Docking by structural similarity at protein-protein interfaces

Rohita Sinha[1], **Petras J. Kundrotas**[1], and **Ilya Vakser**[1,2,*]

[1]Center for Bioinformatics, The University of Kansas, Lawrence, Kansas

[2]Department of Molecular Biosciences, The University of Kansas, Lawrence, Kansas

Abstract

Rapid accumulation of experimental data on protein-protein complexes drives the paradigm shift in protein docking from 'traditional,' template free approaches to template based techniques. Homology docking algorithms based on sequence similarity between target and template complexes can account for up to 20% of known protein-protein interactions. When highly homologous templates for the target complex are not available, but the structure of the target monomers is known, docking by local structural alignment may provide an adequate solution. Such an algorithm was developed based on the structural comparison of monomers to co-crystallized interfaces. A library of the interfaces was generated from co-crystallized protein-protein complexes in PDB. The partial structure alignment algorithm was validated on the DOCKGROUND benchmark sets. The optimal performance of the partial (interface) structure alignment was achieved with the interface residues defined by 12Å distance across the interface. Overall, the partial structural alignment yielded more accurate models than the full structure alignment. Most templates identified by the partial structural alignment had low sequence identity to the target, which makes them hard to detect by sequence-based methods. The results indicate that the structure alignment techniques provide a much needed addition to the docking arsenal, with the combined structural alignment and template free docking success rate significantly surpassing that of the free docking alone.

Keywords

protein recognition; protein modeling; bioinformatics; structure prediction

INTRODUCTION

Protein-protein interactions (PPI) are a key component of life processes at the molecular level. High-throughput methodologies[1⁻3] provide information on PPI primarily at the sequence level. Better understanding of protein function and association requires PPI characterization at the structural level, which is primarily obtained by X-ray crystallography and NMR. However, due to the limitations of these techniques, the experimentally determined structures of protein complexes amount to only a fraction of known PPI. Thus, inherently faster computational methods are of great importance for structural PPI modeling, especially on a genome-wide scale.[4,5]

Following the current paradigm and terminology of the individual proteins structure prediction field, computational methods for PPI structure modeling (docking) can be roughly divided into: (*i*) template-free docking, where sampling of the binding modes is

*Corresponding author: Ilya A. Vakser, Center for Bioinformatics, The University of Kansas, 2030 Becker Drive, Lawrence, Kansas 66047-1620; tel: (785) 864-1057, fax: (785) 864-5558, vakser@ku.edu.

performed with no regard to the possible existence of similar experimentally determined complex structures (templates), and (*ii*) template-based docking, where such similar complexes determine docking predictions.

Free docking methods, initially developed as *ab initio* approaches based on physical potentials (mainly shape complementarity),6 currently are increasingly supplemented by knowledge-based approaches (e.g., docking using statistical potentials,7,8 constraints-driven docking,9 etc.). The increasing availability of the co-crystallize templates allows application of the template-based techniques to docking.10−12 Such approaches follow the paradigm that monomers with similar sequences (structures) form similar complexes. Detecting target-template pairs with sequence identity < 20% by the sequence alignment is an error prone process. Thus there is a need for new concepts in detecting similarity between target and template proteins. The docking problem assumes the *a priori* knowledge of the structures of the participating proteins. Thus the docking templates may be found by the structure alignment of the target monomers to the co-crystallized structures. Another concept is based on the notion that protein functions are attributed to evolutionary conserved surface patches rather than to the full protein scaffold. Thus docking can be performed either by the structure alignment of the full target proteins with the full co-crystallized templates, or by the full target alignment with the interfaces of the co-crystallized templates.

A recent study on protein docking on the basis of interface similarity12 showed promising results on a relatively small test set of complexes (Weng's Docking Benchmark 1.0 set13). However, the structure alignment approach to docking requires comprehensive, large-scale benchmarking and analysis of the methodology, as well as optimized efficiency for adequate application to docking, especially in the high-throughput modeling. We present a comprehensive study of the methodology based on the local structural similarity between the targets and the template interfaces, irrespective of their sequence and global structural similarity. A systematic benchmarking and analysis of the interface alignment was performed on the Dockground benchmark sets14,15 (http://dockground.bioinformatics.ku.edu) with 99 unbound and 372 bound complexes. The performance was compared with the full structure alignment and the template-free docking.

