

PIC: Protein Interactions Calculator

K. G. Tina, R. Bhadra and N. Srinivasan*

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

Received January 31, 2007; Revised April 17, 2007; Accepted May 9, 2007

ABSTRACT

Interactions within a protein structure and interactions between proteins in an assembly are essential considerations in understanding molecular basis of stability and functions of proteins and their complexes. There are several weak and strong interactions that render stability to a protein structure or an assembly. Protein Interactions Calculator (PIC) is a server which, given the coordinate set of 3D structure of a protein or an assembly, computes various interactions such as disulphide bonds, interactions between hydrophobic residues, ionic interactions, hydrogen bonds, aromatic–aromatic interactions, aromatic–sulphur interactions and cation– π interactions within a protein or between proteins in a complex. Interactions are calculated on the basis of standard, published criteria. The identified interactions between residues can be visualized using a RasMol and Jmol interface. The advantage with PIC server is the easy availability of inter-residue interaction calculations in a single site. It also determines the accessible surface area and residue-depth, which is the distance of a residue from the surface of the protein. User can also recognize specific kind of interactions, such as apolar–apolar residue interactions or ionic interactions, that are formed between buried or exposed residues or near the surface or deep inside.

INTRODUCTION

Analyses of atomic interactions in tertiary structures of proteins contribute richly to our understanding of sequence–structure relationships, structural basis of protein stability and protein evolution. Studies on interactions between sidechains are commonly used in designing methods and identifying strategies for remote homology detection, protein fold recognition, protein structural comparisons and comparative protein modelling (1–6). For example, remote homology detection between proteins can rely on conservation of structural motifs involving interacting sidechains (1). Such studies also

serve as guidelines in designing site-directed mutagenesis experiments (7) and in the understanding of the basis for residue conservation in homologous proteins (8).

Interactions between subunits of multimeric proteins and interactions between interacting protein modules are also areas of intense study (9–13). Analyses on nature of sidechain–sidechain interactions across interacting interfaces between protein modules have enlightened us on evolutionary conservation of protein–protein interactions and in distinguishing transient and permanent complexes (14–16).

Different kinds of interactions have been noted in the stabilization of tertiary structures, quaternary structures and assemblies of proteins. Roles and importance of interactions between apolar residues and hydrogen bonds are very well known (2). Importance of interactions such as aromatic–aromatic, aromatic–sulphur, cation– π and ionic interactions in the structure and function of proteins is also well realized (17–20). Observations and analyses of less common features such as exposed cluster of hydrophobic residues, partially buried salt bridges and interactions of buried charged residues are also of specific interest (21–24).

Here we report the development of a web-based service, PIC (Protein Interactions Calculator), to aid recognition and analyses of various kinds of interactions in tertiary structures of proteins and structures of protein–protein complexes. We have also integrated solvent accessibility calculations in PIC to aid recognition of interacting motifs that are exposed or buried. Further, the residue depth calculations are also made possible in PIC so that interactions deep inside the protein structure or near the surface can be recognized. Advantage of using residue depth parameter is that it can distinguish residues with no solvent accessible surface area in terms of how deep they are from the protein surface.

MATERIALS AND METHODS

Organization of PIC

PIC server accepts atomic coordinate set of a protein structure in the standard Protein Data Bank (PDB) format. The user is prompted with selecting one or more of the following interaction types: Interaction between apolar residues, disulphide bridges, hydrogen bond

*To whom correspondence should be addressed. Tel: +91 80 2293 2837; Fax: +91 80 2360 0535; Email: ns@mbu.iisc.ernet.in

between main chain atoms, hydrogen bond between main chain and sidechain atoms, hydrogen bond between two sidechain atoms, interaction between oppositely charged amino acids (ionic interactions), aromatic–aromatic interactions, aromatic–sulphur interactions and cation– π interactions. The input coordinate set is accepted, under each section of the page, for recognition of interactions within a polypeptide chain. If an ensemble of NMR-derived structures is input then the first model in the file is taken as a representative and is used by the PIC server. The output corresponds to the list of residues involved in interaction type of interest. An option is provided, using RasMol (25) interface and Jmol interface, for enabling visualization of structure in the graphics with interactions highlighted. It is possible to get the results by e-mail. It is also possible to download the output files of the original programs.

A separate panel is available for identification of various types of interactions between polypeptide chains when a multi-chain PDB file is subjected to the analysis. All the said interactions could be explored for their occurrence across the inter-polypeptide chain interface. Thus this panel facilitates recognition of interactions between different subunits in a multimeric protein structures or between proteins in a protein–protein complex structure. Figure 1 shows ionic interactions between oppositely charged sidechains across the interface, formed between cyclin-dependent protein kinase and bound cyclin (26), recognized using PIC server.

