

RESEARCH ARTICLE

Open Access

The scoring of poses in protein-protein docking: current capabilities and future directions

Iain H Moal¹, Mieczyslaw Torchala², Paul A Bates² and Juan Fernández-Recio^{1*}

Abstract

Background: Protein-protein docking, which aims to predict the structure of a protein-protein complex from its unbound components, remains an unresolved challenge in structural bioinformatics. An important step is the ranking of docked poses using a scoring function, for which many methods have been developed. There is a need to explore the differences and commonalities of these methods with each other, as well as with functions developed in the fields of molecular dynamics and homology modelling.

Results: We present an evaluation of 115 scoring functions on an unbound docking decoy benchmark covering 118 complexes for which a near-native solution can be found, yielding top 10 success rates of up to 58%. Hierarchical clustering is performed, so as to group together functions which identify near-natives in similar subsets of complexes. Three set theoretic approaches are used to identify pairs of scoring functions capable of correctly scoring different complexes. This shows that functions in different clusters capture different aspects of binding and are likely to work together synergistically.

Conclusions: All functions designed specifically for docking perform well, indicating that functions are transferable between sampling methods. We also identify promising methods from the field of homology modelling. Further, differential success rates by docking difficulty and solution quality suggest a need for flexibility-dependent scoring. Investigating pairs of scoring functions, the set theoretic measures identify known scoring strategies as well as a number of novel approaches, indicating promising augmentations of traditional scoring methods. Such augmentation and parameter combination strategies are discussed in the context of the learning-to-rank paradigm.

Keywords: Docking, Scoring functions, Binding energy, Ranking, SwarmDock

Background

Protein-protein interactions are elemental to almost all biological processes. The atomic-resolution annotation of protein interaction networks can give insights into the kinetics [1-5], thermodynamics [6-10] and organisation [11-13] of the complex systems they constitute, as well as human disease [14,15]. The 3D structures of a protein-protein complex can be used to estimate the effect of mutations [16-21], and thus for protein design [22-29] and determining the functional consequences of mutations associated with diseases (for instance [30]). Further, protein-protein interactions are receiving considerable attention as targets for rational drug design [31-33] and as therapeutic agents [34-37], both endeavours in which structural

information is invaluable. However, in spite of this importance, the rate at which the structures of protein complexes are being solved experimentally lags far behind the rate at which interactions are being discovered. As such, there is a pressing need to fill this ever-growing gap with models derived through computational means such as docking [38-45] and post-docking analysis [46,47].

Despite over three decades of investigations, protein-protein docking remains an unsolved problem. Out of the two critical challenges in docking, the first is sampling, especially in cases with large conformational flexibility. The second, scoring, is the topic of this paper, and is concerned with identifying the correct docking conformations. Scoring attempts to identify the lowest energy poses, and is thus related to the problem of predicting benchmarks of experimental ΔG and $\Delta\Delta G$ values [48,49], an area in which further work is required [50,51]. This is usually achieved by ranking the structures that are

* Correspondence: juanf@bsc.es

¹Joint BSC-IRB Research Program in Computational Biology, Life Science Department, Barcelona Super computing Center, Barcelona 08034, Spain
Full list of author information is available at the end of the article

generated by docking algorithms, and a number of different approaches have been applied to this problem. These range from composite scoring functions using a linear combination of terms, usually models of the underlying physical phenomena at play [52-65], to methods derived from the statistics of structural databases [63,66-70], docking decoys [71-75], experimental binding energies [76], methods based on interface composition and geometry [77-80] or complementarity [81-90], methods based on machine learning [91-97], and methods which account for the characteristics of the binding funnel [98-102]. These approaches span a range of resolution from residue-level to atomic. Further, potentials derived from the field of homology modelling can show promise when applied to interactions [103-105], yet many such methods have not yet had their ability to rank docked structures ascertained.

A large-scale evaluation of the ability of 115 different metrics to rank docked poses using a set of docking decoys generated from the protein-protein docking Benchmark 4.0 [106], using the SwarmDock algorithm [107-109], is presented here. These metrics include docking scores, their constituent terms, molecular mechanics energy functions and methods developed by the protein folding community. We also analyse the union, symmetric difference and relative complement between sets of complexes identified by different methods so as to give an indication of the amount of mutual information embedded in pairs of scoring functions, and the potential for different methods to be combined together synergistically. The results of these investigations give an indication of which approaches are most successful and suggest a number of promising future directions for the improvement of scoring functions.

Results and discussion

Both docking and folding potentials can rank docked poses

The results for the highest performing scoring functions are shown in Figure 1, ordered by top 10 acceptable or better success rate. Numerical values for all the scoring functions are shown in Additional file 1: Table S1, and results ordered by top 1 and top 100 success rates shown in Additional file 1: Figures S1 and S2. Regardless of rank, a medium or better solution could be found for 53% of the complexes, and a high quality for 7%. When ranked and clustered, an acceptable or better solution could be found in the top 100 for up to 92% of the complexes. For top 10 ranked solutions, overall success rates of up to 58% were observed, which dropped to 27% when only the top ranked solution was considered. As expected, methods specifically designed for protein-protein docking feature prominently, with several docking potentials [74,75], pyDock [54], SIPPER [67], DECK [72], PISA [110] and in particular ZRANK2 [58], showing a

very good ability to discriminate near natives from incorrect decoys. The SKOIP intermolecular contact potential [111], which has not been optimised for docking, also performed very well. One consideration that should be made when interpreting the results for the docking-specific methods is that, while these models have been trained on docked structures, none have been specifically trained using SwarmDock decoys. Specifically, using a scoring function outside of the domain used to train it may result in false positives (*e.g.* when encountering an interface with more highly optimised hydrophobic contacts compared to the examples used to train the score) and false negatives (*e.g.* disallowing a near-native structure due to clashes, when the search method for which the scoring function was designed would not produce such contacts). Although SwarmDock does not permit clashes in its solutions, these examples should serve to demonstrate that performance may differ markedly when a different docking algorithm is used, and may be higher on structures generated using the same methods as for training. Thus, the evaluation here is not the same as evaluating the whole docking protocol in the context of the search function used to generate the structures. Nonetheless, the fact that all the docking-specific methods evaluated work well at ranking the SwarmDock decoys indicates that methods designed specifically for one algorithm can also be used to rank poses generated by a different algorithm. This should come as no surprise, as all scoring functions ultimately attempt to identify the structure with the lowest binding energy, and the energy of a given configuration depend only on its coordinates, not on the method used to generate those coordinates. Often, the optimisation acts only to balance the energetic terms. Thus there is a certain degree of interchangeability of scoring functions, and the results can provide information on how well these scoring functions identify the structural aspects that confer affinity, as long as the above caveat is taken into account.

Another consideration is that the performance may be overestimated due to the methods being trained on complexes in the test set. To mitigate such biases and make fair comparisons, the scoring functions were evaluated again using only the updated structures in the Benchmark 4.0, a set which was not used in the training of any of the scoring functions. The results of this are shown in Figure 2. None of the highest performing docking-specific methods perform poorly, indicating that none are drastically over-fitted, while FireDock [59], DARS [73] and SPIDER [78] join the methods above as also being of particular merit. Interestingly, many of the top methods are coarse-grain.

The results harbour a number of surprises. For instance, the fine-grain weighted RosettaDock scoring function performs comparably to the coarse-grain Rosetta environment potential, CG_ENV [112], and the MixRank strategy