METHODS

Docking protocol

Docking was performed by spatial rearrangement of separate 3D structures of the target monomers to overlap with a co-crystallized interface. The target monomers ('receptor' and 'ligand' defined as the larger and the smaller proteins within a complex) were structurally aligned against the library of interface fragments. The alignment was performed by TM-align16 for C^α atoms only. The C^α alignment reduces the effect of conformational changes upon binding, thus enabling proper overlap of the unbound and bound fragments. The alignment quality was assessed by TM-score17 that has values in the 0 to 1 range. TM-scores < 0.2 indicate no fold similarity, whereas scores > 0.5 point to similar folds.16,17 The resulting pool of initial alignment pairs was purged to retain only significant alignments according to the following criteria: (*i*) TM-score of at least one alignment > 0.4, (*ii*) ≥ 50% of aligned residues for both receptor and ligand should be on the surface, and (*iii*) ≥ 40% of residues in both components of the interface should be covered by the alignments. Transformation matrices from each significant alignment were applied to the target receptor and ligand to generate the pool of models for the target.

The docking is currently implemented as a series of Python scripts running highly optimized TM-align procedure.16 The entire docking protocol for two proteins, involving search through the entire library of interfaces/structures and generation of the full set of matches

(the number of matches varies according to the availability of templates), runs for several hours on a single core processor, which is comparable to typical free docking run times.

Library of interface fragments

Success of any template-based method is directly related to the size and diversity of the pool of putative templates. The structures of the complexes were selected from D$_{OCKGROUND}$. The X-ray structures resolution had to be < 3Å, they had to be at least a dimeric biological unit, and the sequence identity between different complexes had to be < 90%. The resulting set consisted of 11,932 complexes. Biological unit coordinates were obtained from ftp://ftp.wwpdb.org/pub/pdb/data/biounit. In each of these complexes, the backbone atoms belonging to the interface residues were extracted and stored in a library of interface fragments. An interface residue was defined as a residue having at least one atom within a distance (varied from 6 to 16Å, see Results section) of any atom across the interface.

Validation sets

Our primarily set of complexes for validating the docking approach was the D$_{OCKGROUND}$ benchmark set containing 99 unbound protein complexes (27 enzyme-inhibitor, 6 antibody-antigen, 2 cytokine or hormone/receptors, and 64 other complexes). The size of the docking benchmark sets is relatively small because they contain only complexes for which the monomers have both bound and unbound structures available. Thus, for the purpose of this study - to enhance statistical reliability of the results, as well as to evaluate how often proteins with different folds have similar interfaces - we also used an extended set of 372 non-redundant bound complexes extracted from D$_{OCKGROUND}$, where each component of the complex is a single-chain protein. For comparison with the earlier results on structural alignment of complex interfaces,[12] we also tested our methodology on the Boston Benchmark 1.0 set[13] containing 59 complexes.

Classification and evaluation of models

The resulting models were classified based on the parameters of structural alignments between the target and the template monomers (Table I). The alignments were performed on the entire structures of both the target and the template, rather than on the interface fragments used to generate the model. If the model was redundant with the template (Table I) then it was considered as a self-match and not counted in the docking success rate (not evaluated).

Quality of the non-redundant models was evaluated based on the interface C$^\alpha$ RMSD (*i*-RMSD) between the ligand model and its native structure, with the receptor model and the native structure optimally aligned. All models with *i*-RMSD < 10Å were considered, consistent with the previous studies on the docking funnel size.[18] The across-the-interface distance for the interface residues definition in *i*-RMSD calculations was 6Å. Sequence identities between targets and templates were calculated by CLUSTALW.[19]

RESULTS AND DISCUSSION

Interface definition

We adopted the interface definition based on the distance between any atoms of the residues across the interface. The values for such distance in different studies vary in a broad range, from ~4 Å up, depending on the purpose of the study. In the structural alignment procedure, using only the residues in physical contact may lead to wrong results due to the loss of significant structural details at protein interface. On the other hand, large distance cutoff may impair the ability to find local structural similarity at the interface due to the presence of large non-interface parts (in the extreme case, the entire protein structure) in the

'interface' fragments. Generally, selection of the distance for the interface definition in the structural alignment protocol can be considered as an optimization process.