Solvent accessibility calculations could be used to identify different kinds of interactions between buried or between solvent exposed residues. Solvent accessibility calculations are performed using NACCESS program (Hubbard, S.J. and Thornton, J.M., 1993, NACCESS

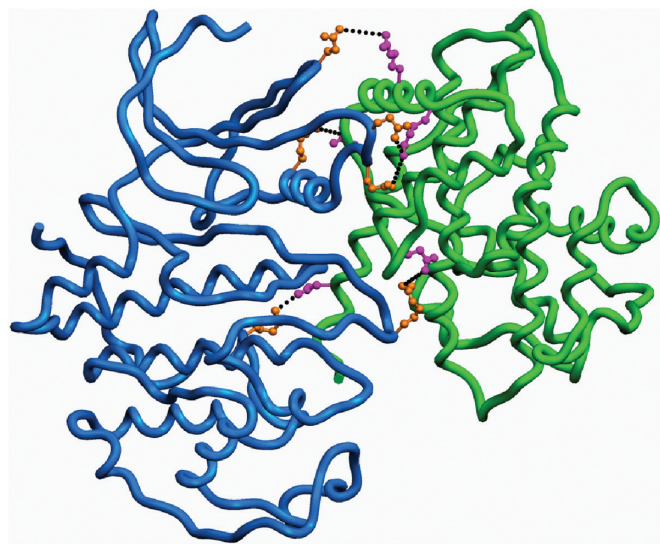


Figure 1. Interactions between oppositely charged amino acid side-chains in the interaction interface of cyclic dependent protein kinase 2 (CDKs) and cyclin identified using PIC server. The folds of CDK2 and cyclin and the charged residues in the interface formed by the two proteins are represented in different colours. The ion pairs are highlighted by black dotted lines. This figure is produced using SETOR (35).

Computer Program, Department of Biochemistry and Molecular Biology, University College London.). The exposed and buried residues are identified by $>7\%$ and $\leq 7\%$ residue accessibility, respectively. Under this facility list of all the interaction types are displayed prompting the user to select list of interaction types of interest. For example, a user may prefer to identify interactions between apolar residues that are exposed. Figure 2 shows interactions between solvent exposed apolar residues, in crambin (27), recognized using PIC server.

Depth of an atom in a protein is defined as the distance from the nearest atom in the surface of the protein structure. Mean depths of atoms of a residue defines the residue depth (28,29). Analogous to the panel on solvent accessibility, panel on residue depth enables the users to identify specific types of interactions near the protein surface or deep inside the core of the structure. Based on the analysis of residue depth parameter by Chakravarty and Varadarajan (28) we consider those residues with depths $\leq 5 \text{ \AA}$ as close to the protein surface and others as deep inside. Using this part of the PIC server it is possible to identify interactions between, say, aromatic residues near the protein structural surface. As calculation of residue depths takes a few minutes for most protein structures, results involving depth calculation are sent by e-mail to the user if a valid e-mail address is provided.

Recognition of interactions

Various types of interactions are recognized from the atomic coordinates using the standard criteria that are published. We used mainly the criteria suggested by NCI server (30) to identify non-canonical interactions in

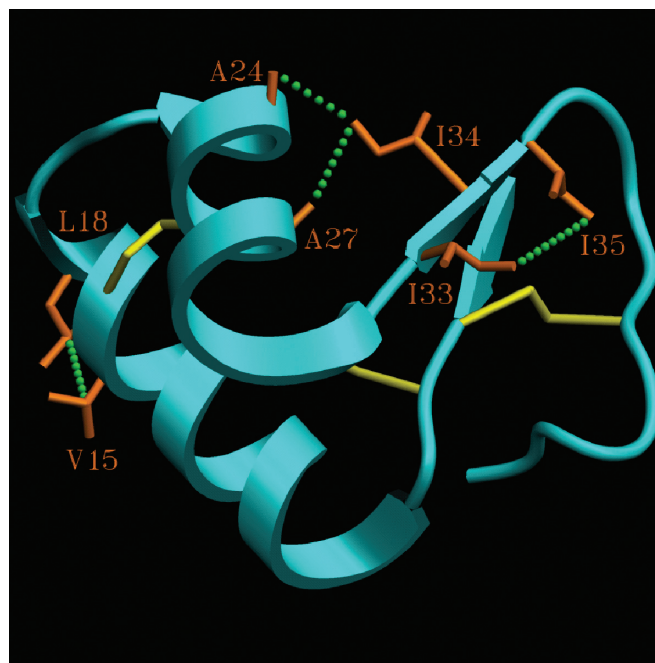


Figure 2. Structure of crambin with solvent exposed and interacting apolar sidechains, recognized using PIC. Interactions between apolar sidechains is shown by green dots. Disulphide bonds are shown in yellow. This figure is produced using SETOR (35).