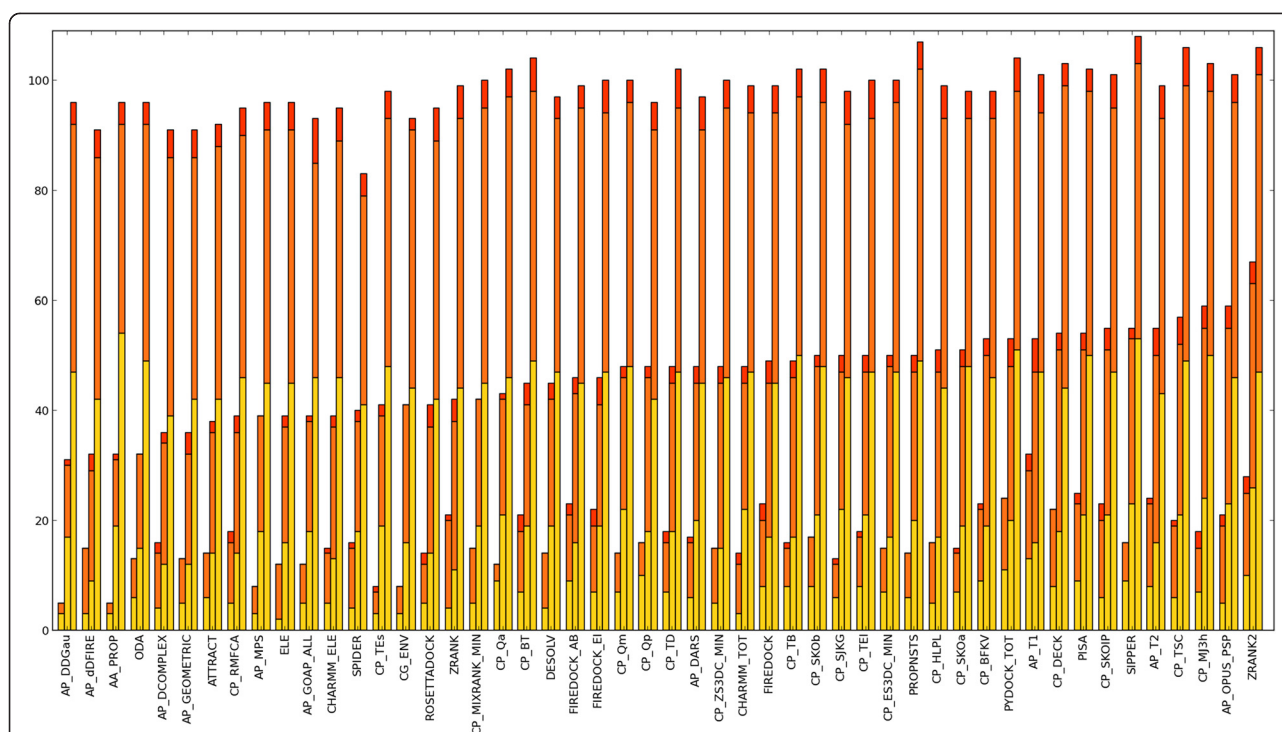


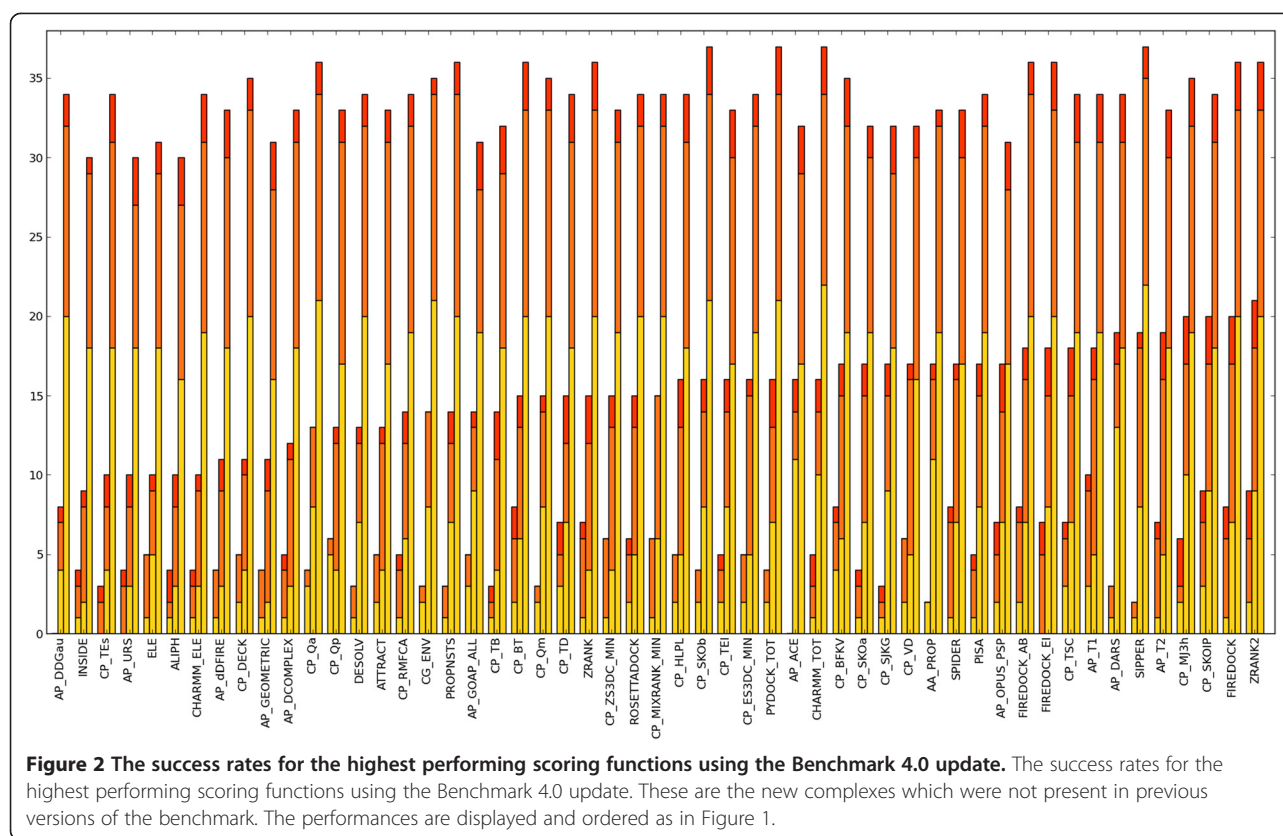
Figure 1 The success rates for the highest performing scoring functions. The success rates for the highest performing scoring functions. The number of complexes for which an acceptable or better solution could be found in the top 1, top 10 and top 100 solutions was calculated for each scoring function, and the best 40 scoring functions for each measure were selected. Acceptable quality solutions are shown in yellow, medium quality solutions in orange, and high quality solutions in red for the three measures (top 1 left, top 10 middle, top 100 right). The functions are ordered by top 10 success rate.

does not perform as well as its constituent S3DC potential [69]. However, perhaps the greatest surprise is the capability for some of the folding potentials to identify near-native solutions. Two methods in particular stand out. The first is the OPUS_PSP potential, a side-chain only orientation-dependent statistical contact potential in which residues are decomposed into rigid and planar chemical moieties [113]. It would be intriguing to see the performance of this potential after training it as an intermolecular potential with crystal structures or near-native docked solution as observations, and/or with docking decoys as the reference state, should a sufficient number of complexes become available. The second is MJ3h, a coarse-grain statistical contact potential which has been corrected for water-to-protein transfer energy [114]. This very simple potential outperforms many more complex scoring functions and, remarkably, has a good ability to distinguish the high quality solutions from other near-native poses.

Difficulty and quality dependent efficacy suggests a need for case-dependent scoring functions

To ascertain whether different scoring methods have different performances depending on difficulty, the analysis

was repeated using only the easier cases, the rigid-body category of the Benchmark 4.0, and the harder cases. As the medium and difficult categories contain fewer complexes, and near-native solutions were found less frequently, they are considered together. The results for the rigid-body cases are shown in Additional file 1: Figures S3, S4 and S5, with high-performing methods similar to those for the whole benchmark. As expected, docking of these cases was more successful than for the flexible cases, with top 1, top 10 and top 100 acceptable or better success rates of up to 30%, 63% and 93% of cases where such a solution is available. For the flexible cases, with results shown in Additional file 1: Figures S6, S7 and S8, only up to 4 complexes were top ranked by any method, too few to reliably distinguish between the different scores. However, for the top 10 and top 100 solutions, success rates of up to 36% and 86% could be achieved. It can be seen that for these complexes, the highest performing methods are highly enriched with coarse-grained scoring functions. This is consistent with the difficulty in correctly predicting the specific atomic contacts for the most flexible cases. While the use of smoothed and coarse-grained functions for the compensation of conformational uncertainty is a common



docking strategy [55,61,62,67,115], the differential scoring performances shown here suggest a role for flexibility-dependent scoring functions, that may be used in conjunction with methods for predicting flexibility [116-120].

In order to further investigate how the ability to score a docked pose depends on the quality of the pose, we investigated the conditional probability of finding a structure of at least a given quality or better given that a solution of at least that quality exists. While an insufficient number of high quality solutions were generated to derive meaningful statistics, this was undertaken for the medium and acceptable quality for all 122 methods, with the results given in Additional file 1: Table S2. When defining a solution as found if it appears in the top 10, 66% of the methods had a greater conditional probability of finding a medium or better solution than an acceptable or better (76/115), indicating a slightly higher success rate for the higher quality solutions. However, the methods evaluated here are biased towards coarse-grain models. When evaluating only the models prefixed with 'AP_', all of which are at or near atomic resolution, this figure rises to 71% (15/21), demonstrating that the high-resolution scoring functions preferentially identify the higher quality solutions. For the scoring functions prefixed with 'CP_', all of which are at residue resolution, the figure is 64% (34/53), indicating a lesser

preferential ability to identify the higher quality solutions. These results further suggest that different scoring strategies are best employed for different docking difficulties.

Differential performance identifies existing and novel scoring strategies

We wished to determine whether the subset of complexes found by any of the methods was significantly different from the subsets found by the other methods. To investigate whether different methods are capable of correctly identifying near-natives in different subsets of the complexes, we looked at all pairs of the methods given in Figure 1. For each pair, we calculated the union (*i.e.* set of complexes found by either methods), symmetric difference (*i.e.* complexes found by only one of the methods) and relative complement (*i.e.* complex found by one method but not the other) of the sets of complexes which were ranked as acceptable or better in the top 10. The numerical results for this analysis are given in Additional file 1: Table S3. While a combined scoring function would not necessarily be able to identify all the correctly docked pose identified by either of the individual scoring functions, nor necessarily miss structures missed by both methods, the cardinalities of the resultant sets can give insights into the extents to which deficiencies in one scoring

function may be compensated by another. For the union and symmetric differences, the larger the size of the resultant sets, the greater the ability of the two methods to identify different complexes. The symmetric difference data is shown in Figure 3. The pairs of scoring functions with highest cardinality are those containing the least mutual information, and are thus most likely to work synergistically together. They suggest some intriguing strategies, including some that have already been developed and applied, as well as novel approaches that might merit from further investigations. For instance, of the 70 complexes which are found using either ODA (33) or PROPNSTS (51), 56 of them are found by only one method and not the other (the other 14 are found by both of them). This indicates that they are detecting different aspects and would work well together. We know this to be true, as the ODA score represents residue and geometry specific exposure of hydrophobic surface [121,122], and PROPNSTS represents chemical complementarity of amino acid pairs [67]. Indeed, the combination of these two terms is the basis for the SIPPER scoring function [67], which routinely performs better than either of the two methods on their own, and can identify near native solutions of acceptable or better quality for 56 complexes. Another common pair of score types with high cardinality is the mix of electrostatics and statistical potentials. Again, this combination

is already exploited in the high performing methods ZRANK [57], ZRANK2 [58] and FireDock [59]. One type of term that is not currently included in the ZRANK and FireDock methods, however, are coarse-grain pair potentials. These results suggest that coarse-grain potentials are capable of finding different subsets of complexes and thus could further enhance these methods. The most promising pairs of methods suggested by the symmetric difference measure, however, are mixtures of SPIDER [78] and other approaches. SPIDER is a novel coarse-grain procedure in which the interfaces of known complexes were decomposed into networks and common motifs found by subgraph mining. When used for scoring, the docking decoys are similarly decomposed and ranked according to the presence of network motifs. While SPIDER is not exceptional on its own, it is good at finding structures which are missed by the other methods. This suggests that this method could be powerful when combined with other techniques. Interestingly, SPIDER distinguishes itself from the other methods in that it explicitly considers multi-body interactions, as opposed to only pairwise interactions, suggesting that other multi-body methods could yield equally promising avenues of exploration. Another interesting aspect of the symmetric difference data is how the scoring functions cluster by similarity. Consider the second principal clusters, shown as the purple subtree in

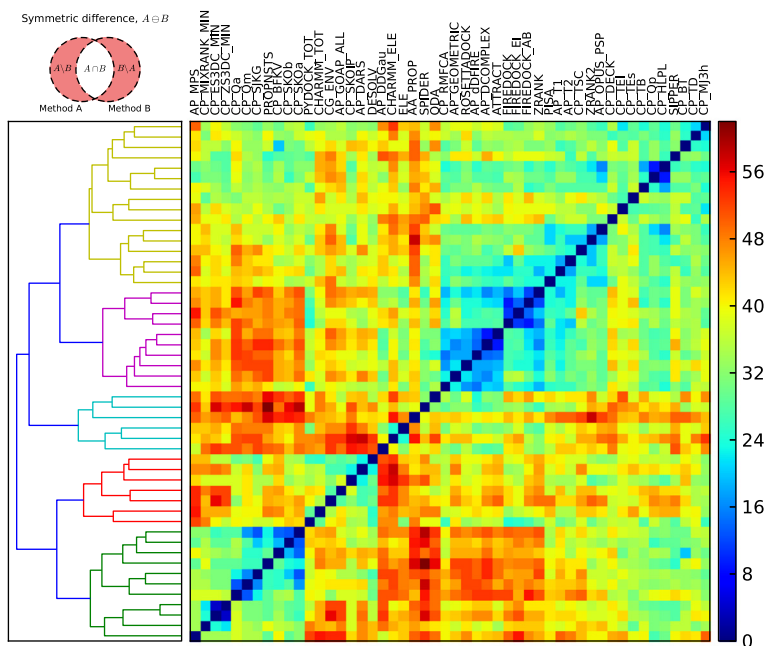


Figure 3 The cardinalities of symmetric differences for pairs of high performing scoring functions. The cardinalities of symmetric differences for pairs of high performing scoring functions. Matrix indices were determined by complete-linkage clustering of the scoring functions, with dissimilarity defined by the cardinality of the symmetric difference sets. The corresponding dendrogram is shown on the left, with the cophenetic distance given by the U-link height. High cardinalities indicate greater ability for the scoring function pairs to identify near-native poses of acceptable or better quality in the top 10 models for different subsets of complexes.