To find the optimal distance, we generated five interface libraries of known protein complexes (see Methods) with different values of the distance: 6Å, 8Å, 10Å, 12Å and 16Å. Models were built and evaluated for 99 unbound complexes from the DOCKGROUND benchmark set using the above five libraries. The docking success rates were defined as the percentage of correctly predicted complexes in top 10, 100 and in all models generated for the target. The criteria for the correct prediction (acceptance criteria) were i-RMSD ≤ 5, 8 and 10Å (earlier analysis of the docking funnels18 suggested that the models with i-RMSD up to 8–10Å can be locally minimized/refined to the near native structures).

The results showed that the success rates for the 10Å, 12Å and 16Å libraries are significantly higher than those for the 6Å and 8Å libraries. Similar results were obtained for the Boston Benchmark 1.0 set. The reason for the relatively poor ranking of models corresponding to the shorter distance libraries was that small fragments lacking well defined secondary structure elements can be aligned to a random place in the target structure (thus generating a model with high TM-score but large i-RMSD). At the same time, the alignment of such fragment of a bound protein (see Methods) to the unbound target interface may have significantly lower TM-score. In the case of relaxed acceptance criteria for the 16Å library, generally, there were always targets for which the 16Å library was unable to generate an acceptable model while the 12Å library (smaller fragments) succeeded. Based on the systematic assessment, it was concluded that the 12Å distance for the interfaces library is the optimal one for the purpose of this study. Thus, all the results presented and discussed below were obtained using the 12Å library.

Global and local structural similarities

The docking approach presented in this paper is designed to detect structural similarities at the interfaces. We also addressed a fundamental question: how often local structural similarity at the interface is *not* accompanied by the structural similarity of the entire monomers? We compared all statistically significant models obtained by partial structural alignments with 12Å library (PSA12) to models obtained by the full-structure alignment (FSA). To ensure statistical significance of our results, the comparison was performed not only for the relatively small DOCKGROUND bound set (99 complexes, DG99) but for a larger DOCKGROUND non-redundant bound set (372 complexes, DG372).

The comparison results are summarized in Table II. Both alignment protocols performed about equally well on both datasets for the higher-accuracy models (i-RMSD < 5Å). Significant parts of the datasets (42 % and 56 % of targets in the DG99 and DG372 datasets, respectively) had the best models produced by both protocols within the same accuracy range. The majority of the best FSA and PSA higher-accuracy models were built using the same template (Table II, numbers in parenthesis for the common models). Thus, local structural similarity at the interfaces of target and template complexes is often accompanied by the global structural similarity between target and template monomers. However, a significant part of both datasets have the best model built by only one of the protocols.

Full structure alignment-only models—PSA12 failure in both cases in Table II for the higher-accuracy DG99 models was related to the absence of residue stretches in the target structures, which lead to the alignment of the template fragments to a non-binding target region. However, for DG372 dataset, there are two other types of worse PSA12 performance for higher-accuracy models. The first one is related to differences in the conformations of interface loops connecting the interface β-strands in the target and the template, leading to a shift in the alignment of the structural fragments. For other targets, PSA12 failed to build the

higher-accuracy models due to the presence of the helix bundle structure motif in the target and the template monomers where only parts of the helices participate in binding. In such case, the interface helix fragments from the template are aligned to a random place on the target helixes, resulting in a wrong model, whereas the FSA protocol correctly aligns the entire helix bundles. In general, for the higher-accuracy models, the target and the template monomers that cause PSA12 failure belong to the same SCOP family/superfamily, but their functions are different.

Out of 21 targets from both datasets with lower-accuracy FSA-only models, only 3 targets have sequence identities with the templates > 20% for one monomer. There were no cases with the sequence identities > 20% for both monomers, implying that such templates are non-trivial for detection by sequence-homology algorithms.

Partial structure alignment-only models—Out of 100 targets for both datasets, for which the best model at all accuracy levels was built by PSA12 only, significant sequence identity (> 20%) between one pair of target-template monomers was observed in just 14 cases. An example is shown in Figure 1A for the target complex of bovine chymotrypsin with eglin C and the template complex of pig trypsin with its inhibitor. The receptors of both complexes have similar conformation (RMSD of aligned structures only 0.9Å) with 45% sequence identity. On the other hand, the ligands have only 5% sequence identity and are so structurally different that FSA did not produce any statistically significant model for this template (TM-score of the global ligand alignment < 0.2). However, both ligands share similar trypsin inhibitor-like loops that make up the entire ligand binding interface. Thus, in this case PSA12 produced an accurate model with i-RMSD = 1.3Å.

