J Mol Model (2011) 17:599–606
DOI 10.1007/s00894-010-0758-9

ORIGINAL PAPER

dockYard–a repository to assist modeling of protein-protein docking

Pralay Mitra · Debnath Pal

Received: 13 April 2010 / Accepted: 12 May 2010 / Published online: 4 June 2010

Abstract In the absence of interlogs, building docking models is a time intensive task, involving generation of a large pool of docking decoys followed by refinement and screening to identify near native docking solutions. This limits the researcher interested in building docking methods with the choice of benchmarking only a limited number of protein complexes. We have created a repository called *dockYard* (http://pallab.serc.iisc.ernet.in/dockYard), that allows modelers interested in protein-protein interaction to access large volume of information on protein dimers and their interlogs, and also download decoys for their work if they are interested in building modeling methods. *dockYard* currently offers four categories of docking decoys derived from: Bound (native dimer co-crystallized), Unbound (individual subunits are crystallized, as well as the target dimer), Variants (match the previous two categories in at least one subunit with 100% sequence identity), and Interlogs (match the previous categories in at least one subunit with ≥90% or ≥50% sequence identity). The web service offers options for full or selective download based on search parameters. Our portal also serves as a repository to modelers who may want to share their decoy sets with the community.

Keywords Decoy · Docking · Interaction · Interlogs · Protein-protein · Repository

P. Mitra · D. Pal (✉)
Bioinformatics Centre, Supercomputer Education Research Centre, Indian Institute of Science,
Bangalore 560 012, India
e-mail: dpal@serc.iisc.ernet.in

Introduction

Protein-protein docking models provide a mechanistic description to interactions between protein molecules in the cell. It is therefore, an important research area that has drawn a wide attention around the world. The forum for the Critical Assessment of Prediction of Interactions (CAPRI) [1], tries to bring together docking modelers in an effort to improve docking models for novel targets. Research in modeling of docking faces inherent difficulty when there is no information on interlogs because a large number of docking decoys need to be generated followed by structural refinement and screening. The average time taken by an Intel(R) Pentium(R) 4 CPU 3.40 GHz Desktop computer to generate docking decoys for a 50 residue homodimeric protein is around two hours, while for large proteins it usually takes days. At a higher resolution search using a smaller rotation and translation step, the time taken is much higher. Also, the time taken to generate the decoys does not scale uniformly with the size of the protein subunits, but is influenced by its shape and surface properties as well. In addition to these if flexibility of side- and main-chain is incorporated into the decoy search, the time requirement is increased for at least additional $(n \log n)^3$ operations, here n is the number of atoms in the system. Under these circumstances, a modeler is restricted to benchmark his/her method on only a handful of test protein complexes (Table 1), unless he/she has an annotated data set and access to supercomputers or large compute clusters. There has been some effort in providing standard test sets to assist modeling method builders, as in Benchmark 3.0 [2] or DOCKGROUND [3, 4], the latter providing only a limited set of decoys for evaluation but not sufficient to create a large test/training set. We offer modelers interested in developing docking model screening methods, but not

Table 1 List of docking algorithms and the size of test sets used for benchmarking

Title / Website	Test set size
3D-Garden: http://www.sbg.bio.ic.ac.uk/3dgarden [18]	84
SOFTDOCK: http://bio.iphy.ac.cn [19].	83
Clustering protein-protein docking predictions [20]	36
ZRANK [21]	62
VorScorE [11]	102
Incorporating intermolecular distance into protein-protein docking [22].	41
Hydrophobic complementarity in protein-protein docking [23].	39
Protein-protein docking with a reduced protein model accounting for side-chain flexibility [24].	18
Zdock [25].	49
A fast empirical approach to binding free energy calculations based on protein interface information [26].	10
Electrostatic contributions to protein-protein interactions: Fast energetic filters for docking and their physical basis [27].	29
Docking unbound proteins using shape complementarity, desolvation, and electrostatics [28].	27
BiGGER: a new (soft) docking algorithm for predicting protein interactions [29].	25
Estimation and filtering of potential protein-protein docking positions [30].	51
PUZZLE: a new method for automated protein docking based on surface shape complementarity [31].	8

willing to devote a lot of time in generating docking decoys, a portal to directly download docking decoys and save significant computing time. Our portal called *dockYard,* is publicly accessible at http://pallab.serc.iisc.ernet.in/dockYard. It also allows modelers who spend time generating their own docking decoys to submit the same in our repository to share them with the community. General docking modelers interested only in modeling of protein-protein docking will find our portal useful for information on protein dimers and their interlogs.

