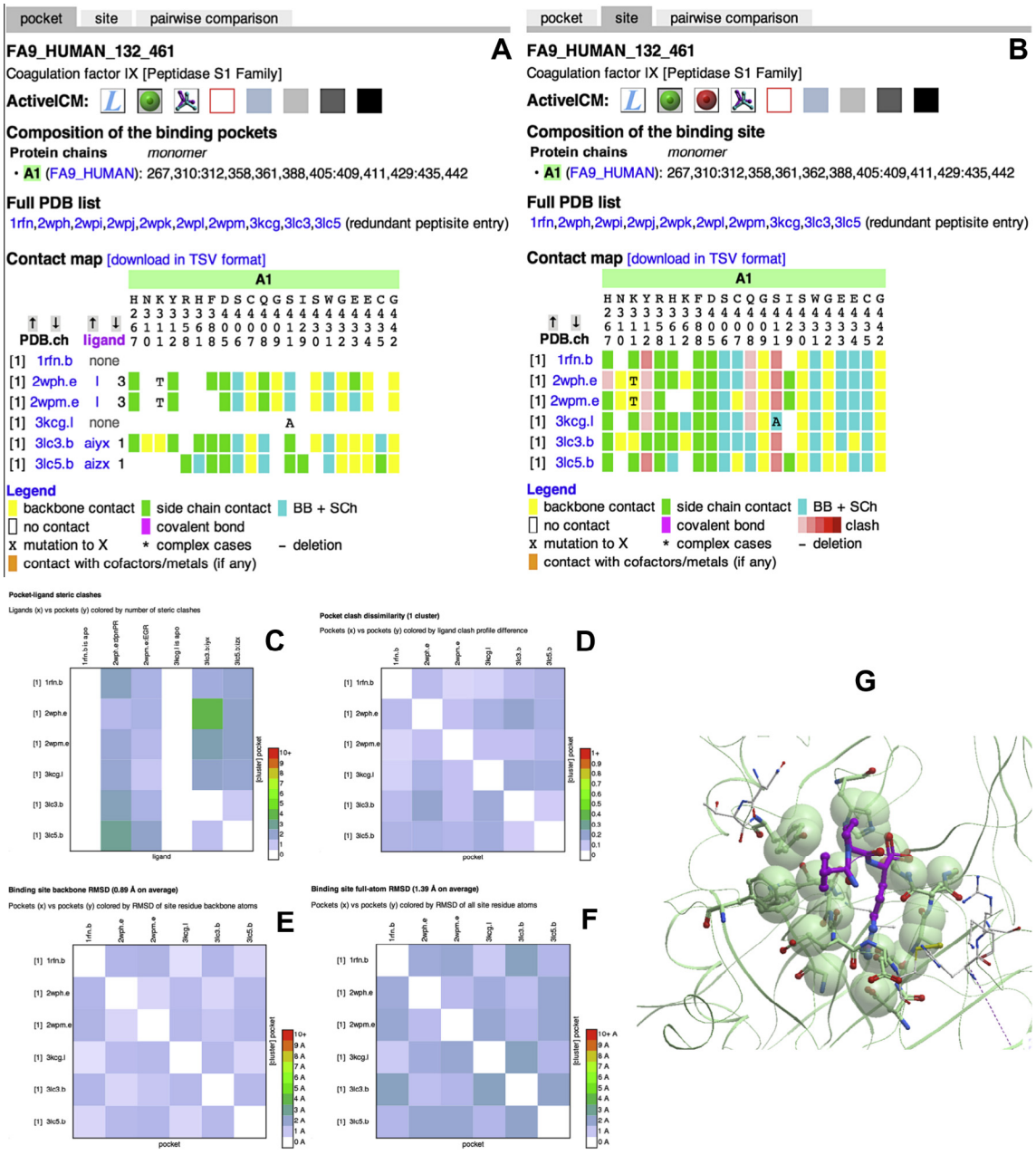

Fig. 4. Summary of a single PeptiSite page content exemplified by the entry for human coagulation factor IX receptor. The binding site is composed of a single polypeptide chain (FA9_HUMAN). The list of binding site residues, full PDB list with links are shown at the top of the page. (A) Pocket contact map; (B) site contact map; (C) pairwise pocket-ligand and (D) pairwise pocket-pocket steric compatibility; (E) and (F) represent the backbone and full-atom RMSD of the binding site residues, respectively. (G) A snapshot of the ActiveICM ensemble visualization window.