The remaining 86 PSA12-only targets had sequence identity with the identified templates < 20% for both monomer pairs. An example is shown in Figure 1B for a PSA12 model of the complex between human cyclophilin and snRNP proteins built using an interface fragment between two chains (out of 4 identical chains in the asymmetric unit) of human transcription factor. The interface fragments used to build the model consist of 71 and 89 residues for the template monomers, but the common structural motif (two short β-strands highlighted in magenta and red, Figure 1B) consists of only 4 residues for both the target and the template. Despite the significant difference in the shape of these β-strands, the PSA12 model has i-RMSD = 4.9Å. The overall structures of the target and the template are very different (with sequence identities 5% and 4% between receptors and ligands, respectively) and the FSA model for this target with the same template has i-RMSD = 37.0Å (i-RMSD = 6.8Å using a different template).

In summary, the results show that the partial and the full structural alignment methods are complementary to each other. Their combination significantly expands the number of identified templates for protein docking. Overall, with the current PDB there is a template to produce a match within acceptable i-RMSD = 10Å accuracy for ~40% of complexes in the benchmark sets.

Ranking of models

Protein docking procedures need adequate scoring functions for the predicted matches. In this study, we used a simple function consisting of TM-scores of the partial (for PSA12) or full (for FSA) structural alignments between receptors and ligands of the target and the template. The results showed that for the lower-accuracy models such simplified function tends to assign low ranks to the near-native predictions generated by either PSA12 or FSA. Lower accuracy models often have structural similarity only between the interfaces of the target and the template, thus decreasing TM-scores of the entire monomers alignments (if any such alignment found at all). At the same time, FSA may find a template complex where

one of the monomers is similar to the target monomer (TM score close to 1), but binds a dissimilar protein at another binding site. This enhances the aggregate TM-score, bringing the incorrect model to the top of the prediction pool. A similar reason causes low ranking of PSA12 models. In addition, there are many small interface fragments in the template library, which may align well (high TM-score) to a non-interface parts of the target complex, thus decreasing the rank of the near-native PSA12 models even further than the corresponding FSA models.

However, the situation is significantly different for the higher-accuracy models, where not only the interfaces of the target and the template complexes are similar but often also the entire structures. For such models, both methods perform equally well in placing the best models at the top of the predictions pool. Still, the PSA12 models had generally lower *i*-RMSD values compared to the corresponding FSA models. Similar structures of one of the target and the template monomers accompanied by dissimilar structures of the other monomer are a common feature for all higher-accuracy PSA12 models. Thus, if it is known that a protein binds different proteins at the same binding site, the partial structural alignment is a better alternative. The superiority of PSA12 was also obvious in the case of domain structure of the target and the template monomers, where the domain orientation in the target and the template is different.

Structure and sequence homology

Structure alignment procedures are computationally demanding (although to a lesser extent than sophisticated multi-template modeling of individual proteins). Thus, for high-throughput structural modeling, where computational speed is essential, it is necessary to understand how many of the structural alignment models can be obtained by a less computationally expensive homology docking approach. For this purpose, we calculated sequence identities between targets and templates (using CLUSTALW). The models were put into three classes of difficulty for the homology modeling: easy (sequence identities of both target-template pairs > 40%), medium (sequence identity of at least one target-template pair from 20% to 40%, with the other not less than 20%), and difficult (sequence identity of at least one target-template pair < 20%).

Distribution of the higher-accuracy models at the different levels of the homology docking complexity showed that the easy cases make up a small part (9.4%) of DG372 dataset, whereas the majority of the models are medium (13.7%) and difficult (19.4%) cases. Interestingly, in a significant number of medium (22 models) and difficult (16 models) cases, the target and the template complexes corresponded to multi-binding proteins, where the same (or similar, with sequence identity > 70%) protein binds dissimilar partners (with sequence identities corresponding to medium or difficult cases for the homology modeling) at the same binding site.