General scheme

The main aim of the database is to offer a researcher an opportunity to obtain an extensive set of docking decoys to assist his/her work. As can be seen from the schematic diagram (Fig. 1), there are four options to query on different categories of molecules (i) single protein, (ii) a class of proteins, (iii) variants of a protein and (iv) models. Accordingly, four categories of information on decoys and models for dimeric protein complexes are filtered and sets of coordinates and transformation matrices satisfying search criteria are available for download from the *dockYard*. The categories are: (i) Docking decoys generated from co-crystallized (bound) protein structures, (ii) Docking decoys generated from unbound protein structures (*i.e.,* protein complexes for which structures of individual subunits are separately crystallized/available), (iii) Variants: structures with 100% sequence identity with the previous two categories, but different in the side-and/or main-chain conformation, and (iv) Interlogs: protein-complex models generated based on homology considerations.

Category (i)

In this category, we have about 9,020,000 docking decoys generated from 902 X-ray crystallography derived co-crystal structures of which 796 are homodimeric, and 106 are heterodimeric quaternary structures. The complexes are filtered using the resolution and R-factor of 2.5 Å and 0.2, respectively. The subunit sizes are restricted to at least 25 residues and with no major errors like chain breaks, specifically at the protein surface. Protein complexes with ligands exceeding the size of 5 heavy atoms at the interface are excluded. The filtered set is nonredundant by selecting only one representative so that no two subunits share >90% sequence identity at the interface. The data set covers a total of 451 structural classification of proteins (SCOP) [5] fold classes for homodimers and 89 SCOP fold classes for heterodimers. Information from the literature, and protein quaternary structure (PQS) database [6] is used to check if the dimers selected is biological—82% of the data on the quaternary structure is confirmed by the PISA server [7]. We did not categorize the data into obligate, non-obligate, transient, permanent and so on since there does not exist sufficient physiological data on the majority of the complexes. As a rule one can use the fact that most homodimers are obligate, while heterodimers are not.

Information on ten thousand docking decoys generated using FTDOCK [8] for each protein complex are stored in the database (for details see section “Data generation” below). We categorize these decoys based on the L_{RMSD} measure, in which the root-mean-square-deviation (RMSD) between the C_α atoms of the native complex and the decoy is calculated after superposing the larger subunits of the two