(b) Loss of receptor solvent accessible surface area (ΔSASA) $\geqslant$ 300 Å^2.

The following equations are used to calculate fracBuried and ΔSASA:

$$\text{fracBuried} = (\text{SASA}_{\text{unbound}} - \text{SASA}_{\text{bound}})/\text{SASA}_{\text{unbound}} \quad (1)$$

$$\Delta\text{SASA} = \text{SASA}_{\text{receptor}} + \text{SASA}_{\text{ligand}} - \text{SASA}_{\text{complex}} \quad (2)$$

where $\text{SASA}_{\text{unbound}}$ is the SASA of the unbound ligand, $\text{SASA}_{\text{bound}}$ is the SASA of the bound ligand, $\text{SASA}_{\text{receptor}}$ is the SASA of the unbound receptor protein, $\text{SASA}_{\text{ligand}}$ is the SASA of the unbound ligand and $\text{SASA}_{\text{complex}}$ is the SASA of the receptor–ligand complex.

2.3. PeptiSite statistics

The current release of the PeptiSite database (November 2013) contains 650 entries, of which 380 (58%) correspond to human proteins. There are 158 entries containing small molecule ligand along with the peptide ligands. There are 585 different receptor proteins originating from 107 different organisms in the current release. Table 1 shows the basic statistics about the database. The current release of PeptiSite contains 7033 unique PDB entries. The number of pocket structures for a full PeptiSite entry varies from 1 to 146 (median 6), and the number liganded structures for a full PeptiSite entry ranges from 1 to 75 (median 2). RMS deviations (RMSD) of the backbone and the side chains of the binding pocket residues range from 0.0032 Å to 16 Å (median 0.57 Å) and from 0.032 Å to 17 Å (median 1.3 Å), respectively. The length of the peptide ligand ranges from 3 to 50 with a median value of 11 residues, while the length of the receptor proteins included in PeptiSite ranges from 45 to 7073 residues with a median of 423 residues. The loss of solvent accessible surface of the peptide ligand upon binding (ΔSASA) ranges from 309 Å^2 to 4170 Å^2 with a median of 1103 Å^2. Panels A–D of Fig. 3 represent the distribution of the hydrophobic (G, A, V, L, I, F, W, C, M, P), polar uncharged (S, T, Y, N, Q), acidic (D, E) and basic (H, R, K) residues forming the entire receptor (green), surface residues of the receptor (red), binding pocket (blue) and peptide ligand (black), respectively. Fig. 3E represents a pie chart of the distribution of the organisms present in the database.

3. PeptiSite Access

3.1. Search options

The online version of the PeptiSite is available at http://peptisite.ucsd.edu/. In the current release this resource allows search by:

- Receptor swissProt ID.
- Peptide ligand swissProt ID.
- Protein name, family or domain.
- PDB ID.

The search results are always returned in the form of a hit list represented by the site identifier, protein name and family, the list of all associated PDB entries. Clicking on the site identifier takes the user to the individual entry page with peptide–protein contact analysis, pocket pair-wise comparison, and interactive visualization of the site ensemble (Fig. 4).