the dendrogram of Figure 3. This cluster splits into two subclusters. The corresponding blue squares indicate that RMFCA, GEOMETRIC, RosettaDock, dDFIRE, DComplex, and ATTRACT are very similar in terms of the near-natives they can correctly identify. In turn, they are also similar to FireDock and ZRANK, which are even more similar to one another. Looking at the first principle cluster, we find similarities between these functions and other methods including ZRANK2, SIPPER, PISA, DECK, other docking potentials and the two highest performing homology modelling potentials, OPUS_PSP and MJ3h. Moreover, these two clusters contain all but one of the composite scoring functions and most of the atomic resolution functions. This makes a stark contrast to the last principle cluster, shown as the green subtree in Figure 3. With only one exception, this is made up only of coarse-grain potentials. These interactions tend to have high cardinalities when paired with the second principle cluster. This is, in part, due to the generally higher performance of the methods in the second cluster, but also due to the fact that these potentials are capable of finding the higher flexibility cases and lower quality near-native solutions, as noted in the previous section. The third and fourth principal clusters, shown in cyan and red, contain an assortment of potentials, including the CHARMM energy, two electrostatics models (ELE and CHARMM_ELE), two residue-level desolvation terms (ODA and CG_ENV), SPIDER, the DARS and GOAP potentials, and a potential extracted from energy changes upon mutation (AP_DDGAU). These disparate potentials have high cardinalities amongst themselves and with the functions in the other principal clusters, indicating that they may be capturing aspects which are overlooked by the other functions.

While the cardinalities of symmetric difference are highly informative, especially when the methods being compared have comparable success rates, focussing on the differences between methods only gives some of the picture. For instance, large symmetric difference sets can be observed when comparing a very high-performing scoring function with a less successful method, not due to their synergistic value but because of deficiencies in the latter. Further, two pairs of methods could have equally sized symmetric differences yet have significantly different success rates due to differences in their intersection. However, such overlap is desirable as it indicates that the two methods can reinforce one another. For these reasons, we investigated the cardinalities of two different sets, the union set and the relative complement. The union data is shown in Figure 4. The data forms two principal clusters shown as red and green subtrees in the dendrogram of Figure 4. The first consists of methods which identify near natives in similar sets of complexes, and thus combining their sets amongst

themselves does not largely expand the range of complexes correctly identified. The second cluster consists of the very high performing scoring functions, in particular the two subclusters which form the first nine scoring functions in the bottom left corner, which can be significantly enhanced if combined with each other and with many of the other functions. Within these two subclusters, one of the methods that performs the least on its own is the total CHARMM energy, yet this it is this method which makes the greatest unions within this cluster, particularly with ZRANK2, AP_OPUS_PSP and CP_TSC, suggesting that this energy function can complement these highly performing methods well. Other intriguing pairs within this cluster include CP_TSC with CP_SKOIP, both very computationally efficient contact potentials, ZRANK2 with CP_MJ3h, and AP_DARS with CP_TSC and ZRANK2. The first primary cluster, corresponding to the red subtree, consists of all the other methods, with great variation amongst themselves and with the second clusters. Within this cluster, the potentials which tend to form the highest cardinalities are PISA, AP_T1, AP_T2, SPIDER, the FireDock potentials and the CHARMM electrostatics potential, all but one of which are atomic resolution. These potentials also form high cardinalities when combined with the potentials in the second principle cluster, particularly with the coarse-grain potentials running from CP_Qa to SIPPER. Also of note are the Rosetta coarse-grain environment potential, CG_ENV, the DESOLVE solvation term and the AP_MPS potential.

The third and final set theoretic comparison method used is the relative complement, as shown in Figure 5. This asymmetric method can be seen as a decomposition of the symmetric difference measure. As the results are ordered by individual success rates, it can be clearly seen that the highest performing methods have the least to gain should they be able to identify the near-natives identified by the other methods (blue left hand side), and *vice versa* (red right hand side). This visualisation allows the identification of the methods which could contribute the most to other methods, in general, by finding rows with incongruously high values. These rows include some methods that have already been identified, such as ELE, CHARMM_ELE, CHARMM_TOT, SPIDER, AP_DARS, FIREDOCK and AP_MPS. Further, it can identify methods which could be used to further improve the already high performing scoring functions, by finding incongruously high values in the leftmost columns. For instance, it also suggests that the ZRANK2 method could be combined with CHARMM_TOT or CG_ENV, or that CP_MJ3h could be combined with ZRANK2, CP_TSC or CHARMM_ELE, or AP_OPUS_PSP with CP_BFKV. More significantly, CP_TSC could be profitably combined with a number of methods, such as CP_DECK, PYDOCK_TOT, CP_SKOa, CP_BVKV,

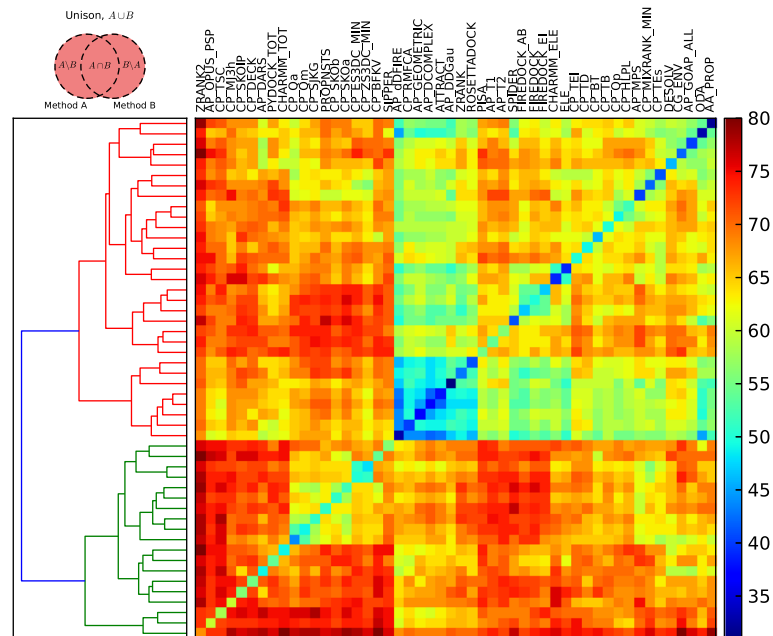


Figure 4 The cardinalities of unions for pairs of high performing scoring functions. The cardinalities of unions for pairs of high performing scoring functions. Clustering was performed as described in Figure 3, with the union defining the matrix. High cardinalities indicate that if a scoring function could be created from the two methods capable of identifying all the near natives correctly identified in the top 10 models for both methods, then it would identify a large proportion of the benchmark complexes.

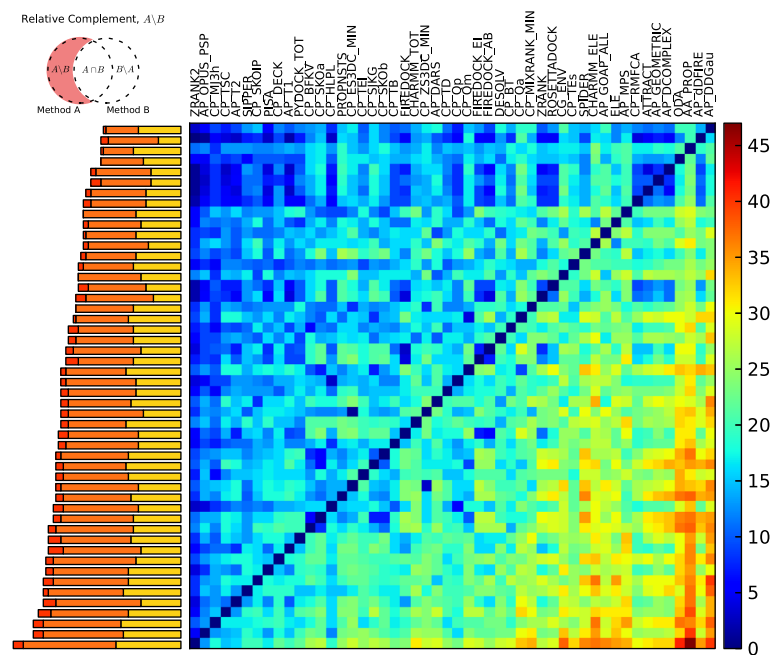


Figure 5 The cardinalities of relative complements for pairs of high performing scoring functions. The cardinalities of relative complements for pairs of high performing scoring functions. Indices are ordered by individual top10 acceptable or better success rate, as shown in the leftmost histogram, with acceptable, medium and high quality success rates shown in yellow, orange and red respectively. This matrix indicates the extent to which the method corresponding to each column can benefit from being able to identify the near-native solutions identified by the methods corresponding to each row. Equivalently, each row indicates the extent to which its method could contribute to the methods in each respective column.

SIPPER, CP_SKOIP, CP_SJKG, CHARMM_TOT, AP_DARS, CG_ENV, CP_Qm or CP_Qa.

Conclusions

Here we have evaluated 115 different scoring functions using a recent docking benchmark [106], yielding acceptable or better solutions in the top 10 for up to 58% of cases. All of the docking-specific scoring functions evaluated performed well, vindicating a range of approaches, including pair potentials trained with docking decoys [72-75,110], composite scoring functions [54,57-59,61,67,112], and a novel approach based on the identification of common motifs in interacting residue networks [78]. Interestingly, some methods taken from the field of homology modelling also performed exceptionally well. In particular, two methods stood out. The first is a simple residue contact potential [114], which can be used to quickly evaluate thousands of structures [123], and thus would be suitable for the initial filtering of poses determined with algorithms that generated a large initial set of decoys, such as Fourier transform docking [89]. The second is a novel statistical potential [113], with accuracy that may be further enhanced if reparameterised as an intermolecular potential or using docking decoys.