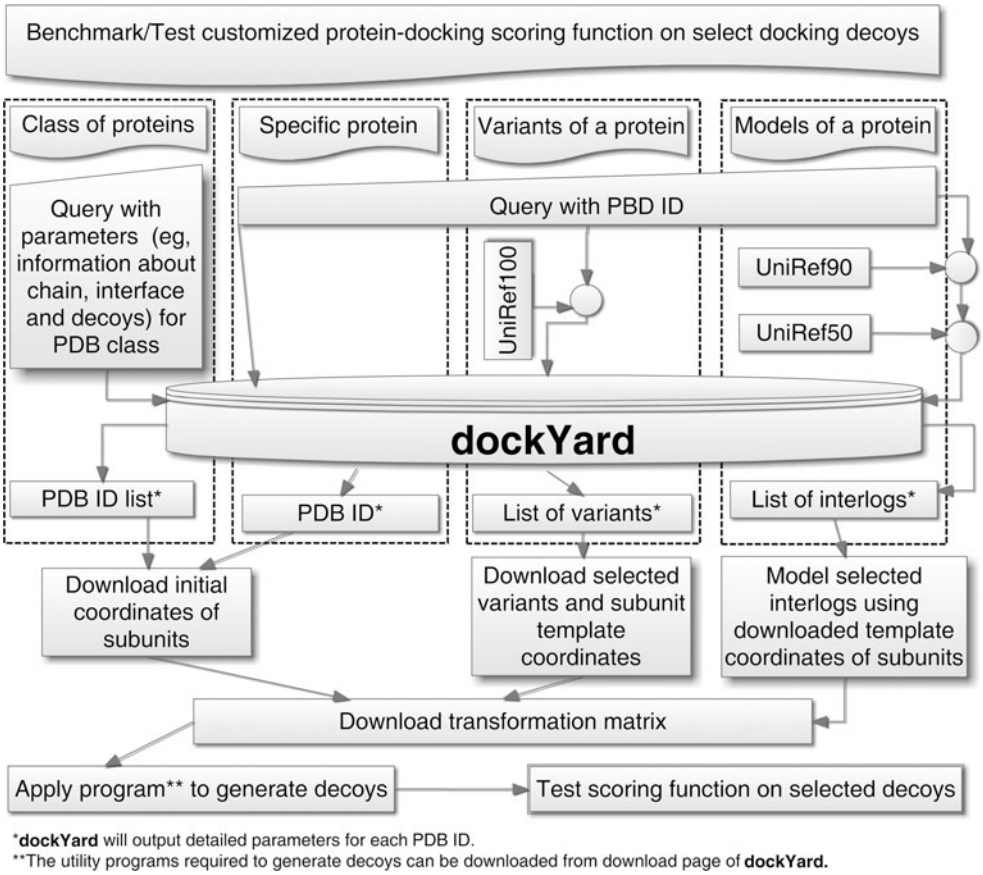


Fig. 1 Schematic diagram showing the general flow of tasks during the various docking decoy searches available from the web interface

quaternary structures. The minimum L_{RMSD} in the set is 0.7 Å and the maximum 144.1 Å. Of the 902 dimers, all have at least one decoy within L_{RMSD} <5.0 Å (Fig. 2). 400 dimers have at least one decoy within L_{RMSD} 2.0 Å. 728 and 864 dimers are there with at least one decoy within L_{RMSD} 3.0 Å and 4.0 Å, respectively. Detailed L_{RMSD} information for each protein complex can also be obtained from the web server (see Fig. 3b). Using the definition based on I_{RMSD}, one can also categorize the decoys [2]. I_{RMSD} is same as L_{RMSD} except that only the interface C_α coordinates are used for RMSD calculation. I_{RMSD}<1.5 Å between dimer and its decoys is termed as a good match, medium matches have 1.5 Å<I_{RMSD}<2.2 Å, and the poor matches > 2.2 Å. For 36% of dimers in the data set, there exists at least one decoy with I_{RMSD}<1.5 Å, 47% cases in 1.5 Å<I_{RMSD}<2.2 Å range and 17% cases with I_{RMSD}>2.2 Å.

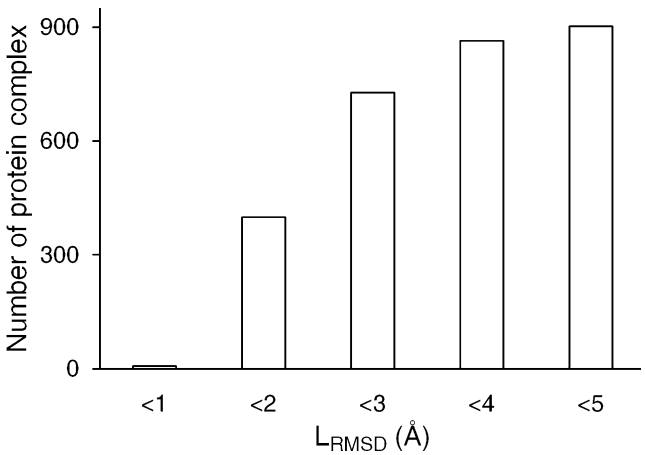


Fig. 2 Plot showing the distribution of minimum L_{RMSD} in the decoy set obtained for each protein complex