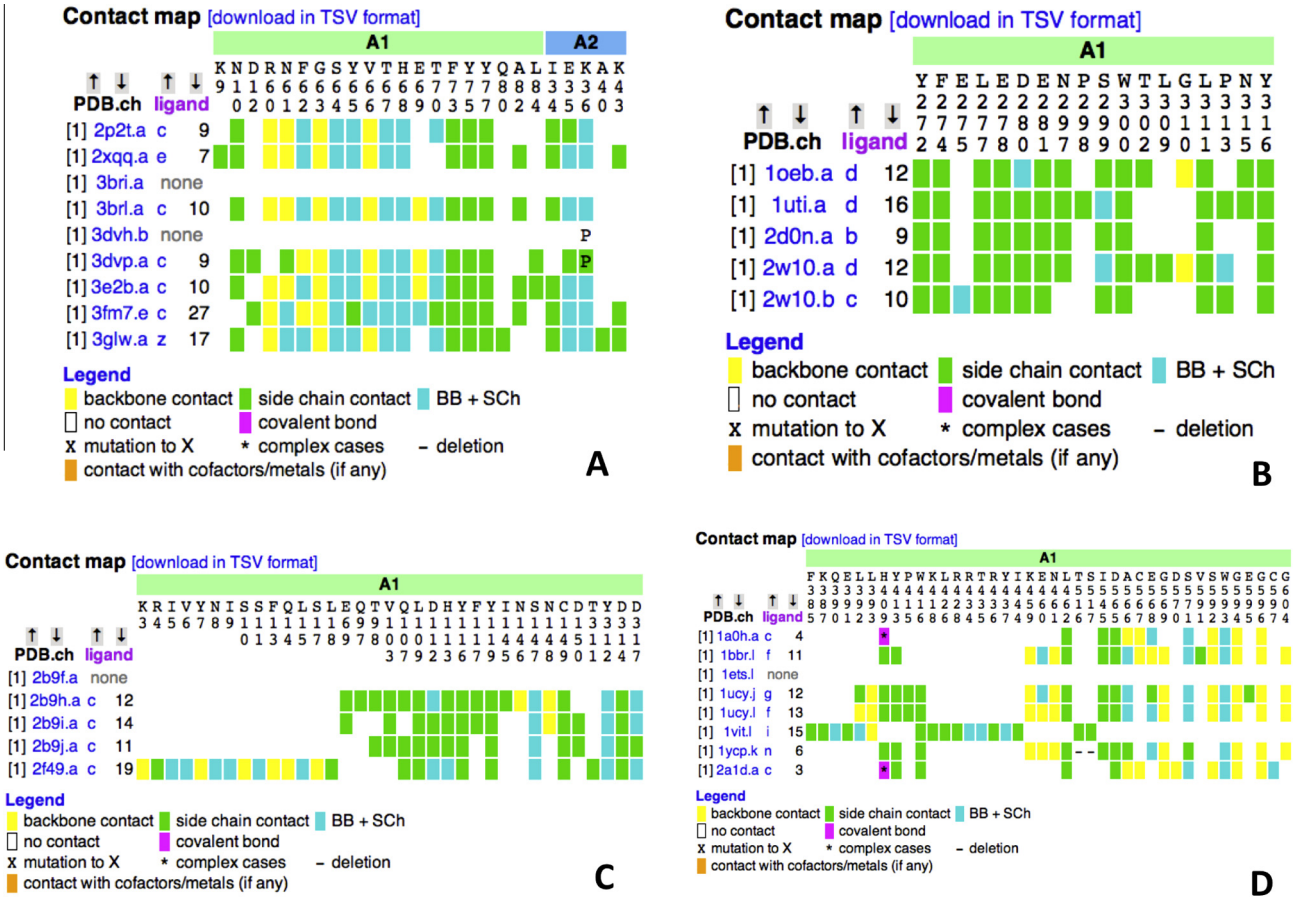

Fig. 5. The contact map obtained from the *pocket* tab of (A) DYL2_HUMAN_1_89 and (B) GRAP2_MOUSE_264_322 showing conserved receptor–peptide interaction pattern; similar maps of (C) FUS3_YEAST_1_353 and (D) THRB_BOVIN_367_518 indicating diverse interaction patterns between the receptor binding site and different peptide ligands.

3.2. Peptide-pocket contact analysis

Two of the three tabs in the text frame of a PeptiSite entry are named *pocket* and *site*. In addition to the summary information on the pocket/site composition, both tabs present the results of analysis of ligand contacts with the neighboring protein residues and other molecules in the site (Fig. 4A and B). We consider two atoms being in contact if their centers are separated by less than 120% of the sum of their van der Waals radii. A residue is said to be in contact with the peptide ligand if at least one non-hydrogen atom of the ligand contacts this residue. In the *pocket* tab, the contacts of each individual ligand with its binding pocket are described, allowing quick identification of conserved or ligand-specific interactions. In contrast, the *site* tab summarizes the ligand ensemble information, most importantly, the contacts that other ligands from the ensemble could make if they were placed into each selected pocket structure. Contacts are color-coded as backbone, side-chain, or both. As ligands from other structures may be sterically incompatible with a selected pocket, the *site* tab uses an additional type of contact not present in the *pocket* tab, a ligand-residue *steric clash*. A clashing residue is colored red with the color intensity corresponding to the number of ligands spatially overlapping with that residue.