Comparison to free docking

As shown above, the structural alignment is a useful tool in finding templates hardly detectable by fast sequence based methods. On the other hand, it is important to understand where the structural alignment stands with respect to the well established and widely used free docking techniques. Since the docking techniques are usually tested on sets of unbound structures, we compared the performance of PSA12 and the free docking GRAMM-X server20 on the DG99 unbound set (Docking Benchmark). The results are shown in Figure 2. A significant part of the targets successfully docked by GRAMM-X was modeled by PSA12 as well, in the case of both higher- and lower-accuracy models (60% and 71% of all successful free docking models for higher- and lower-accuracy models, respectively). In

turn, PSA12 produced 14 higher-accuracy and 4 lower-accuracy models for targets where GRAMM-X failed in any acceptable-accuracy docking.

The structure alignment approach was also tested on previous CAPRI targets, with limited success, which is in sharp contrast with the significantly higher success rate for the docking benchmark sets. The likely reason is that CAPRI targets have high representation of novel structures, reflecting the effort of the crystallographers providing the CAPRI targets to move beyond easily crystallizable proteins. However, in typical 'real case' modeling of protein-protein complexes of biological interest, the novelty of the structure usually is not a consideration and the existence of homologous co-crystallized complexes, of course, is welcomed. Thus, the docking benchmarks, which follow the increasing availability of co-crystallized homologous complexes, are representative of the biological community needs. In this respect, both the CAPRI target pool and the automatically generated benchmark sets reflect complementary requirements of the docking methodology development.

The structural alignment algorithm is generally more reliable than the free docking methodology. Its utility is increasing with more structural templates being determined by crystallography and NMR. Thus the emerging docking strategy should involve search for available docking templates prior to the free docking modeling. This paradigm is especially valid in genome-wide high-throughput modeling, where most structures of the monomers will be models, with structural accuracy lower than that of the X-ray/NMR.

CONCLUDING REMARKS

We explored the template-based protein-protein docking approach that takes advantage of the structural similarity of the interfaces. A library of 11,932 interfaces was generated from co-crystallized protein-protein complexes in PDB. The optimal performance of the partial (interface) structure alignment was achieved with the interface residues defined by 12Å distance across the interface. The structure alignment algorithm was validated on the Dᴏᴄᴋɢʀᴏᴜɴᴅ benchmark sets.

Templates for higher-accuracy models often shared not only local, but also global structural similarity with the targets, regardless of the target-template sequence identity. However, the templates for lower-accuracy models typically had only local structural similarity with the target structures. Overall, the partial structural alignment approach yielded more accurate models than the full structure alignment. Most templates identified by the partial structural alignment had low sequence identity to the target, which makes them hard to detect by sequence-based methods.

The results show that the structure alignment techniques provide a much needed addition to the docking arsenal, with the combined success rate significantly surpassing that of the template free and homology docking alone. The partial structure alignment is a useful tool in detecting remote docking homologues. The utility of the approach is increasing with the greater availability of the docking templates - co-crystallized protein complexes. Our future work involves development of more sophisticated scoring and computational performance optimization.

Acknowledgments

REFERENCES

1. Schwartz AS, Yu J, Gardenour KR, Finley RL, Ideker T. Cost-effective strategies for completing the interactome. Nature Methods. 2009; 6:55–61. [PubMed: 19079254]

2. Piehler J. New methodologies for measuring protein interactions in vivo and in vitro. Curr Opin Struct Biol. 2005; 15:4–14. [PubMed: 15718127]

3. Bork P, Jensen LJ, Von Mering C, Ramani AK, Lee I, Marcotte EM. Protein interaction networks from yeast to human. Curr Opin Struct Biol. 2004; 14:292–299. [PubMed: 15193308]

4. Aloy P, Russell RB. Structural systems biology: Modelling protein interactions. Nature Rev Mol Cell Biol. 2006; 7:188–197. [PubMed: 16496021]

5. Russell RB, Alber F, Aloy P, Davis FP, Korkin D, Pichaud M, Topf M, Sali A. A structural perspective on protein-protein interactions. Curr Opin Struct Biol. 2004; 14:313–324. [PubMed: 15193311]