Category (ii)

In the unbound category, there is a total of 40 dimeric complexes. We culled unique dimers from Benchmark 2.0 [9], Benchmark 3.0 [2], Gottschalk et al. [10] and Bernaur et al. [11]. To confirm the oligomeric state, we used information from both PQS [6] and PISA [7] servers. As with the decoys in category (i), the side chains / main-chain of the decoys are not refined, but only rigid body docking applied using the FTDOCK search as outlined above. In 5 out of the 40 cases there is at least one decoy with $L_{RMSD}\leq$ 2.0 Å., 23 cases ≤5 Å and all of them ≤10 Å. Additional unbound cases in the database can be accessed through query on models and variants. For an example see the section “An example run”.

a

Home Search Download Help Reference Contact

Result of your query

List of filtered PDB

PDB id	PDB Title	Resol.	R fact	Space group	AERO SPACI	Complex	Chain*	SCOP Class**	DIP	Interface area (A^2)	Number of interface residues #	Number of docking decoys ##
12as	ASPARAGINE SYNTHETASE MUTANT C51A, C315A COMPLEXED WITH L- ASPARAGINE AND AMP	2.2	0.16	P1211	0.38	Homo	B:330 A:330	A(d.104.1.1) B(d.104.1.1)	EXIST	1880	52, 50	10000
1a05	CRYSTAL STRUCTURE OF THE COMPLEX OF 3-ISOPROPYLMALATE DEHYDROGENASE FROM THIOBACILLUS FERROOXIDANS WITH 3-ISOPROPYLMALATE	2	0.2	P21212	0.48	Homo	B:358 A:358	A(c.77.1.1) B(c.77.1.1)	-	2629	65, 64	9998
1a0f	CRYSTAL STRUCTURE OF GLUTATHIONE S-TRANSFERASE FROM ESCHERICHIA COLI COMPLEXED WITH GLUTATHIONESULFONIC ACID	2.1	0.18	P212121	0.49	Homo	B:201 A:201	A(a.45.1.1) B(a.45.1.1) A(c.47.1.5) B(c.47.1.5)	EXIST	1656	43, 43	10000
1a4i	HUMAN TETRAHYDROFOLATE DEHYDROGENASE / CYCLOHYDROLASE	1.5	0.2	P212121	0.74	Homo	B:301 A:301	A(c.2.1.7) B(c.2.1.7) A(c.58.1.2) B(c.58.1.2)	-	1359	38, 35	10000
1a78	COMPLEX OF TOAD OVARY GALECTIN WITH THIO-DIGALACTOSE	2	0.19	P1211	0.47	Homo	B:134 A:134	A(b.29.1.3) B(b.29.1.3)	-	510	14, 14	10000
1a[illegible]	[illegible]	[illegible]	0.19	[illegible]	0.41	[illegible]ero	B:267	B(b.42.2.1) B(b.[illegible])	-	1741	50, 51	10000

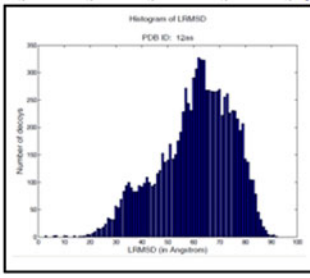
b

```
PDB_ID1(DIP_ID1) <-> List of PDB_ID(DIP_ID) reported as interacting partner of PDB_ID1 in DIP
=============================================================================================
12AS(DIP-9176N)   <->   11AS(DIP-9176N)   12AS(DIP-9176N)
```

Click to know about the interaction partners

PDB id	PDB Title	Resol.	R fact	Space group	AERO SPACI	Complex	Chain*	SCOP Class**	DIP	Interface area (A^2)	Number of interface residues #	Number of docking decoys ##
12as	ASPARAGINE SYNTHETASE MUTANT C51A, C315A COMPLEXED WITH L- ASPARAGINE AND AMP	2.2	0.16	P1211	0.38	Homo	B:330 A:330	A (d.104.1.1) B (d.104.1.1)	EXIST	1880	52, 50	10000