Cases of residue mutations from the reference swissProt sequence, as well as cases of residue covalent modifications, deletions, or insertions are marked in the *pocket* and *site* matrices. The contact maps are interactively clickable leading to changes in the 3D graphics window on the right.

3.3. Pairwise comparison of binding pockets

Understanding binding site flexibility and induced fit effects is one of the primary goals of the PeptiSite database. Therefore, each PeptiSite entry is complemented by a *pairwise comparison* tab, with the results of the steric compatibility of the corresponding ligand ensembles (Fig. 4C and D). For that comparison, (i) each pocket was described by a vector of steric clashes that it makes with the ligand ensemble, (ii) pocket distances were calculated as normalized absolute difference between the vectors, and (iii) all pockets within the entry were clustered according to the calculated distance to form sterically cross-compatible subsets. This tab also includes the straightforward plots showing the pairwise backbone and full-atom RMSD between the pockets (Fig. 4E and F). These plots give an idea about the degree of the conformational variability of the sites.

The receptor–peptide contact map on the *pocket* tab of each entry may be used to assess the flexibility of the binding pocket. For example, the contact map of DYL2_HUMAN_1_89 (Fig. 5A) displays that 17 out of 25 binding-site residues are in contact with most of the peptide ligands indicating a highly conserved interaction between this binding site and its ligands. Similarly, the contact map of GRAP2_MOUSE_264_322 (Fig. 5B) displays that 10 out of 18 binding-site residues are in contact with all of the peptide ligands. In contrast, FUS3_YEAST_1_353 and THRB_BOVIN_367_518 (Fig. 5C and D) display that their corresponding binding sites can yield different interaction patterns between the binding-site residues and the different peptides. The contact map of FUS3_YEAST_1_353 shows that peptide "*c*" occupies an extended region of the binding site (PDBID 2F49), which is not accessed by any other ligands. Similarly, THRB_BOVIN_367_518 shows that peptide "*i*" (PDBID 1VIT) yields a unique interaction pattern, which is different form the rest of the ligands. Such information about a peptide-binding site can be combined with the corresponding ICM binary file (available at http://peptisite.ucsd.edu) to analyze more specific details (interacting receptor–peptide residue pairs, binding mode etc.) that can, in turn, be used for peptide docking studies.

3.4. ActiveICM visualization

PeptiSite employs the ActiveICM [71] (version 1.1–6 or higher) for the interactive visualization of the binding site ensembles. ActiveICM is a freely available plug-in for all major web-browsers with unique capabilities not accessible in other similar tools. ActiveICM is widely accepted and used by the computational and structural biology communities as an advanced format for 3D data visualization [72,73]. In the PeptiSite pages, the molecular entities forming the binding site are consistently color-coded according to their type (e.g. polypeptide receptor chains, co-factors, metal ions, or the ligands) and are presented in several convenient and informative views, which are controlled by the links on the left side of the screen, in the *pocket* and *site* tabs. Generic ActiveICM controls enable view manipulations like rotation, zooming in and out, centering on selected parts, clipping, rocking, etc. An added advantage of using the ActiveICM and data format is that the PeptiSite pages are viewable on all Apple portable devices using the sister iMolview application.

Funding

This work is partially supported by NIH Grants R01 GM071872, U01 GM094612, U54 GM094618, and RC2 LM010994.