6. Vakser IA, Kundrotas P. Predicting 3D structures of protein-protein complexes. Curr Pharm Biotech. 2008; 9:57–66.

7. Kozakov D, Brenke R, Comeau SR, Vajda S. PIPER: An FFT-based protein docking program with pairwise potentials. Proteins. 2006; 65:392–406. [PubMed: 16933295]

8. Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R, Weng Z. Integrating statistical pair potentials into protein complex prediction. Proteins. 2007; 69:511–520. [PubMed: 17623839]

9. de Vries SJ, van Dijk ADJ, Bonvin AMJJ. WHISCY: What information does surface conservation yield? Application to data-driven docking. Proteins. 2006; 63:479–489. [PubMed: 16450362]

10. Keskin O, Nussinov R, Gursoy A. PRISM: Protein-protein interaction prediction by structural matching. Methods Mol Biol. 2008; 484:505–521. [PubMed: 18592198]

11. Lu L, Lu H, Skolnick J. MULTIPROSPECTOR: An algorithm for the prediction of protein-protein interactions by multimeric threading. Proteins. 2002; 49:350–364. [PubMed: 12360525]

12. Gunther S, May P, Hoppe A, Frommel C, Preissner R. Docking without docking: ISEARCH - prediction of interactions using known interfaces. Proteins. 2007; 69:839–844. [PubMed: 17803236]

13. Chen R, Mintseris J, Janin J, Weng Z. A protein-protein docking benchmark. Proteins. 2003; 52:88–91. [PubMed: 12784372]

14. Douguet D, Chen HC, Tovchigrechko A, Vakser IA. DOCKGROUND resource for studying protein-protein interfaces. Bioinformatics. 2006; 22:2612–2618. [PubMed: 16928732]

15. Gao Y, Douguet D, Tovchigrechko A, Vakser IA. DOCKGROUND system of databases for protein recognition studies: Unbound structures for docking. Proteins. 2007; 69:845–851. [PubMed: 17803215]

16. Zhang Y, Skolnick J. TM-align: A protein structure alignment algorithm based on the TM-score. Nucl Acid Res. 2005; 33:2303–2309.

17. Zhang Y, Skolnick J. Scoring function for automated assessment of protein structure template quality. Proteins. 2004; 57:702–710. [PubMed: 15476259]

18. Hunjan J, Tovchigrechko A, Gao Y, Vakser IA. The size of the intermolecular energy funnel in protein-protein interactions. Proteins. 2008; 72:344–352. [PubMed: 18214966]

19. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, H M, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23:2947–2948. [PubMed: 17846036]