Click to download initial coordinates and transformation matrix



Click to know the distribution of L_{RMSD}

Fig. 3 Example of a query-result of bound complexes in dockYard. (**a**). Users can use a combination of search parameters like PDB identifier, title, resolution, R-factor, Space group, interface area, number of interface residues and so on for selective downloads. (**b**). Query results are hyperlinked to provide more detailed information on initial PDB coordinates, matrices, DIP and L_{RMSD} distribution

Category (iii)

In the category of Variants, we have included structures with 100% sequence identity with category (i) and (ii) above. The modelers using information from this category will find it useful in performing studies that focus on conformation variation. The UniRef100 database is used to find the structures with identical sequence [12]. For category (i) above, we have 797 more protein subunits for 95 dimers derived from UniRef100 database. For category (ii), we have 138 more protein subunits for 14 dimers derived from UniRef100 database.

Category (iv)

In the category of interlogs, modelers are provided with homology related information using UniRef90 and UniRef50 databases. They have the opportunity to apply homology modeling and model decoys for family based comparison. These could guide better use of evolutionary information in improving modeling of docking. The interlogs have been provided for both bound and unbound complexes. For bound cases, at 90% sequence identity, 4515 protein subunits representing 428 dimers are derived from UniRef90 database and 6951 protein subunits representing 609 dimers are derived at 50% sequence identity from UniRef50 database. For unbound cases, 1380 protein subunits representing 34 target dimeric complexes and 1830 protein subunits representing 40 target dimeric complexes are derived from UniRef90 and UniRef50 databases, respectively. It may be noted that although the conservation of interface residues is more of a consequence than the overall similarity measure in the interlogs, we have on purpose based our selection criteria on overall sequence identity because template matching for model building of interlogs requires an overall match for satisfactory implementation.

Data generation

The algorithm of Katzir and coworkers [13] implemented by Gabb et al. in FTDOCK [8] is used for generation of the docking decoys. The coordinates of the interacting subunits are first downloaded from the PDB. Thereafter an arbitrary transformation (translation and rotation) is given to one subunit (called the mobile molecule) with respect to other (called the static molecule) to give the initial coordinates for decoy generation. The FTDOCK program uses a grid box of 0.875 Å to discretize the protein subunits into coarse elements. An additional surface of 1.5 Å in the grid-space is defined for the static molecule. A rotational sampling step of 12° is applied on the mobile molecule for generating new orientations. All decoys generated by the sampling are scored by the grid correlation function [13] using default parameters. The 12° rotation and default translations result in 9240 iterations, from which three best decoys are kept at each step. These result in 27720 decoys of which the top 10,000 ranked by the grid correlation function are accepted. The L_{RMSD} of these decoys is computed and their transformation matrices along with the initial coordinates of the static and mobile molecules are saved. These are directly used as the data resource for our database.

Data transfer method

To facilitate quick downloading of data, and bypass the downloading of the raw coordinates whose average size per protein complex is in the order of gigabytes, we have provided for the users to download the initial PDB coordinates and the transformation matrices of 10000 decoys. A program/script available as Utilities in the download page need to be downloaded to run in the client machine to generate all the coordinates of the docking decoys as per user defined criteria. This way the user can generate the docking decoys in their system locally saving a great amount of time and internet bandwidth. The whole process takes a few minutes to few hours based on the number of decoys to be generated and the size of the initial PDB-coordinate file(s). The time requirement is negligible in comparison to generating the docking decoys through sampling search. Please note that supplying the refined structures would need the raw coordinates and may unnecessarily overwhelm the internet bandwidth.