We have only considered complexes for which SwarmDock could generate near-native poses. Nevertheless, for at least a third of the Benchmark 4.0, acceptable solutions were generated but not identified by any single scoring function on its own. However, near-native solutions missed by one method were often found by different methods. For instance some functions, particularly the coarse-grained, were better at identifying correct poses for the more difficult cases and where the quality of the docked solutions was lower. On the other hand, atomic potentials were superior for the less flexible cases and higher quality solution. This indicates that flexibility prediction could be used for the selection of the most appropriate scoring functions on a case-by-case basis, or incorporated into the scoring scheme. To investigate which potentials could be combined together, we identified pairs of scoring functions capable of correctly identifying near-natives in different subsets of the benchmark. To achieve this, we used three different set operators, which give complementary pictures of the data. This analysis identified a number of general strategies, such as the combination of amino acid propensities with hydrophobic burial, statistical potentials with electrostatics, and atomic-resolution functions with residue-level potentials. It also allowed the identification of individual methods which appear to capture aspects missed by traditional scoring functions, such as the SPIDER algorithm and the Rosetta environment potential. As SPIDER captures multi-body interactions, and these have not received significant attention in the field of protein-protein docking, this indicates computational

characterisation of cooperative interactions across the interface should be a focus of future research. Finally, these methods also identify specific combinations of terms which may be promising, such as merging the CP_TSC potentials with the AP_DARS potential. For the scoring functions evaluated here, SwarmDock decoys were used. As the scoring functions are easily available, other groups could extend the analysis using structures generated with their own algorithms. However, the fact that methods optimised using different decoys sets still perform well on these structures indicates that scoring functions can be transferred from one docking algorithm to another. Subsequently, the insights garnered here should still apply.

Although this work suggests promising combinations of terms, we have not yet considered how these terms can be combined. A common approach is to take them in linear combination. This makes sense as a first approximation when the terms consist of energy models of physical phenomena, due to the additivity of thermodynamic cycles. However when physical phenomena are coupled, or when features are used that are not rooted in physical phenomena, non-linear relationships between the terms and their utility in ranking arise. Further, it would be desirable to account for heterogeneous data sources, such as predicted flexibility measures as suggested here, but also terms such as sequence conservation data, cluster sizes or agreement with experimental data, or polynomial features such as complex type or, if attempting to merge docking solutions from different algorithms, the provenance of each pose. In these cases, flexible machine learning algorithms capable of inferring these relationships can be used to tailor scoring functions to the structures produced by the sampling methods [124], and indeed examples of this approach can be found in the literature [91-97]. Surprisingly, however, all of these examples have treated the ranking of docked poses not as a ranking problem but as a classification problem, resulting in two issues. Firstly, they are narrow in terms of the models that they use. For instance, when ranking is undertaken in the pointwise approach, it can be reformulated as a classification, regression or ordinal classification problem [125-128]. Similarly a pairwise ranking approach can be formulated as pairwise classification or pairwise regression [129-136], for instance where a model is trained on restraints derived from the fact that acceptable docking poses are superior to incorrect poses, medium are superior to acceptable and incorrect poses, and so on. Further, listwise ranking methods could be employed where, instead of viewing each pose independently and assigning it a score, and instead of comparing pairs of poses to determine which is superior, the whole list is ranked simultaneously as an inseparable set [137-142]. The exploration of how docking ranking performs when

reformulated along these lines remains to be seen. The second issue is that current machine learning based docking scoring methods do not directly address the needs of those who wish to rank docked solutions. For instance, they fail to distinguish high quality solutions from those that are merely acceptable. Additionally, for the purpose of docking, the difference between a top ranked pose and a pose ranked 20th is much greater than the difference between poses ranked 101st and 120th, and this should be reflected in the associated loss function. We suggest that terms such as those explored in this work should be combined in a way which more closely resembles how search engines rank documents. Just as page ranking strives to order lists of documents according to relevance with the most relevant at the top, docking strives to order poses according to their quality with the highest quality at the top. Similarly, just as only the first page of documents is generally of interest in web page ranking, only the top 10 or so docking poses are usually considered for further investigations. Machine-learned ranking has received considerable attention in recent years due to its importance to search engines, with fundamental developments spurred on by initiatives such as the 2009 Internet Mathematics contest, the \$30,000 Yahoo! Learning to Rank Challenge [143] and the ICDM 2013 Expedia Challenge. We believe that such approaches, with loss functions based on measures such as the discounted cumulative gain, and constructed and validated with completely blind features selected within an outer leave-complex-out cross-validation wrapper, will considerably improve our future ability to identify correctly docked structures.

Methods

Generating and evaluating the docking decoys

Ideally, all the scoring functions would be evaluated on different docking decoy sets generated using a number of different methods. This would allow the evaluation of not just scoring functions, but of whole docking protocols/scoring function combinations [124]. However, in order for the calculations to remain tractable, the consideration of increasing numbers of decoys would place a limit on the number of scoring functions which could be evaluated. For this reason, we have chosen to limit our evaluation to a large number of functions using a single decoy set generated using SwarmDock [107,108], a flexible protein-protein docking algorithm which has shown one of the top predictive performances in CAPRI [144]. SwarmDock uses normal modes to model conformational changes. It locates minima on the energy landscape using a hybrid global/local search algorithm, in which translational, orientational and normal coordinate space is simultaneously optimised. As this method

produces relatively few structures, it allows us to evaluate a large number of different scoring methodologies. The decoys were generated using the SwarmDock server [109]. Briefly, non-standard residues were reverted to their precursor amino acid, missing atoms were repaired, missing residue were modelled as alanine residues, atoms were reordered to ensure agreement with the standard PDB atom ordering, and the first location was selected for atoms with multiple location indicators. The algorithm was run using default parameters [107] on a set of 118 of the 176 complexes taken from the Benchmark 4.0 [106]. Four of the Benchmark 4.0 complexes were omitted due to their large size and subsequent difficulties in evaluating a number of the scoring functions (1DE4, 1N2C, 2HMI and 2VIS). The remainder of the complexes were omitted because no solution of acceptable or better quality was generated. For each complex, around 500 decoys were generated. The decoys were scored using the 115 metrics outlined below. Some of the metrics were calculated directly from the structures. For others, the receptor, ligand and complex were evaluated separately, and the final score calculated as $E_{\text{complex}} - (E_{\text{receptor}} + E_{\text{ligand}})$. For each metric, the decoys were reranked and clustered at 3Å resolution in ascending order of energy as described previously [107,108]. For three of the metrics, NHB, SIPPER and PROPSTTS, a positive value corresponds to the most promising solutions, so these were clustered in descending order. For the ranked list of clusters, all but the first (lowest energy) member of each cluster was discarded, leaving a ranked list of structures. For these, the standard CAPRI docking quality measures were calculated: fraction of native contacts (f_{nat}), interface RMSD (IRMSD) and ligand RMSD (LRMSD). These were used to classify the solutions as incorrect ($f_{\text{nat}} < 0.1$ or $(\text{LRMSD} > 10 \text{ \AA}$ and $\text{IRMSD} > 4 \text{ \AA})$), acceptable ($(f_{\text{nat}} \geq 0.3$ and $\text{LRMSD} > 5 \text{ \AA}$ and $\text{IRMSD} > 2 \text{ \AA})$ or $(f_{\text{nat}} \geq 0.1$ and $f_{\text{nat}} < 0.3$ and $(\text{LRMSD} \leq 10 \text{ \AA}$ or $\text{IRMSD} \leq 4 \text{ \AA}))$), medium quality ($(f_{\text{nat}} \geq 0.5$ and $\text{LRMSD} > 1 \text{ \AA}$ and $\text{IRMSD} > 1 \text{ \AA})$ or $(f_{\text{nat}} \geq 0.3$ and $f_{\text{nat}} < 0.5$ and $(\text{LRMSD} \leq 5 \text{ \AA}$ or $\text{IRMSD} \leq 2 \text{ \AA}))$) or high quality ($f_{\text{nat}} \geq 0.5$ and $(\text{LRMSD} \leq 1 \text{ \AA}$ or $\text{IRMSD} \leq 1 \text{ \AA})$), in ascending order of accuracy, in accordance with the CAPRI criteria [145].

Methods evaluated

The scoring functions evaluated are shown in Table 1. Although often experimental, biological and evolutionary information can be used to aid in the scoring of docked poses, this is not always available and here we restrict the analysis to the scoring of global docking solutions using functions which can be calculated from structure alone. Among the functions, there are many contact and distance-dependent residue-level potentials, which are prefixed with 'CP_'

Table 1 A summary of the scoring functions evaluated

CP_DECK [72]	r	The DECK potential, reimplemented based on the original source code.
CP_RMFCFA [146]	r	An α -carbon potential.
CP_RMFCEN1 [147]	r	A 6 bin distance-dependent centroid-centroid potential.
CP_RMFCEN2 [147]	r	A 7 bin distance-dependent centroid-centroid potential.
CP_SKOIP [111]	r	A statistical intermolecular contact potential.
CP_TB [75]	r	A docking contact potential.
CP_TSC [74]	r	A 2 bin docking potential.
PAIR [69]	p	Residue potentials that have been factorised into different energetic contributions
LOCAL [69]	p	(E _{pair} , E _{local} , E _{ZS3DC} , E _{3DC} and E _{3D} respectively). These are prefixed with either 'CP_E' for energies or 'CP_Z' for z-scores, and
S3DC [69]	p	suffixed with '_CB' for the β -carbon potential and '_MIN' for the minimum inter-residue
3DC [69]	p	distance potential. The combination of these
3D [69]	p	into the MixRank ranking strategy is also included. For this method, the 5 largest
CP_MIXRANK [69]	p	complexes failed to produce scores and are thus omitted.
CP_DDGrw [76]	r	The weighted intermolecular contact potential extracted from $\Delta\Delta G$ data, a preliminary model.
CP_DDGrU [76]	r	The unweighted intermolecular contact potential extracted from $\Delta\Delta G$ data, a preliminary model.
CP_BFVK [148]	r	A number of residue-level contact potentials which have been used for protein folding
CP_BL [150]	r	studies. For these, the naming scheme and
CP_BT [151]	r	descriptions can be found elsewhere
CP_GKS [152]	r	[123,149]. Contact energy matrices were
CP_HLPL [153]	r	downloaded from the Potentials'R'Us server.
CP_MJ1 [154]	r	
CP_MJ2 [155]	r	
CP_MJ2h [155]	r	
CP_MJ3h [114]	r	
CP_MJPL [153]	r	
CP_MS [156]	r	
CP_MSBM [157,158]	r	
CP_Qa [159]	r	
CP_Qm [159]	r	
CP_Qp [159]	r	
CP_RO [160]	r	
CP_SJKG [161]	r	
CP_SKOa [162]	r	
CP_SKOb [162]	r	
CP_TD [163]	r	
CP_TEI [164]	r	
CP_TEs [164]	r	
CP_TS [165]	r	
CP_VD [166]	r	