proteins. The aromatic–aromatic, aromatic–sulphur and cation– π interactions are recognized between appropriate sidechains using the criteria proposed by Burley and Petsko (17), Reid *et al.* (18) and Satyapriya and Vishveshwara (19), respectively. Disulphide bonds are recognized using the distance criteria employed originally in the MODIP program (31). Hydrogen bonds are recognized using HBOND routine developed by Overington *et al.* (32) and described in Mizuguchi *et al.* (33). The hydrogen bonds are categorized as main chain–main chain, main chain–sidechain and sidechain–sidechain. Only standard hydrogen bonds are recognized in PIC as NCI server (30) is available for identification of interactions such as C–H...O. Interactions between hydrophobic sidechains are identified using a distance cut-off of 5 Å between apolar groups in the apolar sidechains. Though various interactions are recognized using the standard criteria, user has an option of changing the distance cut-off in recognizing any of the types of interactions.

CONCLUSIONS

Analysing the interactions that stabilize tertiary and quaternary structures and protein–protein complexes is a common situation in structural biology. For example, a structure just solved using X-ray analysis or nuclear magnetic resonance may be subjected to such an analysis. In protein engineering and design experiments, a good understanding of the structural roles of various residues is essential before taking decisions on residues to mutate by site-directed mutagenesis and the replacing residue. Use of combinations of features available in PIC such as various kinds of interactions and solvent exposure or buried nature or depths of residues is also expected to aid recognition of common and uncommon structural features in a given protein structure. Analysis of such interactions in homologous protein structures (34) enables recognition of evolutionary constraints critical for the retention of fold of the protein family.

The PIC server is available at: <http://crick.mbu.iisc.ernet.in/~PIC>

ACKNOWLEDGEMENTS

This research is supported by the Department of Biotechnology (DBT), Government of India. K.G.T. and R.B. are supported by grants to N.S. by DBT. Funding to pay the Open Access publication charges for this article was provided by Oxford University Press.

Conflict of interest statement. None declared.