Usage

To access all the information in the webserver, we have provided web pages for browsing and searching to enable selective downloading, in addition to options for bulk download. The database provides a search tool to customize the list of proteins by protein-specific, protein-complex-specific and docking-decoy-specific parameters. To inform the modelers on the accuracy of stereo-chemical data in the native complex, we have included information on resolution, R-factor and AEROSPACI score. Other protein specific parameters include the PDB identifier, UniRef information, PDB header, space group, SCOP class, and complex type. The dimeric complex related information like chain length of dimers, interface area and number of residues interacting at the interface is also provided. We have provided the modelers a way to filter docking decoys by specifying the range of L_{RMSD} from the native complex structure. The result of the query will show the number of docking decoys present for that PDB within the specified L_{RMSD} range at the *Number of docking decoys* column (Fig. 3a). Modelers interested only in generating interlog-

based models can suitably use the L_{RMSD} criterion to extract near native templates to build homology models. The initial coordinates of the templates and the transformation matrices can be obtained by accessing the PDB hyperlink (Fig. 3b); experimental protein-protein interaction can also be accessed from the DIP hyperlink. The user can choose various parameters outlined above to prepare a customized list for searching and selectively downloading decoys. In all cases, the user may use his/her choice of a refinement algorithm to optimize interface packing geometry of the downloaded decoy/model. To facilitate smooth usage of the website, we have further provided help pages and tutorial pages.

Integrating information from various databases

We have incorporated information from UniRef, in order to extend the coverage of information on protein complexes irrespective of resolution and R-factor. The concept of interlogs is used in exploiting the available information on clustered sets of sequence in the UniRef database. We have been able to include 7003 interlogs, based on the 902 bound complexes, and 1830 interlogs based on 40 target complexes in unbound category. In the search page, the modeler can input PDB name of his/her choice and get representative docking decoys/models based on sequence identity measure derived from the UniRef100, UniRef90 and UniRef50 databases.

Sequence of an individual subunit in hetero-complex belongs to separate clusters in the UniRef database. Therefore, in our query for hetero-complex, we show results for complexes in all those clusters where the sequence for at least one subunit shows a match. The information on output complex is useful in doing various categories of predictive docking: (i) native, (ii) pseudo-native, and (iii) modeled. While the information on native structures is directly available from our database, UniRef can assist in finding pseudo-native interacting partners based on homology. UniRef may also be used to find templates for modeling subunits that can be docked.

Getting an idea about the similarity of protein structure of the interlogs from sequence perspective alone may not be sufficient; therefore, we have also included the SCOP information. Corresponding to each chain there may exist one or more fold classes according to SCOP classification, which are listed in query result page under the column SCOP (Fig. 3a). There is also an option of selecting proteins having at least one occurrence of a particular fold class.

To allow users to have an idea if a certain complex listed in our database has been validated by large-scale experimental assay such as yeast two-hybrid, etc., we have incorporated the protein interaction data from the Database of Interacting Proteins (DIP) [14]. Wherever

Table 2 A comparison of our database *dockYard* with docking resource–DOCKGROUND

Parameter/feature	Database	
	dockYard	DOCKGROUND
Resolution	Available	Available
R-factor	Available	Not available
Aerospaci Score	Available	Available
Mutimeric state and homo/hetero N-ary	Not applicable	Available
Complex type	Available	Available
Interface area	Available	Available
Number of interface residue	Available	Available
SCOP number	Available	Available
GI number	Not available	Available
Space group info	Available	Not available
Option for alternative binding mode, disordered, DNA/RNA, ligand	Not applicable	Available
Option for S-S bonds between chains	Not available	Available
Option for membrane associated	Not available	Available
Option for tangled	Not applicable	Available
Interlogs (using UniRef from UNIPROT database)	Available	Not available
Docking decoys	Available	Limited
L_{RMSD} screening for docking decoys	Available	Limited
DIP information	Available	Not available
Community participation	Available	Not applicable

the interacting partner as reported by DIP has PDB identifier, we have listed them along with DIP identifier hyper-linked under the column DIP in query-result page (Fig. 3b).