Table 1 A summary of the scoring functions evaluated (Continued)

AP_DCOMPLEX [105]	r	The DComplex potential, reimplementation based on original data file.
AP_dDFIRE [167]	d	The dDFIRE potential.
AP_DFIRE2 [168]	d	The DFIRE 2.0 potential.
AP_T1 [74]	r	The first of two two-step docking potentials.
AP_T2 [74]	r	The second of two two-step docking potentials.
AP_DOPE [169]	r	The standard DOPE potential.
AP_DOPE_HR [169]	r	The high-resolution potentials implemented in MODELLER [170,171].
AP_ACE [172]	d	The atomic contact energy desolvation score, calculated using FireDock [59].
AP_OPUS_PSP [113]	d	The OPUS_PSP folding potential.
AP_GEOMETRIC	d	The geometric potential reported in Li and Liang: Geometric packing potential function for model selection in protein structure and protein-protein binding predictions, unpublished.
AP_DARS [73]	r	The DARS decoys-as-reference-state statistical potential.
AP_URS [73]	r	The URS statistical potential.
AP_MPS [73]	r	The MFP statistical potential.
AP_WENG [173]	r	An atomic contact potential.
AP_calRW [174]	d	The distance-dependent calRW potential.
AP_calRWp [174]	d	The orientation-dependent calRWplus potential.
AP_GOAP_ALL [175]	d	The GOAP potential and its two constituent terms.
AP_GOAP_DF [175]	d	
AP_GOAP_G [175]	d	
AP_PISA [110]	d	The PISA score.
AP_DDGrw [76]	r	The weighted intermolecular contact potential extracted from $\Delta\Delta G$ data.
AP_DDGrU [76]	r	The unweighted intermolecular contact potential extracted from $\Delta\Delta G$ data.
ATTRACT [61]	d	The ATTRACT scoring function, as calculated in PTools [176].
PYDOCK_TOT [54]	i	The PyDock scoring function and the electrostatics, van der Waals and desolvation terms it is composed from.
ELE [54]	i	
VDW [54]	i	
DESOLV [177]	i	
FIREDOCK [59]	d	The general purpose, enzyme-inhibitor and antibody-antigen FireDock scores and the insiderness concavity score and hydrogen-
FIREDOCK_EI [59]	d	bonding, π - π , cation- π and aliphatic potentials they are composed from.
FIREDOCK_AB [59]	d	
INSIDE [59]	d	
HBOND [59]	d	
PI_PI [59]	d	
CAT_PI [59]	d	
ALIPH [59]	d	
SIPPER [67]	i	The SIPPER score and its amino-acid propensity and desolvation constituents.

Table 1 A summary of the scoring functions evaluated (Continued)

PROPST [67]	i
ODA [121,122]	i
ZRANK [57]	d The original ZRANK scoring function.
ZRANK2 [58]	d The reoptimised ZRANK scoring function.
NIP [79]	d Interface packing score.
NSC [79]	d Surface complementarity score.
ROSETTA [112]	d The unweighted Rosetta energy, calculated using PyRosetta.
ROSETTADOCK [112]	d The optimised RosettaDock energy, calculated using PyRosetta.
CG_PP [112]	d The coarse-grain PyRosetta pair-potential, van der Waals, environment potential and β -potential.
CG_VDW [112]	d
CG_ENV [112]	d
CG_BETA [112]	d
HBOND2 [112]	d The atomic-resolution PyRosetta hydrogen bonding potential, amino-acid propensity scores, attractive and repulsive van der Waals energies, pair potential and desolvation energy.
AA_PROP [112]	d
FA_ATR [112]	d
FA_REP [112]	d
PA_PP [112]	d
LK_SOLV [178]	d
NHB [112]	d The total number of hydrogen bonds, calculated using PyRosetta.
CHARMM_TOT [179]	d The total CHARMM energy, electrostatic energy, SASA energy and van der Waals, as calculated using the enerCHARMM script in the MMTSB toolset.
CHARMM_ELE [179]	d
CHARMM_SASA [179]	d
CHARMM_VDW [179]	d
SPIDER [78]	d The sub-graph mining based SPIDER score. As the SPIDER program only allowed scoring using a fixed receptor molecule, the unbound receptor conformation was used for this method, with a relaxed parameter set (dRMSD_CutOff = 1.0, intrCvrAbs_CutOff = 20, intrCvrPer_CutOff = 0.3, intrNumPat_CutOff = 10 and intrAveOcc_CutOff = 2).

Shown are the name of the scoring function and reference, how it was calculated (r for reimplemented, d for downloaded, p for personal communication, i for in-house), and a description/notes.

as well as a number of atomic and near-atomic potentials, which are prefixed with 'AP_'. A number of molecular mechanics terms were included, as well as terms obtained from docking programs and other software. Where scores are composed of multiple terms (RosettaDock, FireDock, ZRANK, ZRANK2, SIPPER, PyDock and ATTRACT), the program provided by the authors was used to calculate and weight the terms using the correct weighting scheme. Most of the functions can be either easily reprogrammed from publically available data, or can be accessed from freely available and well documented programs.

Additional file

Additional file 1: Supplementary Information: This document contains additional figures and tables regarding the success rates and conditional probabilities for each scoring function, and cardinalities for each scoring function pair.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

IM performed the initial conception, scoring, ranking and analysis, and drafted the manuscript. MT performed the docking calculations and classification of the docked poses. IM, MT, PAB and JFR participated in further developing the conception of the work, interpreting the data, performing critical revisions, and have read and approved the final manuscript.

Acknowledgements

The methods evaluated here have been developed through many hours of research in many laboratories around the world. The authors would like to extend their gratitude to all the researchers who have provided their scoring function via their published works, on their web pages and through personal contact. The research leading to these results has received funding from the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement PIEF-GA-2012-327899. This research has also been funded by the Spanish Ministry of Science (BIO2010-22324) and Cancer Research UK.

Author details

¹Joint BSC-IRB Research Program in Computational Biology, Life Science Department, Barcelona Super computing Center, Barcelona 08034, Spain. ²Biomolecular Modelling Laboratory, Cancer Research UK London Research Institute, London WC2A 3LY, UK.

Received: 18 June 2013 Accepted: 25 September 2013

Published: 1 October 2013

References

- Schreiber G, Haran G, Zhou HX: **Fundamental aspects of protein-protein association kinetics.** *Chem Rev* 2009, **109**(3):839–860.
- Gabdouline RR, Wade RC: **Biomolecular diffusional association.** *Curr Opin Struct Biol* 2002, **12**(2):204–213.
- Elcock AH, Sept D, McCammon JA: **Computer simulation of protein-protein interactions.** *J Phys Chem B* 2001, **105**(8):1504–1518.
- Mereghetti P, Kokh D, McCammon JA, Wade RC: **Diffusion and association processes in biological systems: theory, computation and experiment.** *BMC Biophys* 2011, **4**:2.
- Moal IH, Bates PA: **Kinetic rate constant prediction supports the conformational selection mechanism of protein binding.** *PLOS Comput Biol* 2012, **8**:e1002351.
- Vreven T, Hwang H, Pierce BG, Weng Z: **Prediction of protein-protein binding free energies.** *Protein Sci* 2012, **21**(3):396–404.
- Moal IH, Agius R, Bates PA: **Protein-protein binding affinity prediction on a diverse set of structures.** *Bioinformatics* 2011, **27**(21):3002–3009.
- Tian F, Lv Y, Yang L: **Structure-based prediction of protein-protein binding affinity with consideration of allosteric effect.** *Amino Acids* 2012, **43**(2):531–543.
- Audie J, Swanson J: **Advances in the prediction of protein-peptide binding affinities: implications for peptide-based drug discovery.** *Chem Biol Drug Des* 2013, **81**:50–60.
- Zhou P, Wang C, Tian F, Ren Y, Yang C, Huang J: **Biomacromolecular quantitative structure-activity relationship (BioQSAR): a proof-of-concept study on the modeling, prediction and interpretation of protein-protein binding affinity.** *J Comput Aided Mol Des* 2013, **27**:67–78.
- Bhardwaj N, Abyzov A, Clarke D, Shou C, Gerstein MB: **Integration of protein motions with molecular networks reveals different mechanisms for permanent and transient interactions.** *Protein Sci* 2011, **20**(10):1745–1754.