References

[1] C.W. Murray, D.C. Rees, The rise of fragment-based drug discovery, Nat. Chem. 1 (3) (2009) 187–192.
[2] M.M. Hann, A.R. Leach, G. Harper, Molecular complexity and its impact on the probability of finding leads for drug discovery, J. Chem. Inf. Comput. Sci. 41 (3) (2001) 856–864.
[3] S.H. Gellman, Introduction: molecular recognition, Chem. Rev. 97 (5) (1997) 1231–1232.
[4] K.K. Frederick et al., Conformational entropy in molecular recognition by proteins, Nature 448 (7151) (2007) 325–329.
[5] D.D. Boehr, R. Nussinov, P.E. Wright, The role of dynamic conformational ensembles in biomolecular recognition, Nat. Chem. Biol. 5 (11) (2009) 789–796.
[6] E. Petsalaki et al., Accurate prediction of peptide binding sites on protein surfaces, PLoS Comput. Biol. 5 (3) (2009) e1000335.
[7] P. Vanhee et al., PepX: a structural database of non-redundant protein–peptide complexes, Nucleic Acids Res. 38 (Database issue) (2010) D545–D551.
[8] T.R. Kumar, K. Soppimath, S.K. Nachaegari, Novel delivery technologies for protein and peptide therapeutics, Curr. Pharm. Biotechnol. 7 (4) (2006) 261–276.
[9] S. Lien, H.B. Lowman, Therapeutic peptides, Trends Biotechnol. 21 (12) (2003) 556–562.
[10] J. An, M. Totrov, R.A. Abagyan, Pocketome via comprehensive identification and classification of ligand binding envelopes, Mol. Cell. Proteomics 4 (6) (2005) 752–761.
[11] G. Schneider, U. Fechner, Computer-based de novo design of drug-like molecules, Nat. Rev. Drug Discovery 4 (8) (2005) 649–663.
[12] T. Honma, Recent advances in de novo design strategy for practical lead identification, Med. Res. Rev. 23 (5) (2003) 606–632.
[13] J.B. Moon, W.J. Howe, Computer design of bioactive molecules: a method for receptor-based de novo ligand design, Proteins 11 (4) (1991) 314–328.
[14] J. Greer et al., Application of the three-dimensional structures of protein target molecules in structure-based drug design, J. Med. Chem. 37 (8) (1994) 1035–1054.
[15] R.E. Babine, S.L. Bender, Molecular recognition of proteinminus signLigand complexes: applications to drug design, Chem. Rev. 97 (5) (1997) 1359–1472.
[16] I.D. Kuntz, Structure-based strategies for drug design and discovery, Science 257 (5073) (1992) 1078–1082.
[17] N.M. Glykos, G. Cesareni, M. Kokkinidis, Protein plasticity to the extreme: changing the topology of a 4-alpha-helical bundle with a single amino acid substitution, Structure 7 (6) (1999) 597–603.
[18] M. Huse, J. Kuriyan, The conformational plasticity of protein kinases, Cell 109 (3) (2002) 275–282.
[19] A. Morton, B.W. Matthews, Specificity of ligand binding in a buried nonpolar cavity of T4 lysozyme: linkage of dynamics and structural plasticity, Biochemistry 34 (27) (1995) 8576–8588.