An example run

dockYard presents a significant advantage over other decoy databases by presenting a large volume of information available through integration of SCOP, UniRef and DIP information. The presence of UniRef100, UniRef90 and UniRef50 information extends *dockYard* beyond the 902 bound and 40 unbound complexes by allowing users to test and model new quaternary structures. For example, PDB ID: 1AAR and PDB ID: 2JNH, which are not present in *dockYard* database when queried with UniRef90 as a search option will output the same information corresponding to PDB ID: 2OOB. 2OOB is a hetero complex with its B and A subunit having 100% sequence identity with 1AAR:A/B and 2JNH:A, respectively. While 1AAR contains a pair of chains from ubiquitin [15], 2JNH contains only a single chain for UBA Domain from Cbl-b [16]. *dockYard* therefore, easily allows a user to find a template on which a complex of Ubiquitin and UBA Domain be modeled. To be able to use the decoys in *dockYard* for these two chains, one needs to transform the coordinates such that they superpose onto the initial coordinates of 2OOB. Any open source softwares like ProFit (http://www.bioinf.org.uk/software/profit/), POLYPOSE [17] may be used. The transformed and saved coordinates, which serve as the initial coordinates for 1AAR and 2JNH can now be used with the transformation matrices available from 2OOB to generate the decoys. These decoys can be used to test a scoring function for predictive docking where 1AAR:A/B and 2JNH:A are the unbound docking partners and 2OOB is the target protein complex. A similar exercise can be undertaken for any other protein whose coordinates are not known, but there exist a homolog in our database whose coordinates are known and can be used as a template to generate a new structure model.

Comparison with other resources

Our database is unique that it can provide comprehensive information on protein dimers, including the coordinates for their 10000 decoys. Two other database with a similar objective, called BENCHMARK 3.0 [2] and, DOCKGROUND [3, 4] attempts to provide similar information like ours. BENCHMARK 3.0 provides a total of 25 unbound decoys, which are dimers as per PISA server [7]. There is no search and query system, only flat-file download of decoy sets for individual complexes and sequence information. In comparison, we provide 40 dimers in our dataset, and a search and query system. DOCKGROUND also provides a limited number of decoys, which is not very helpful in creating large test/training sets required to perform comprehensive docking model studies. After the first release of this database, we plan to further streamline decoy submission by external users to enhance community-based efforts. Below we summarize a comparison of parameters/features available from our website and DOCKGROUND (Table 2). Important differences are: (i) ours is a database for dimeric complexes, with a plan to upgrade on multimeric complexes in the future. Dimeric complexes form the bulk of the complexes currently available in PDB, and therefore, we have focused on them in the current release. (ii) We also restrict our focus to protein-protein complexes alone and exclude information on heteroatoms, ligands, DNA/RNA. These are also potential additions in a future release. (iii) Keeping with our focus to aid modelers with more docking decoys, we have introduced interlogs in our databases, which is not available in any other publicly accessible database.

Acknowledgments PM thanks All India Council for Technical Education, New Delhi for the national doctoral fellowship. The work was supported by funds from the Centre of Excellence Grants in Bioinformatics and Tuberculosis Research, Department of Biotechnology, New Delhi. There is no conflict of interest in publication of this work.

References

1. Janin J (2007) The targets of CAPRI rounds 6–12. Proteins 69:699–703
2. Hwang H, Pierce B, Mintseris J, Janin J, Weng Z (2008) Protein-protein docking benchmark, version 3.0. Proteins 73:705–709
3. Douguet D, Chen HC, Tovchigrechko A, Vakser IA (2006) DOCKGROUND resource for studying protein-protein interfaces. Bioinformatics 22:2612–2618
4. Gao Y, Douguet D, Tovchigrechko A, Vakser IA (2007) DOCKGROUND system of databases for protein recognition studies: unbound structures for docking. Proteins 69:845–851
5. Andreeva A, Howorth D, Chandonia JM, Brenner SE, Hubbard TJ, Chothia C, Murzin AG (2008) Data growth and its impact on the SCOP database: new developments. Nucleic Acids Res 36: D419–D425
6. Henrick K, Thornton JM (1998) PQS: a protein quaternary structure file server. Trends Biochem Sci 23:358–361
7. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774–797
8. Gabb HAG, Jackson RM, Sternberg MJ (1997) Modelling protein docking using shape complementarity, electrostatics and biochemical information. J Mol Biol 272:106–120
9. Mintseris J, Wiehe K, Pierce B, Anderson R, Chen R, Janin J, Weng Z (2005) Protein–protein docking benchmark 2.0: an update. Proteins: Struct Funct Bioinf 60:214–216
10. Gottschalk KE, Neuvirth H, Schreiber G (2004) A novel method for scoring of docked protein complexes using predicted protein-protein binding sites. Protein Eng Des Sel 17:183–189