12. Clarke D, Bhardwaj N, Gerstein MB: **Novel insights through the integration of structural and functional genomics data with protein networks.** *J Struct Biol* 2012, **179**(3):320–326.
13. Tuncbag N, Kar G, Gursay A, Keskin O, Nussinov R: **Towards inferring time dimensionality in protein-protein interaction networks by integrating structures: the p53 example.** *Mol Biosyst* 2009, **5**(12):1770–1778.
14. Wang X, Wei X, Thijssen B, Das J, Lipkin SM, Yu H: **Three-dimensional reconstruction of protein networks provides insight into human genetic disease.** *Nat Biotechnol* 2012, **30**(2):159–164.
15. Zhong Q, Simonis N, Li QR, Charlotheaux B, Heuze F, Klitgord N, Tam S, Yu H, Venkatesan K, Mou D, Swearingen V, Yildirim MA, Yan H, Dricot A, Szteto D, Lin C, Hao T, Fan C, Milstein S, Dupuy D, Brasseur R, Hill DE, Cusick ME, Vidal M: **Edgetic perturbation models of human inherited disorders.** *Mol Syst Biol* 2009, **5**:321.
16. Moretti R, Fleishman SJ, Agius R, Torchala M, Bates PA, Kastriitis PL, Rodrigues JP, Trellet M, Bonvin AM, Cui M, Rooman M, Gillis D, Dehouck Y, Moal I, Romero-Durana M, Perez-Cano L, Pallara C, Jimenez B, Fernandez-Recio J, Flores S, Pacella M, Praneeth Kilambi K, Gray JJ, Popov P, Grudinin S, Esquivel-Rodriguez J, Kihara D, Zhao N, Korkin D, Zhu X, et al: **Community-wide evaluation of methods for predicting the effect of mutations on protein-protein interactions.** *Proteins* 2013. doi: 10.1002/prot.24356.
17. Kamisetty H, Ramanathan A, Bailey-Kellogg C, Langmead CJ: **Accounting for conformational entropy in predicting binding free energies of protein-protein interactions.** *Proteins* 2011, **79**(2):444–462.
18. Benedix A, Becker CM, de Groot BL, Caffisch A, Bockmann RA: **Predicting free energy changes using structural ensembles.** *Nat Methods* 2009, **6**:3–4.
19. Brandsdal BO, Smalas AO: **Evaluation of protein-protein association energies by free energy perturbation calculations.** *Protein Eng* 2000, **13**(4):239–245.
20. Almlöf M, Aqvist J, Smalas AO, Brandsdal BO: **Probing the effect of point mutations at protein-protein interfaces with free energy calculations.** *Biophys J* 2006, **90**(2):433–442.
21. Guerois R, Nielsen JE, Serrano L: **Predicting changes in the stability of proteins and protein complexes: a study of more than 1000 mutations.** *J Mol Biol* 2002, **320**(2):369–387.
22. Kortemme T, Baker D: **Computational design of protein-protein interactions.** *Curr Opin Chem Biol* 2004, **8**:91–97.
23. Chen TS, Keating AE: **Designing specific protein-protein interactions using computation, experimental library screening, or integrated methods.** *Protein Sci* 2012, **21**(7):949–963.
24. Baker D: **Prediction and design of macromolecular structures and interactions.** *Philos Trans R Soc Lond B Biol Sci* 2006, **361**(1467):459–463.
25. Lippow SM, Tidor B: **Progress in computational protein design.** *Curr Opin Biotechnol* 2007, **18**(4):305–311.
26. Fleishman SJ, Whitehead TA, Strauch EM, Corn JE, Qin S, Zhou HX, Mitchell JC, Demerdash ON, Takeda-Shitaka M, Terashi G, Moal IH, Li X, Bates PA, Zacharias M, Park H, Ko JS, Lee H, Seok C, Bourquard T, Bernauer J, Poupon A, Aze J, Soner S, Ovali SK, Ozbek P, Tal NB, Haliloglu T, Hwang H, Vreven T, Pierce BG, et al: **Community-wide assessment of protein-interface modeling suggests improvements to design methodology.** *J Mol Biol* 2011, **414**(2):289–302.
27. Mandell DJ, Kortemme T: **Computer-aided design of functional protein interactions.** *Nat Chem Biol* 2009, **5**(11):797–807.
28. Ahmad JN, Li J, Biedermannova L, Kuchar M, Sipova H, Semeradtova A, Cerny J, Petrokova H, Mikulecky P, Polinek J, Stanek O, Vondrasek J, Homola J, Maly J, Osicka R, Sebo P, Maly P: **Novel high-affinity binders of human interferon gamma derived from albumin-binding domain of protein G.** *Proteins* 2012, **80**(3):774–789.
29. Sammond DW, Eletr ZM, Purbeck C, Kimple RJ, Siderovski DP, Kuhlman B: **Structure-based protocol for identifying mutations that enhance protein-protein binding affinities.** *J Mol Biol* 2007, **371**(5):1392–1404.
30. Domene C, Illingworth CJ: **Effects of point mutations in pVHL on the binding of HIF-1 α .** *Proteins* 2012, **80**(3):733–746.
31. Metz A, Ciglia E, Gohlke H: **Modulating protein-protein interactions: from structural determinants of binding to druggability prediction to application.** *Curr Pharm Des* 2012, **18**(30):4630–4647.
32. Gonzalez-Ruiz D, Gohlke H: **Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding.** *Curr Med Chem* 2006, **13**(22):2607–2625.
33. Nisius B, Sha F, Gohlke H: **Structure-based computational analysis of protein binding sites for function and druggability prediction.** *J Biotechnol* 2012, **159**(3):123–134.
34. Zhou P, Wang C, Ren Y, Yang C, Tian F: **Computational peptidology: a New and promising approach to therapeutic peptide design.** *Curr Med Chem* 2013, **20**(15):1985–1996.
35. Szymkowski DE: **Creating the next generation of protein therapeutics through rational drug design.** *Curr Opin Drug Discov Devel* 2005, **8**(5):590–600.
36. Hwang I, Park S: **Computational design of protein therapeutics.** *Drug Discov Today Technol* 2008, **5**(2–3):e43–e48.
37. Wanner J, Fry DC, Peng Z, Roberts J: **Druggability assessment of protein-protein interfaces.** *Future Med Chem* 2011, **3**(16):2021–2038.
38. Kiel C, Serrano L: **Structural data in synthetic biology approaches for studying general design principles of cellular signaling networks.** *Structure* 2012, **20**(11):1806–1813.
39. Aloy P, Russell RB: **Structural systems biology: modelling protein interactions.** *Nat Rev Mol Cell Biol* 2006, **7**(3):188–197.
40. Kiel C, Beltrao P, Serrano L: **Analyzing protein interaction networks using structural information.** *Annu Rev Biochem* 2008, **77**:415–441.
41. Stein A, Mosca R, Aloy P: **Three-dimensional modeling of protein interactions and complexes is going 'omics.** *Curr Opin Struct Biol* 2011, **21**(2):200–208.
42. Beltrao P, Kiel C, Serrano L: **Structures in systems biology.** *Curr Opin Struct Biol* 2007, **17**(3):378–384.
43. Stein M, Gabdoulline RR, Wade RC: **Bridging from molecular simulation to biochemical networks.** *Curr Opin Struct Biol* 2007, **17**(2):166–172.
44. Dell'Orco D: **Fast predictions of thermodynamics and kinetics of protein-protein recognition from structures: from molecular design to systems biology.** *Mol Biosyst* 2009, **5**(4):323–334.
45. Melquiond ASJ, Karaca E, Kastriitis PL, Bonvin AMJJ: **Next challenges in protein-protein docking: from proteome to interactome and beyond.** *WIREs Comput Mol Sci* 2012, **2**(4):642–651.
46. Zhou P, Tian F, Shang Z: **2D depiction of nonbonding interactions for protein complexes.** *J Comput Chem* 2009, **30**(6):940–951.
47. Zhou P, Shang Z: **2D molecular graphics: a flattened world of chemistry and biology.** *Brief. Bioinformatics* 2009, **10**(3):247–258.
48. Kastriitis PL, Moal IH, Hwang H, Weng Z, Bates PA, Bonvin AM, Janin J: **A structure-based benchmark for protein-protein binding affinity.** *Protein Sci* 2011, **20**(3):482–491.
49. Moal IH, Fernandez-Recio J: **SKEMPI: a structural kinetic and energetic database of mutant protein interactions and its use in empirical models.** *Bioinformatics* 2012, **28**(20):2600–2607.
50. Moal IH, Moretti R, Baker D, Fernandez-Recio J: **Scoring functions for protein-protein interactions.** *Curr Opin Struct Biol* 2013. doi: 10.1016/j.sbi.2013.06.017.
51. Kastriitis PL, Bonvin AM: **Are scoring functions in protein-protein docking ready to predict interactomes? Clues from a novel binding affinity benchmark.** *J Proteome Res* 2010, **9**(5):2216–2225.
52. Pallara C, Jimenez-Garcia B, Perez-Cano L, Romero M, Solernou A, Grosdidier S, Pons C, Moal IH, Fernandez-Recio J: **Expanding the frontiers of protein-protein modelling: From docking and scoring to binding affinity predictions and other challenges.** *Proteins* 2013. doi: 10.1002/prot.24387.
53. Demir-Kavuk O, Krull F, Chae MH, Knapp EW: **Predicting protein complex geometries with linear scoring functions.** *Genome Inform* 2010, **24**:21–30.
54. Cheng TM, Blundell TL, Fernandez-Recio J: **pyDock: electrostatics and desolvation for effective scoring of rigid-body protein-protein docking.** *Proteins* 2007, **68**(2):503–515.
55. Lyskov S, Gray JJ: **The RosettaDock server for local protein-protein docking.** *Nucleic Acids Res* 2008, **36**(Web Server issue):W233–238.
56. Dominguez C, Boelens R, Bonvin AM: **HADDOCK: a protein-protein docking approach based on biochemical or biophysical information.** *J Am Chem Soc* 2003, **125**(7):1731–1737.
57. Pierce B, Weng Z: **ZRANK: reranking protein docking predictions with an optimized energy function.** *Proteins* 2007, **67**(4):1078–1086.
58. Pierce B, Weng Z: **A combination of rescoring and refinement significantly improves protein docking performance.** *Proteins* 2008, **72**:270–279.
59. Andrusier N, Nussinov R, Wolfson HJ: **FireDock: fast interaction refinement in molecular docking.** *Proteins* 2007, **69**:139–159.
60. Mashiah E, Nussinov R, Wolfson HJ: **FiberDock: Flexible induced-fit backbone refinement in molecular docking.** *Proteins* 2010, **78**(6):1503–1519.