20. Tovchigrechko A, Vakser IA. GRAMM-X public web server for protein-protein docking. Nucleic Acids Res. 2006; 34:W310–W314. [PubMed: 16845016]
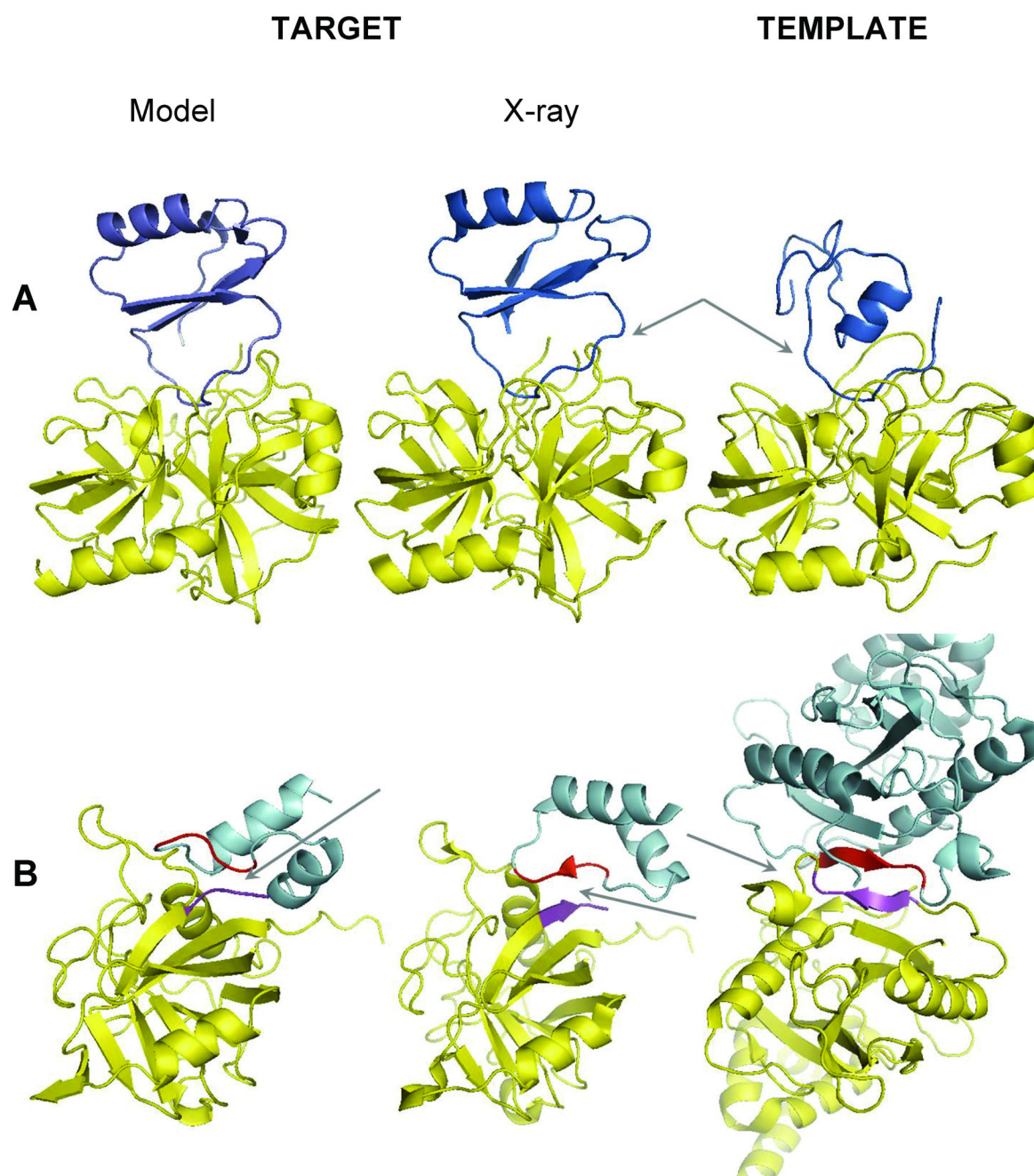
TARGET   **TEMPLATE**