11. Bernauer J, Aze J, Janin J, Poupon A (2007) A new protein-protein docking scoring function based on interface residue properties. Bioinformatics 23:555–562
12. Jain E, Bairoch A, Duvaud S, Phan I, Redaschi N, Suzek BE, Martin MJ, McGarvey P, Gasteiger E (2009) Infrastructure for the life sciences: design and implementation of the UniProt website BMC. Bioinformatics 10:136
13. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA (1992) Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 89:2195–2199
14. Salwinski L, Miller CS, Smith AJ, Pettit FK, Bowie JU, Eisenberg D (2004) The database of interacting proteins: 2004 update-. Nucleic Acids Res 32:D449-D451
15. Cook WJ, Jeffrey LC, Carson M, Chen Z, Pickart CM (1992) Structure of a diubiquitin conjugate and a model for interaction with ubiquitin conjugating enzyme (E2). J Biol Chem 267:16467–16471
16. Zhou ZR, Gao HC, Zhou CJ, Chang YG, Hong J, Song AX, Lin DH, Hu HY (2008) Differential ubiquitin binding of the UBA domains from human c-Cbl and Cbl-b: NMR structural and biochemical insights. Protein Sci 17:1805–1814
17. Diamond R (1992) On the multiple simultaneous superposition of molecular structures by rigid body transformations. Protein Sci 1:1279–1287
18. Lesk VI, Sternberg MJ (2008) 3D-Garden: a system for modelling protein-protein complexes based on conformational refinement of ensembles generated with the marching cubes algorithm. Bioinformatics 24:1137–1144
19. Li N, Sun Z, Jiang F (2007) SOFTDOCK application to protein-protein interaction benchmark and CAPRI. Proteins 69:801–808
20. Tong W, Weng Z (2004) Clustering protein-protein docking predictions Conf Proc IEEE. Eng Med Biol Soc 4:2999–3002
21. Pierce B, Weng Z (2007) ZRANK: reranking protein docking predictions with an optimized energy function. Proteins 67:1078–1086
22. Lei H, Duan Y (2004) Incorporating intermolecular distance into protein-protein docking. Protein Eng Des Sel 17:837–845
23. Berchanski A, Shapira B, Eisenstein M (2004) Hydrophobic complementarity in protein-protein docking. Proteins 56:130–142
24. Zacharias M (2003) Protein-protein docking with a reduced protein model accounting for side-chain flexibility. Protein Sci 12:1271–1282
25. Chen R, Weng Z (2003) A novel shape complementarity scoring function for protein-protein docking. Proteins 51:397–408
26. Ma XH, Wang CX, Li CH, Chen WZ (2002) A fast empirical approach to binding free energy calculations based on protein interface information. Protein Eng 15:677–681
27. Norel R, Sheinerman F, Petrey D, Honig B (2001) Electrostatic contributions to protein-protein interactions: fast energetic filters for docking and their physical basis. Protein Sci 10:2147–2161
28. Chen R, Weng Z (2002) Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 47:281–294
29. Palma PN, Krippahl L, Wampler JE, Moura JJ (2000) BiGGER: a new (soft) docking algorithm for predicting protein interactions. Proteins 39:372–384
30. Ackermann F, Herrmann G, Posch S, Sagerer G (1998) Estimation and filtering of potential protein-protein docking positions. Bioinformatics 14:196–205
31. Helmer-Citterich M, Tramontano A (1994) PUZZLE: a new method for automated protein docking based on surface shape complementarity. J Mol Biol 235:1021–1031