61. Zacharias M: Protein-protein docking with a reduced protein model accounting for side-chain flexibility. *Protein Sci* 2003, **12**(6):1271–1282.
62. Tovchigrechko A, Vakser IA: Development and testing of an automated approach to protein docking. *Proteins* 2005, **60**(2):296–301.
63. Liang S, Liu S, Zhang C, Zhou Y: A simple reference state makes a significant improvement in near-native selections from structurally refined docking decoys. *Proteins* 2007, **69**(2):244–253.
64. Murphy J, Gatchell DW, Prasad JC, Vajda S: Combination of scoring functions improves discrimination in protein-protein docking. *Proteins* 2003, **53**(4):840–854.
65. Mitra P, Pal D: Using correlated parameters for improved ranking of protein-protein docking decoys. *J Comput Chem* 2011, **32**(5):787–796.
66. Huang SY, Zou X: An iterative knowledge-based scoring function for protein-protein recognition. *Proteins* 2008, **72**(2):557–579.
67. Pons C, Talavera D, de la Cruz X, Orozco M, Fernandez-Recio J: Scoring by intermolecular pairwise propensities of exposed residues (SIPPER): a new efficient potential for protein-protein docking. *J Chem Inf Model* 2011, **51**(2):370–377.
68. Moont G, Gabb HA, Sternberg MJ: Use of pair potentials across protein interfaces in screening predicted docked complexes. *Proteins* 1999, **35**(3):364–373.
69. Felix E, Aloy P, Oliva B: On the analysis of protein-protein interactions via knowledge-based potentials for the prediction of protein-protein docking. *Protein Sci* 2011, **20**(3):529–541.
70. Geppert T, Proschak E, Schneider G: Protein-protein docking by shape-complementarity and property matching. *J Comput Chem* 2010, **31**(9):1919–1928.
71. Muller W, Sticht H: A protein-specifically adapted scoring function for the reranking of docking solutions. *Proteins* 2007, **67**:98–111.
72. Liu S, Vakser IA: DECK: Distance and environment-dependent, coarse-grained, knowledge-based potentials for protein-protein docking. *BMC Bioinformatics* 2011, **12**:280.
73. Chuang GY, Kozakov D, Brenke R, Comeau SR, Vajda S: DARS (Decoys As the Reference State) potentials for protein-protein docking. *Biophys J* 2008, **95**(9):4217–4227.
74. Tobi D: Designing coarse grained-and atom based-potentials for protein-protein docking. *BMC Struct Biol* 2010, **10**:40.
75. Tobi D, Bahar I: Optimal range of protein docking potentials: efficiency and limitations. *Proteins* 2006, **62**(4):970–981.
76. Moal IH, Fernandez-Recio J: Intermolecular contact potentials for protein-protein interactions extracted from binding free energy changes upon mutation. *J Chem Theory Comput* 2013, **9**(8):3715–3727.
77. Chang S, Jiao X, Li CH, Gong XQ, Chen WZ, Wang CX: Amino acid network and its scoring application in protein-protein docking. *Biophys Chem* 2008, **134**(3):111–118.
78. Khashan R, Zheng W, Tropsha A: Scoring protein interaction decoys using exposed residues (SPIDER): a novel multibody interaction scoring function based on frequent geometric patterns of interfacial residues. *Proteins* 2012, **80**(9):2207–2217.
79. Mitra P, Pal D: New measures for estimating surface complementarity and packing at protein-protein interfaces. *FEBS Lett* 2010, **584**(6):1163–1168.
80. Pons C, Glaser F, Fernandez-Recio J: Prediction of protein-binding areas by small-world residue networks and application to docking. *BMC Bioinformatics* 2011, **12**:378.
81. Duhovny D, Nussinov R, Wolfson H: Efficient Unbound Docking of Rigid Molecules. In *Lecture Notes in Computer Science, Volume 2452: Algorithms in Bioinformatics*. Edited by Guigó R, Gusfield D. Berlin, Heidelberg: Springer; 2002:185–200.
82. Shentu Z, Al Hasan M, Bystroff C, Zaki MJ: Context shapes: Efficient complementary shape matching for protein-protein docking. *Proteins* 2008, **70**(3):1056–1073.
83. Chen R, Li L, Weng Z: ZDOCK: an initial-stage protein-docking algorithm. *Proteins* 2003, **52**:80–87.
84. Axenopoulos A, Daras P, Papadopoulos GE, Houstis EN: SP-dock: protein-protein docking using shape and physicochemical complementarity. *IEEE/ACM Trans Comput Biol Bioinform* 2013, **10**:135–150.
85. Gu S, Koehl P, Hass J, Amenta N: Surface-histogram: a new shape descriptor for protein-protein docking. *Proteins* 2012, **80**:221–238.
86. Venkatraman V, Yang YD, Sael L, Kihara D: Protein-protein docking using region-based 3D Zernike descriptors. *BMC Bioinformatics* 2009, **10**:407.
87. Ritchie DW, Kemp GJ: Protein docking using spherical polar Fourier correlations. *Proteins* 2000, **39**(2):178–194.
88. Gabb HA, Jackson RM, Sternberg MJ: Modelling protein docking using shape complementarity, electrostatics and biochemical information. *J Mol Biol* 1997, **272**:106–120.
89. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA: Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. *Proc Natl Acad Sci USA* 1992, **89**(6):2195–2199.
90. Mandell JG, Roberts VA, Pique ME, Kotlovyi V, Mitchell JC, Nelson E, Tsigelny I, Ten Eyck LF: Protein docking using continuum electrostatics and geometric fit. *Protein Eng* 2001, **14**(2):105–113.
91. Chae MH, Krull F, Lorenzen S, Knapp EW: Predicting protein complex geometries with a neural network. *Proteins* 2010, **78**(4):1026–1039.
92. Palma PN, Krippahl L, Wampler JE, Moura JJ: BIGGER: a new (soft) docking algorithm for predicting protein interactions. *Proteins* 2000, **39**(4):372–384.
93. Aze J, Bourquard T, Hamel S, Poupon A, Ritchie DW: Using Kendall- τ Meta-Bagging to Improve Protein-Protein Docking Predictions. In *Lecture Notes in Computer Science, Volume 7036: Pattern Recognition in Bioinformatics*. Edited by Loog M, Wessels L, Reinders M, Ridder D. Berlin, Heidelberg: Springer; 2011:284–295.
94. Bourquard T, Bernauer J, Aze J, Poupon A: A collaborative filtering approach for protein-protein docking scoring functions. *PLoS ONE* 2011, **6**(4):e18541.
95. Bernauer J, Aze J, Janin J, Poupon A: A new protein-protein docking scoring function based on interface residue properties. *Bioinformatics* 2007, **23**(5):555–562.
96. Fink F, Hochrein J, Wolowski V, Merkl R, Gronwald W: PROCOS: computational analysis of protein-protein complexes. *J Comput Chem* 2011, **32**(12):2575–2586.
97. Bordner AJ, Gorin AA: Protein docking using surface matching and supervised machine learning. *Proteins* 2007, **68**(2):488–502.
98. Kozakov D, Schueler-Furman O, Vajda S: Discrimination of near-native structures in protein-protein docking by testing the stability of local minima. *Proteins* 2008, **72**(3):993–1004.
99. Torchala M, Moal IH, Chaleil RA, Agius R, Bates PA: A Markov-chain model description of binding funnels to enhance the ranking of docked solutions. *Proteins* 2013. doi: 10.1002/prot.24369.
100. London N, Schueler-Furman O: FunHunt: model selection based on energy landscape characteristics. *Biochem Soc Trans* 2008, **36**:1418–1421.
101. London N, Schueler-Furman O: Funnel hunting in a rough terrain: learning and discriminating native energy funnels. *Structure* 2008, **16**(2):269–279.
102. London N, Schueler-Furman O: Assessing the energy landscape of CAPRI targets by FunHunt. *Proteins* 2007, **69**(4):809–815.
103. Ravikumar KM, Huang W, Yang S: Coarse-grained simulations of protein-protein association: an energy landscape perspective. *Biophys J* 2012, **103**(4):837–845.
104. Zheng W, Schafer NP, Davtyan A, Papoian GA, Wolynes PG: Predictive energy landscapes for protein-protein association. *Proc Natl Acad Sci USA* 2012, **109**(47):19244–19249.
105. Liu S, Zhang C, Zhou H, Zhou Y: A physical reference state unifies the structure-derived potential of mean force for protein folding and binding. *Proteins* 2004, **56**:93–101.
106. Hwang H, Vreven T, Janin J, Weng Z: Protein-protein docking benchmark version 4.0. *Proteins* 2010, **78**(15):3111–3114.
107. Moal IH, Bates PA: SwarmDock and the Use of normal modes in protein-protein docking. *Int J Mol Sci* 2010, **11**(10):3623–3648.
108. Li X, Moal IH, Bates PA: Detection and refinement of encounter complexes for protein-protein docking: taking account of macromolecular crowding. *Proteins* 2010, **78**(15):3189–3196.
109. Torchala M, Moal IH, Chaleil RA, Fernandez-Recio J, Bates PA: SwarmDock: a server for flexible protein-protein docking. *Bioinformatics* 2013, **29**(6):807–809.
110. Viswanath S, Ravikant DV, Elber R: Improving ranking of models for protein complexes with side chain modeling and atomic potentials. *Proteins* 2012, **81**(4):592–606.
111. Lu H, Lu L, Skolnick J: Development of unified statistical potentials describing protein-protein interactions. *Biophys J* 2003, **84**(3):1895–1901.
112. Chaudhury S, Lyskov S, Gray JJ: PyRosetta: a script-based interface for implementing molecular modeling algorithms using Rosetta. *Bioinformatics* 2010, **26**(5):689–691.