Model   X-ray

Figure 1.
Examples of docking results by partial structural alignment. (A) Non-homologous ligands: target 1acb, chains E and I, and template 1ldt, chains T and L; match i-RMSD = 1.3Å. (B) Non-homologous receptors and ligands: target 1mzw, chains A and B, and template 1m1l, chains B and C; match i-RMSD = 4.9Å. Structural elements responsible for the alignment are in magenta and red and/or are indicated by arrows.
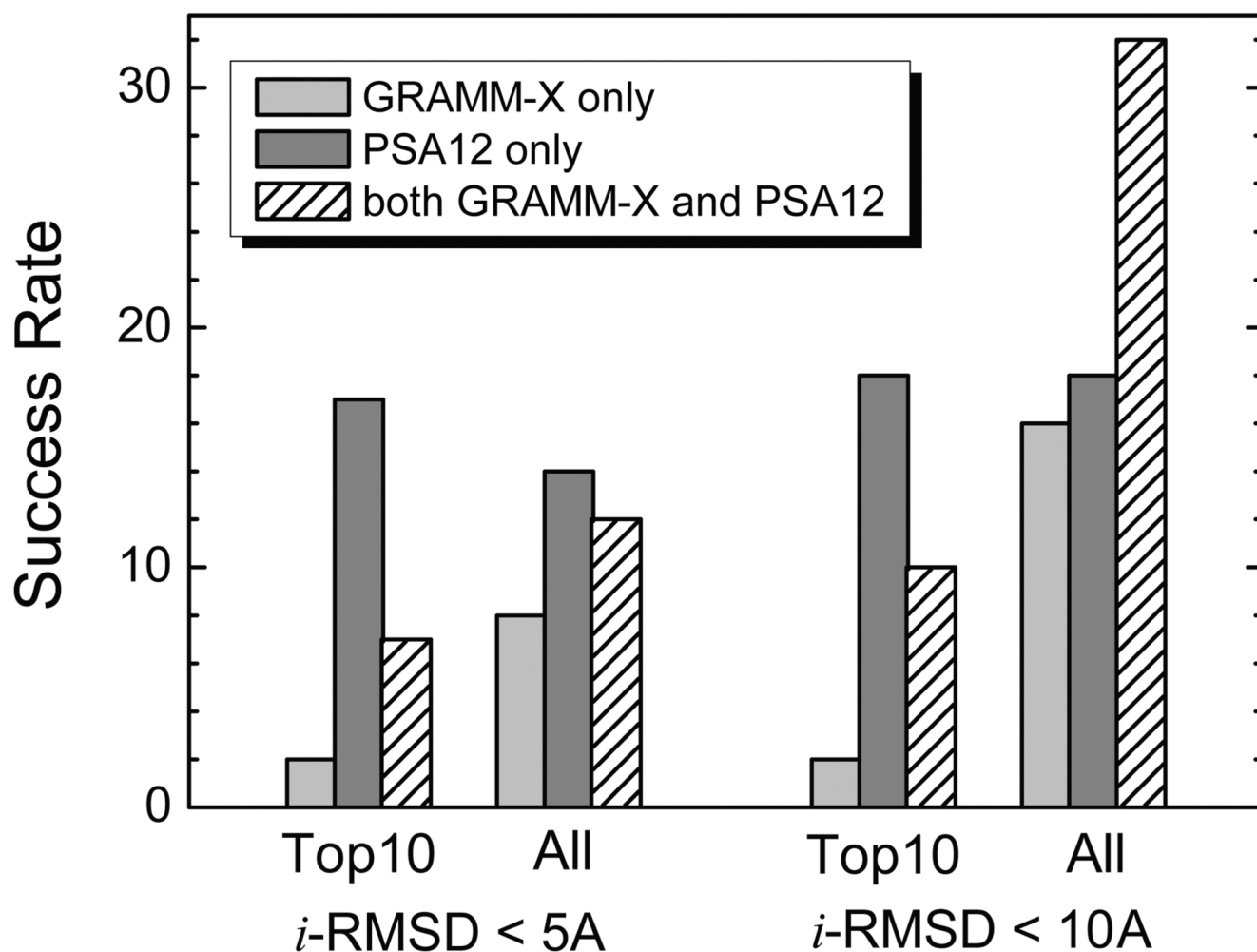
Figure 2.
Comparison of the success rates in template-based and free docking. The success rates are defined as the percentage of targets in DG99 unbound dataset for which higher-accuracy only (*i*-RMSD < 5Å) and all acceptable (*i*-RMSD < 10Å) models were produced by free docking only (GRAMM-X), template-based only (PSA12), and both.

Table I

Classification of Models

Model class	TM-score	Alignment coverage, %	Sequence identity, %[a]
Redundant	0.9 – 1	80 – 100	95 – 100
Structural homolog	0.5 – 0.9	80 – 100	–
Partial structural homolog	0.5 – 0.9	0 – 80	–
Non-homolog	< 0.5	–	–

[a]Sequence identity by TM-align corresponding to the optimal structural alignment of proteins.

Table II

Comparison of Full and Partial Structure Alignment

| Model *i*-RMSD | Number of targets modeled by | | | | | | | |
| | both PSA12 and FSA[a] | | PSA12 only[b] | | FSA only[b] | |
	DG99[c]	DG372	DG99	DG372	DG99	DG372
0 – 5 Å	26 (26)	130 (124)	0	13 (11)	2 (0)	15 (14)
5 – 10 Å	10 (4)	38 (2)	14	73	5	16

[a] Number of targets for which the best models produced by both partial structure alignment using the 12Å library (PSA12) and full-structure alignment (FSA) protocols using the same (number in parenthesis) or different templates have *i*-RMSD in a given accuracy range.

[b] Number of targets for which the best model produced by one of the protocols (PSA12 or FSA) has *i*-RMSD value in a given accuracy range, whereas the other protocol either yielded the best model (based on the same or different template) with *i*-RMSD value in a lower-accuracy range (number in parenthesis) or failed to produce any statistically significant structure alignment for one or both target monomers.

[c] DOCKGROUND unbound benchmark (99 complexes, DG99) and bound benchmark (372 complexes, DG372).

Comparison was performed for all statistically significant models.