[20] L. Birch et al., Sensitivity of molecular docking to induced fit effects in influenza virus neuraminidase, J. Comput. Aided Mol. Des. 16 (12) (2002) 855–869.
[21] S.B. Nabuurs, M. Wagener, J. de Vlieg, A flexible approach to induced fit docking, J. Med. Chem. 50 (26) (2007) 6507–6518.
[22] W. Sherman et al., Novel procedure for modeling ligand/receptor induced fit effects, J. Med. Chem. 49 (2) (2006) 534–553.
[23] B. Sandak, H.J. Wolfson, R. Nussinov, Flexible docking allowing induced fit in proteins: insights from an open to closed conformational isomers, Proteins 32 (2) (1998) 159–174.
[24] R. Najmanovich et al., Side-chain flexibility in proteins upon ligand binding, Proteins 39 (3) (2000) 261–268.
[25] S.J. Teague, Implications of protein flexibility for drug discovery, Nat. Rev. Drug Discovery 2 (7) (2003) 527–541.
[26] M.E. Fraser et al., Crystallographic analysis of transition-state mimics bound to penicillopepsin: phosphorus-containing peptide analogues, Biochemistry 31 (22) (1992) 5201–5214.
[27] H.J. Hecht et al., Three-dimensional structure of a recombinant variant of human pancreatic secretory trypsin inhibitor (Kazal type), J. Mol. Biol. 225 (4) (1992) 1095–1103.
[28] A.M. Lesk, C. Chothia, Elbow motion in the immunoglobulins involves a molecular ball-and-socket joint, Nature 335 (6186) (1988) 188–190.
[29] I. Kufareva, R. Abagyan, Type-II kinase inhibitor docking, screening, and profiling using modified structures of active kinase states, J. Med. Chem. 51 (24) (2008) 7921–7932.
[30] R. Abagyan, I. Kufareva, The flexible pocketome engine for structural chemogenomics, Methods Mol. Biol. 575 (2009) 249–279.
[31] C. de Graaf, D. Rognan, Selective structure-based virtual screening for full and partial agonists of the beta2 adrenergic receptor, J. Med. Chem. 51 (16) (2008) 4978–4985.
[32] C. de Graaf, D. Rognan, Customizing G protein-coupled receptor models for structure-based virtual screening, Curr. Pharm. Des. 15 (35) (2009) 4026–4048.
[33] J.J. Irwin et al., Automated docking screens: a feasibility study, J. Med. Chem. 52 (18) (2009) 5712–5720.
[34] V. Katritch et al., Analysis of full and partial agonists binding to beta2-adrenergic receptor suggests a role of transmembrane helix V in agonist-specific conformational changes, J. Mol. Recognit. 22 (4) (2009) 307–318.
[35] R.L. Dunbrack Jr., M. Karplus, Conformational analysis of the backbone-dependent rotamer preferences of protein sidechains, Nat. Struct. Biol. 1 (5) (1994) 334–340.
[36] P. Kallblad, P.M. Dean, Efficient conformational sampling of local side-chain flexibility, J. Mol. Biol. 326 (5) (2003) 1651–1665.
[37] R.M. Knegtel, I.D. Kuntz, C.M. Oshiro, Molecular docking to ensembles of protein structures, J. Mol. Biol. 266 (2) (1997) 424–440.
[38] D.S. Goodsell, G.M. Morris, A.J. Olson, Automated docking of flexible ligands: applications of AutoDock, J. Mol. Recognit. 9 (1) (1996) 1–5.
[39] M.J. Betts, M.J. Sternberg, An analysis of conformational changes on protein–protein association: implications for predictive docking, Protein Eng. 12 (4) (1999) 271–283.
[40] J.W. Nissink et al., A new test set for validating predictions of protein–ligand interaction, Proteins 49 (4) (2002) 457–471.
[41] H. Claussen et al., FlexE: efficient molecular docking considering protein structure variations, J. Mol. Biol. 308 (2) (2001) 377–395.
[42] G. Jones et al., Development and validation of a genetic algorithm for flexible docking, J. Mol. Biol. 267 (3) (1997) 727–748.
[43] I.D. Kuntz et al., A geometric approach to macromolecule–ligand interactions, J. Mol. Biol. 161 (2) (1982) 269–288.
[44] M. Totrov, R. Abagyan, Flexible ligand docking to multiple receptor conformations: a practical alternative, Curr. Opin. Struct. Biol. 18 (2) (2008) 178–184.
[45] J.A. Erickson et al., Lessons in molecular recognition: the effects of ligand and protein flexibility on molecular docking accuracy, J. Med. Chem. 47 (1) (2004) 45–55.
[46] H.A. Carlson, Protein flexibility and drug design: how to hit a moving target, Curr. Opin. Chem. Biol. 6 (4) (2002) 447–452.
[47] H.A. Carlson, Protein flexibility is an important component of structure-based drug discovery, Curr. Pharm. Des. 8 (17) (2002) 1571–1578.