113. Lu M, Dousis AD, Ma J: **OPUS-PSP: an orientation-dependent statistical all-atom potential derived from side-chain packing.** *J Mol Biol* 2008, **376**:288–301.
114. Miyazawa S, Jernigan RL: **Self-consistent estimation of inter-residue protein contact energies based on an equilibrium mixture approximation of residues.** *Proteins* 1999, **34**:49–68.
115. Camacho CJ, Vajda S: **Protein docking along smooth association pathways.** *Proc Natl Acad Sci USA* 2001, **98**(19):10636–10641.
116. Bornot A, Etchebest C, de Brevern AG: **Predicting protein flexibility through the prediction of local structures.** *Proteins* 2011, **79**(3):839–852.
117. Pandey BP, Zhang C, Yuan X, Zi J, Zhou Y: **Protein flexibility prediction by an all-atom mean-field statistical theory.** *Protein Sci* 2005, **14**(7):1772–1777.
118. Gunasekaran K, Nussinov R: **How different are structurally flexible and rigid binding sites? Sequence and structural features discriminating proteins that do and do not undergo conformational change upon ligand binding.** *J Mol Biol* 2007, **365**:257–273.
119. Dobbins SE, Lesk VI, Sternberg MJ: **Insights into protein flexibility: The relationship between normal modes and conformational change upon protein-protein docking.** *Proc Natl Acad Sci USA* 2008, **105**(30):10390–10395.
120. Karaca E, Borvin AM: **A multidomain flexible docking approach to deal with large conformational changes in the modeling of biomolecular complexes.** *Structure* 2011, **19**(4):555–565.
121. Fernandez-Recio J, Totrov M, Skorodumov C, Abagyan R: **Optimal docking area: a new method for predicting protein-protein interaction sites.** *Proteins* 2005, **58**:134–143.
122. Arafat Y, Kamruzzaman J, Karmakar GC, Fernandez-Recio J: **Predicting protein-protein interfaces as clusters of optimal docking area points.** *Int J Data Min Bioinform* 2009, **3**:55–67.
123. Feng Y, Kloczkowski A, Jernigan RL: **Potentials 'R' Us web-server for protein energy estimations with coarse-grained knowledge-based potentials.** *BMC Bioinformatics* 2010, **11**:92.
124. Vajda S, Hall DR, Kozakov D: **Sampling and scoring: A marriage made in heaven.** *Proteins* 2013. doi: 10.1002/prot.24343.
125. Cossack D, Zhang T: **Subset ranking using regression.** In *Proceedings of the 19th annual conference on Learning Theory, COLT'06: 22–25 June 2006; Pittsburgh.* Edited by Simons HU, Lugosi G. Berlin, Heidelberg: Springer; 2006:605–619.
126. Li P, Burges C, Wu Q, Platt JC, Koller D, Singer Y, Roweis S: **McRank: Learning to Rank Using Multiple Classification and Gradient Boosting.** In *Proceedings of the 21st Annual Conference on Neural Information Processing Systems: 3–6 December 2007; Vancouver.* Edited by Platt JC, Koller D, Singer Y, Roweis ST. New York: Curran Associates Inc; 2008:1176–1183.
127. Crammer K, Singer Y: **Pranking with Ranking.** In *Advances in Neural Information Processing Systems 14.* Edited by Dietterich TG, Becker S, Thrun S, Obermayer K. Cambridge: MIT Press; 2001:641–647.
128. Shashua A, Levin A: **Ranking with large margin principle: Two approaches.** In *Advances in Neural Information Processing Systems 15.* Edited by Becker S, Thrun S, Obermayer K. Cambridge: MIT Press; 2003:937–944.
129. Herbrich R, Graepel T, Obermayer K: **Large margin rank boundaries for ordinal regression.** In *Advances in Large Margin Classifiers.* Edited by Smola AJ, Bartlett PL, Schölkopf B, Schuurmans D. Cambridge: MIT Press; 2000:115–132.
130. Freund Y, Iyer R, Schapire RE, Singer Y: **An efficient boosting algorithm for combining preferences.** *J Mach. Learn. Res.* 2003, **4**:933–969.
131. Burges C, Shaked T, Renshaw E, Lazier A, Deeds M, Hamilton N, Hullender G: **Learning to rank using gradient descent.** In *Proceedings of the 22nd international conference on Machine learning, ICML '05: 7–11 August 2005; Bonn.* Edited by De Raedt L, Wrobel S. New York: ACM; 2005:89–96.
132. Zheng Z, Zha H, Zhang T, Chapelle O, Chen K, Sun G: **A General Boosting Method and its Application to Learning Ranking Functions for Web Search.** In *Proceedings of the 21st Annual Conference on Neural Information Processing Systems: 3–6 December 2007; Vancouver.* Edited by Platt JC, Koller D, Singer Y, Roweis ST. New York: Curran Associates Inc; 2008:1697–1704.
133. Cao Y, Xu J, Yan Liu T, Li H, Huang Y, Wuen Hon H: **Adapting ranking SVM to document retrieval.** In *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '06: 6–10 August 2007; Seattle.* Edited by Efthimiadis EN, Dumais ST, Hawking D, Järvelin K. New York: ACM; 2006:186–193.
134. Burges CJC, Ragno R, Le QV: **Learning to Rank with Nonsmooth Cost Functions.** In *Proceedings of the 20th Annual Conference on Neural Information Processing Systems: 4–7 December 2006; Vancouver.* Edited by Schölkopf B, Platt JC, Hoffman T. New York: Curran Associates Inc; 2006:193–200.
135. Wu Q, Burges CJ, Svore KM, Gao J: **Adapting boosting for information retrieval measures.** *Inf. Retr.* 2010, **13**(3):254–270.
136. Li H: **A short introduction to learning to rank.** *IEICE Transactions* 2011, **94-D**(10):1854–1862.
137. Cao Z, Qin T, Liu TY, Tsai MF, Li H: **Learning to rank: from pairwise approach to listwise approach.** In *Proceedings of the 24th international conference on Machine learning, ICML '07: 20–24 June 2007; Corvallis.* Edited by Zoubin G. New York: ACM; 2007:129–136.
138. Xia F, Liu TY, Wang J, Zhang W, Li H: **Listwise approach to learning to rank: theory and algorithm.** In *Proceedings of the 25th international conference on Machine learning, ICML '08: 5–9 July; Helsinki.* Edited by Cohen W. New York: ACM; 2008:1192–1199.
139. Xu J, Li H: **AdaRank: a boosting algorithm for information retrieval.** In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '07: 23–27 July; Amsterdam.* Edited by Wessel K, de Vries AP. New York: ACM; 2007:391–398.
140. Yue Y, Finley T, Radlinski F, Joachims T: **A support vector method for optimizing average precision.** In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '07: 23–27 July; Amsterdam.* Edited by Wessel K, de Vries AP. New York: ACM; 2007:271–278.
141. Taylor M, Guiver J, Robertson S, Minka T: **SoftRank: optimizing non-smooth rank metrics.** In *Proceedings of the 2008 International Conference on Web Search and Data Mining, WSDM '08: 11–12 February; Palo Alto.* Edited by Najork M. New York: ACM; 2008:77–86.
142. Xu J, Yan Liu T, Lu M, Li H, Ying Ma W: **Directly optimizing evaluation measures in learning to rank.** In *Proceedings of the 31th annual international ACM SIGIR conference on Research and development in information retrieval, SIGIR '08: 20–24 July; Singapore.* Edited by Chua TS, Leong M. New York: ACM; 2008:107–114.
143. Chapelle O, Chang Y, Liu TY: **Future directions in learning to rank.** *J Mach Learn Res* 2011, **14**:91–100.
144. Janin J: **Assessing predictions of protein-protein interaction: the CAPRI experiment.** *Protein Sci* 2005, **14**(2):278–283.
145. Lensink MF, Mendez R, Wodak SJ: **Docking and scoring protein complexes: CAPRI 3rd Edition.** *Proteins* 2007, **69**(4):704–718.
146. Rajgaria R, McAllister SR, Floudas CA: **A novel high resolution C^α-C^α distance dependent force field based on a high quality decoy set.** *Proteins* 2006, **65**(3):726–741.
147. Rajgaria R, McAllister SR, Floudas CA: **Distance dependent centroid to centroid force fields using high resolution decoys.** *Proteins* 2008, **70**(3):950–970.
148. Bastolla U, Farwer J, Knapp EW, Vendruscolo M: **How to guarantee optimal stability for most representative structures in the Protein Data Bank.** *Proteins* 2001, **44**(2):79–96.
149. Pokarowski P, Kloczkowski A, Jernigan RL, Kothari NS, Pokarowska M, Kolinski A: **Inferred ideal amino acid interaction forms from statistical protein contact potentials.** *Proteins* 2005, **59**:49–57.
150. Bryant SH, Lawrence CE: **An empirical energy function for threading protein sequence through the folding motif.** *Proteins* 1993, **16**:92–112.
151. Betancourt MR, Thirumalai D: **Pair potentials for protein folding: choice of reference states and sensitivity of predicted native states to variations in the interaction schemes.** *Protein Sci* 1999, **8**(2):361–369.
152. Godzik A, Kolinski A, Skolnick J: **Are proteins ideal mixtures of amino acids? Analysis of energy parameter sets.** *Protein Sci* 1995, **4**(10):2107–2117.
153. Park B, Levitt M: **Energy functions that discriminate X-ray and near native folds from well-constructed decoys.** *J Mol Biol* 1996, **258**(2):367–392.
154. Miyazawa S, Jernigan RL: **Estimation of effective interresidue contact energies from protein crystal structures: quasi-chemical approximation.** *Macromolecules* 1985, **18**(3):534–552.
155. Miyazawa S, Jernigan RL: **Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading.** *J Mol Biol* 1996, **256**(3):623–644.
156. Mirny LA, Shakhnovich EI: **How to derive a protein folding potential? A new approach to an old problem.** *J Mol Biol* 1996, **264**(5):1164–1179.
157. Simons KT, Kooperberg C, Huang E, Baker D: **Assembly of protein tertiary structures from fragments with similar local sequences using simulated annealing and Bayesian scoring functions.** *J Mol Biol* 1997, **268**:209–225.
158. Simons KT, Ruczinski I, Kooperberg C, Fox BA, Bystroff C, Baker D: **Improved recognition of native-like protein structures using a combination of sequence-dependent and sequence-independent features of proteins.** *Proteins* 1999, **34**:82–95.

159. Boniecki M, Rotkiewicz P, Skolnick J, Kolinski A: **Protein fragment reconstruction using various modeling techniques.** *J Comput Aided Mol Des* 2003, **17**(11):725–738.
160. Robson B, Osguthorpe DJ: **Refined models for computer simulation of protein folding. Applications to the study of conserved secondary structure and flexible hinge points during the folding of pancreatic trypsin inhibitor.** *J Mol Biol* 1979, **132**:19–51.
161. Skolnick J, Jaroszewski L, Kolinski A, Godzik A: **Derivation and testing of pair potentials for protein folding. When is the quasichemical approximation correct?** *Protein Sci* 1997, **6**(3):676–688.
162. Skolnick J, Kolinski A, Ortiz A: **Derivation of protein-specific pair potentials based on weak sequence fragment similarity.** *Proteins* 2000, **38**:3–16.
163. Thomas PD, Dill KA: **An iterative method for extracting energy-like quantities from protein structures.** *Proc Natl Acad Sci USA* 1996, **93**(21):11628–11633.
164. Tobi D, Shafran G, Linial N, Elber R: **On the design and analysis of protein folding potentials.** *Proteins* 2000, **40**:71–85.
165. Tanaka S, Scheraga HA: **Medium- and long-range interaction parameters between amino acids for predicting three-dimensional structures of proteins.** *Macromolecules* 1976, **9**(6):945–950.
166. Vendruscolo M, Domany E: **Pairwise contact potentials are unsuitable for protein folding.** *J Chem Phys* 1998, **109**:11101–11108.
167. Yang Y, Zhou Y: **Specific interactions for ab initio folding of protein terminal regions with secondary structures.** *Proteins* 2008, **72**(2):793–803.
168. Yang Y, Zhou Y: **Ab initio folding of terminal segments with secondary structures reveals the fine difference between two closely related all-atom statistical energy functions.** *Protein Sci* 2008, **17**(7):1212–1219.
169. Shen MY, Sali A: **Statistical potential for assessment and prediction of protein structures.** *Protein Sci* 2006, **15**(11):2507–2524.
170. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using MODELLER.** *Curr Protoc Protein Sci* 2007, **2**:2.9.
171. Eswar N, Webb B, Marti-Renom MA, Madhusudhan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using Modeller.** *Curr Protoc Bioinformatics* 2006, **5**:5.6.
172. Zhang C, Vasmatzis G, Cornette JL, DeLisi C: **Determination of atomic desolvation energies from the structures of crystallized proteins.** *J Mol Biol* 1997, **267**(3):707–726.
173. Mintseris J, Pierce B, Wiehe K, Anderson R, Chen R, Weng Z: **Integrating statistical pair potentials into protein complex prediction.** *Proteins* 2007, **69**(3):511–520.
174. Zhang J, Zhang Y: **A novel side-chain orientation dependent potential derived from random-walk reference state for protein fold selection and structure prediction.** *PLoS ONE* 2010, **5**(10):e15386.
175. Zhou H, Skolnick J: **GOAP: a generalized orientation-dependent, all-atom statistical potential for protein structure prediction.** *Biophys J* 2011, **101**(8):2043–2052.
176. Saladin A, Fiorucci S, Poulain P, Prevost C, Zacharias M: **PTools: an opensource molecular docking library.** *BMC Struct. Biol.* 2009, **9**:27.
177. Fernandez-Recio J, Totrov M, Abagyan R: **Identification of protein-protein interaction sites from docking energy landscapes.** *J Mol Biol* 2004, **335**(3):843–865.
178. Lazaridis T, Karplus M: **Effective energy function for proteins in solution.** *Proteins* 1999, **35**(2):133–152.
179. Feig M, Karanicolas J, Brooks CL: **MMTSB Tool Set: enhanced sampling and multiscale modeling methods for applications in structural biology.** *J Mol Graph Model* 2004, **22**(5):377–395.

doi:10.1186/1471-2105-14-286

Cite this article as: Moal et al.: The scoring of poses in protein-protein docking: current capabilities and future directions. *BMC Bioinformatics* 2013 **14**:286.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