REFERENCES

- Bhaduri, A., Ravishankar, R. and Sowdhamini, R. (2004) Conserved spatially interacting motifs of protein superfamilies: application to fold recognition and function annotation of genome data. *Proteins*, **54**, 657–670.
- Gromiha, M.M. and Selvaraj, S. (2004) Inter-residue interactions in protein folding and stability. *Prog. Biophys. Mol. Biol.*, **86**, 235–277.
- Shao, X. and Grishin, N.V. (2000) Common fold in helix-hairpin-helix proteins. *Nucl. Acids Res.*, **28**, 2643–2650.
- Reva, B.A., Finkelstein, A.V., Sanner, M., Olson, A.J. and Skolnick, J. (1997) Recognition of protein structure on coarse lattices with residue-residue energy functions. *Prot. Eng.*, **10**, 1123–1130.
- Russell, R.B. and Barton, G.J. (1994) Structural features can be unconserved in proteins with similar folds. An analysis of side-chain to side-chain contacts secondary structure and accessibility. *J. Mol. Biol.*, **244**, 332–350.
- Sali, A. and Blundell, T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779–815.
- Van Roey, P., Pereira, B., Li, Z., Hiraga, K., Belfort, M. and Derbyshire, V. (2006) Crystallographic and mutational studies of mycobacterium tuberculosis recA mini-inteins suggest a pivotal role for a highly conserved aspartate residue. *J. Mol. Biol.* [Epub ahead of print].
- Krupa, A., Preethi, G. and Srinivasan, N. (2004) Structural modes of stabilization of permissive phosphorylation sites in protein kinases: distinct strategies in Ser/Thr and Tyr kinases. *J. Mol. Biol.*, **339**, 1025–1039.
- Saha, R.P., Bhattacharyya, R. and Chakrabarti, P. (2007) Interaction geometry involving planar groups in protein-protein interfaces. *Proteins*. [Epub ahead of print].
- Nooren, I.M. and Thornton, J.M. (2003) Diversity of protein-protein interactions. *EMBO J.*, **22**, 3486–3492.
- Jones, S. and Thornton, J.M. (1995) Protein-protein interactions: a review of protein dimer structures. *Prog. Biophys. Mol. Biol.*, **63**, 31–65.
- Ofran, Y. and Rost, B. (2003) Analysing six types of protein-protein interfaces. *J. Mol. Biol.*, **325**, 377–387.
- Rekha, N., Machado, S.M., Narayanan, C., Krupa, A. and Srinivasan, N. (2005) Interaction interfaces of protein domains are not topologically equivalent across families within superfamilies: Implications for metabolic and signaling pathways. *Proteins*, **58**, 339–353.
- Bahadur, R.P., Chakrabarti, P., Rodier, F. and Janin, J. (2004) A dissection of specific and non-specific protein-protein interfaces. *J. Mol. Biol.*, **336**, 943–955.
- De, S., Krishnadev, O., Srinivasan, N. and Rekha, N. (2005) Interaction preferences across protein-protein interfaces of obligatory and non-obligatory components are different. *BMC Struct. Biol.*, **5**, 15.
- Valdar, W.S. and Thornton, J.M. (2001) Conservation helps to identify biologically relevant crystal contacts. *J. Mol. Biol.*, **313**, 399–416.
- Burley, S.K. and Petsko, G.A. (1985) Aromatic-aromatic interaction: a mechanism of protein structure stabilization. *Science*, **229**, 23–28.
- Reid, K.S.C., Lindley, P.F. and Thornton, J.M. (1985) Sulphur-aromatic interactions in proteins. *FEBS Lett.*, **190**, 209–213.
- Satyapriya, R. and Vishveshwara, S. (2004) Interaction of DNA with clusters of amino acids in proteins. *Nucleic Acids Res.*, **32**, 4109–4118.
- Barlow, D.J. and Thornton, J.M. (1983) Ion-pairs in proteins. *J. Mol. Biol.*, **168**, 867–885.
- Ahn, H.C., Juranic, N., Macura, S. and Markley, J.L. (2006) Three-dimensional structure of the water-insoluble protein crambin in dodecylphosphocholine micelles and its minimal solvent-exposed surface. *J. Am. Chem. Soc.*, **128**, 4398–4404.
- Anderson, T.A., Cordes, M.H. and Sauer, R.T. (2005) Sequence determinants of a conformational switch in a protein structure. *Proc. Natl Acad. Sci. USA*, **102**, 18344–18349.
- Dong, F. and Zhou, H.X. (2002) Electrostatic contributions to T4 lysozyme stability: solvent-exposed charges versus semi-buried salt bridges. *Biophys. J.*, **83**, 1341–1347.
- Balaji, S., Aruna, S. and Srinivasan, N. (2003) Tolerance to the substitution of buried apolar residues by charged residues in the homologous protein structures. *Proteins*, **53**, 783–791.
- Sayle, R.A. and Milner-White, E.J. (1995) RasMol: Biomolecular graphics for all. *Trends Biochem. Sci.*, **20**, 374–376.
- Brown, N.R., Noble, M.E., Endicott, J.A. and Johnson, L.N. (1999) The Structural Basis for Specificity of Substrate and Recruitment Peptides for Cyclin-Dependent Kinases. *Nat. Cell. Biol.*, **1**, 438–443.

27. Teeter, M.M. (1984) Water Structure of a Hydrophobic Protein at Atomic Resolution. Pentagon Rings of Water Molecules in Crystals of Crambin. *Proc. Natl Acad. Sci. USA*, **81**, 6014–6018.
28. Chakravarty, S. and Varadarajan, R. (1999) Residue depth: a novel parameter for the analysis of protein structure and stability. *Structure*, **7**, 723–732.
29. Pintar, A., Carugo, O. and Pongor, S. (2003) Atom depth as a descriptor of the protein interior. *Biophys. J.*, **84**, 2553–2561.
30. Babu, M.M. (2003) NCI: A server to identify non-canonical interactions in protein structures. *Nucl. Acids Res.*, **31**, 3345–3348.
31. Sowdhamini, R., Srinivasan, N., Shoichet, B., Santi, D.V., Ramakrishnan, C. and Balaram, P. (1989) Stereochemical modelling of disulfide bridges: Criteria for introduction into proteins by site-directed mutagenesis. *Prot. Eng.*, **3**, 95–103.
32. Overington, J.P., Johnson, M.S., Sali, A. and Blundell, T.L. (1990) Tertiary structural constraints on protein evolutionary diversity: templates, key residues and structure prediction. *Proc. Roy. Soc. Biol. Sci.*, **241**, 132–145.
33. Mizuguchi, K., Deane, C.M., Blundell, T.L., Johnson, M.S. and Overington, J.P. (1998) JOY: protein sequence-structure representation and analysis. *Bioinformatics*, **14**, 617–623.
34. Gowri, V.S., Pandit, S.B., Karthik, P.S., Srinivasan, N. and Balaji, S. (2003) Integration of related sequences with protein three-dimensional structural families in an updated version of PALI database. *Nucleic Acids Res.*, **31**, 486–488.
35. Evans, S.V. (1993) SETOR: hardware lighted three-dimensional solid model representations of macromolecules. *J. Mol. Graphics*, **11**, 134–138.