[48] C.N. Cavasotto, R.A. Abagyan, Protein flexibility in ligand docking and virtual screening to protein kinases, J. Mol. Biol. 337 (1) (2004) 209–225.
[49] H. Gohlke, L.A. Kuhn, D.A. Case, Change in protein flexibility upon complex formation: analysis of Ras–Raf using molecular dynamics and a molecular framework approach, Proteins 56 (2) (2004) 322–337.
[50] M.I. Zavodszky, L.A. Kuhn, Side-chain flexibility in protein–ligand binding: the minimal rotation hypothesis, Protein Sci. 14 (4) (2005) 1104–1114.
[51] M.J. Bower, F.E. Cohen, R.L. Dunbrack Jr., Prediction of protein side-chain rotamers from a backbone-dependent rotamer library: a new homology modeling tool, J. Mol. Biol. 267 (5) (1997) 1268–1282.
[52] A.M. Ferrari et al., Soft docking and multiple receptor conformations in virtual screening, J. Med. Chem. 47 (21) (2004) 5076–5084.
[53] R.L. Dunbrack Jr., M. Karplus, Backbone-dependent rotamer library for proteins. Application to side-chain prediction, J. Mol. Biol. 230 (2) (1993) 543–574.
[54] O.V. Kalinina et al., Combinations of protein–chemical complex structures reveal new targets for established drugs, PLoS Comput. Biol. 7 (5) (2011) e1002043.
[55] M.L. Benson et al., Binding MOAD, a high-quality protein–ligand database, Nucleic Acids Res. 36 (Database issue) (2008) D674–D678.
[56] B.A. Shoemaker et al., Inferred biomolecular interaction server – a web server to analyze and predict protein interacting partners and binding sites, Nucleic Acids Res. 38 (Database issue) (2010) D518–D524.
[57] J. Gunther et al., Utilising structural knowledge in drug design strategies: applications using Relibase, J. Mol. Biol. 326 (2) (2003) 621–636.
[58] E.I. Juritz, S.F. Alberti, G.D. Parisi, PCDB: a database of protein conformational diversity, Nucleic Acids Res. 39 (Database issue) (2011) D475–D479.
[59] J. Ponomarenko et al., IEDB-3D: structural data within the immune epitope database, Nucleic Acids Res. 39 (Database issue) (2011) D1164–D1170.
[60] I. Kufareva, A.V. Ilatovskiy, R. Abagyan, Pocketome: an encyclopedia of small-molecule binding sites in 4D, Nucleic Acids Res. 40 (Database issue) (2012) D535–D540.
[61] H.M. Berman et al., The protein data bank, Nucleic Acids Res. 28 (2000) 235–242.
[62] G. Bottegoni et al., Four-dimensional docking: a fast and accurate account of discrete receptor flexibility in ligand docking, J. Med. Chem. 52 (2) (2009) 397–406.
[63] A. Ceol et al., DOMINO: a database of domain–peptide interactions, Nucleic Acids Res. 35 (Database issue) (2007) D557–D560.
[64] K.R. Govindarajan et al., MPID: MHC–peptide interaction database for sequence-structure-function information on peptides binding to MHC molecules, Bioinformatics 19 (2) (2003) 309–310.
[65] B.A. Shoemaker et al., Inferred biomolecular interaction server: a web server to analyze and predict protein interacting partners and binding sites, Nucleic Acids Res. 38 (Suppl. 1) (2010) D518–D524.
[66] G. Bottegoni et al., Systematic exploitation of multiple receptor conformations for virtual ligand screening, PLoS One 6 (5) (2011) e18845.
[67] J. Carlsson et al., Structure-based discovery of A2A adenosine receptor ligands, J. Med. Chem. 53 (9) (2010) 3748–3755.
[68] V. Katritch et al., Structure-based discovery of novel chemotypes for adenosine A2A receptor antagonists, J. Med. Chem. 53 (4) (2010) 1799–1809.
[69] D.S. Wishart et al., DrugBank: a knowledgebase for drugs, drug actions and drug targets, Nucleic Acids Res. 36 (Database issue) (2008) D901–D906.
[70] G.P.S. Raghava, BioTherapi: Bioinformatics for Therapeutic Peptides and Proteins. Available from: <http://www.imtech.res.in/raghava/biotherapi/pdrug.html>.
[71] E. Raush et al., A new method for publishing three-dimensional content, PLoS One 4 (10) (2009).
[72] W.H. Lee et al., Interactive JIMD articles using the iSee concept: turning a new page on structural biology data, J. Inherit. Metab. Dis. 34 (3) (2011) 565–567.
[73] R. Abagyan et al., Disseminating structural genomics data to the public: from a data dump to an animated story, Trends Biochem. Sci. 31 (2) (2006) 76–78.