Virtually all biological processes are dependent on protein–promising avenue and necessary task is the structural determination protein interactions (Russell et al., 2004). Toward this goal, one ultimate goal of current research is to identify and elucidate all protein–proteins. First, the individual corresponding proteins from two complexes are sufficiently similar in structure. Second, the combination of individual protein alignments produces a good interface alignment. However, these assumptions are often not true in biologically interesting cases. One example is shown in Figure 1. Two-headed tomato inhibitor-II (TI-II) simultaneously inhibits two molecules of complex structures will be further accelerated (Bravo and Aloy, 2006; Strong et al., 2006). These rich structural data allow us to characterize protein–protein interactions at the atomic level. By analyzing various properties of a protein–protein interface, such as hydrophobicity, hydrogen bonding, buried surface area, topology, planarity, compositions and so on, one can gain insights into the mechanisms of protein–protein recognition (Janin et al., 2005) and TM-align (Zhang and Skolnick, 2005). Extensive classifications of protein domains based on their structural, sequence and functional similarity have also become available, such as SCOP (Murzin et al., 1995) and CATH (Orengo et al., 1997). Since protein–protein interactions are responsible for the stability and/or function of protein complexes, it makes sense to directly compare the interaction modes of protein complexes and to categorize these complexes according to their interaction modes. One simple strategy is to utilize the alignments of individual protein structures and define interaction modes by the orientation of two complexes and/or the overlap in the interfaces (Aloy et al., 2006; Shoemaker et al., 2006). While these studies are informative, they have two assumptions: first, the individual corresponding proteins from two complexes are sufficiently similar in structure. Second, the combination of individual protein alignments produces a good interface alignment. However, these assumptions are often not true in biologically interesting cases. One example is shown in Figure 1. Two-headed tomato inhibitor-II (TI-II) simultaneously inhibits two molecules of complex structures provides an unprecedented opportunity for comparative studies of protein–protein interactions. To facilitate such studies, it is necessary to develop an accurate and efficient computational algorithm for the comparison of protein–protein interaction modes. While there are many structural comparison approaches developed for individual proteins, very few methods are available for protein–protein complexes. Interaction modes. While there are many structural comparison approaches developed for individual proteins, very few methods are available for protein–protein complexes. Interaction modes. While there are many structural comparison approaches developed for individual proteins, very few methods are available for protein–protein complexes.
protease subtilisin (Barrette-Ng et al., 2003). The two interaction sites, located in the two subdomains of TI-II, are well separated (Fig. 1A). Obviously, these two interaction sites cannot be aligned in the comparison of the inhibitor against itself. Moreover, due to the different topologies exhibited in the two TI-II subdomains (Fig. 1B), most structural alignment programs only detect weak similarity between them, and cannot properly align the two interaction sites even when one attempts to align the two subdomains. In contrast, it is straightforward for iAlign to identify highly significant similarity between the two protein–protein interfaces (Fig. 1C), suggesting that the same inhibition mechanism is employed.

In a second strategy, one can extend a structural alignment algorithm for individual proteins to the entire protein complex, e.g. MM-align (Mukherjee and Zhang, 2009). An intrinsic limitation of such an approach is that it does not differentiate between interface regions from non-interface regions. As a result, a significant similarity score does not necessarily mean that two complexes have similar interaction modes. An illustration of this issue is the following: Suppose that a protein A forms complexes with two different partners B and C. Assuming no significant structural changes of A, the comparison of two complexes A/B and A/C will lead to a significant score because of the alignment of A to itself, yet A may associate with B and C in totally different ways. This is particularly problematic when the size of protein A is much bigger than the sizes of B and C. Thus, the similarity between these two A proteins dominates the structural comparison, regardless of the actual interaction modes, which could in fact be quite different. In addition, the alignment of full complexes does not necessarily provide the best alignment of interfaces. In the TI-II/subtilisin example mentioned above, the two interaction sites on the inhibitor still cannot be properly aligned by MM-align.

Thus, the development of a dedicated method for comparing protein–protein interfaces is necessary for studying protein–protein interaction modes. An early approach employed a geometric hashing algorithm (Tsai et al., 1996). More recently, methods that compare physical chemical interactions, non-covalent interactions or contact maps have been proposed in I2I-SiteEngine (Shulman-Peleg et al., 2004), Galinter (Zhu et al., 2008) and CMAPI (Pulim, 2008), respectively. None of these studies, however, provides an assessment of statistical significance of the interface similarity. Moreover, these methods were tested on small datasets, and it is not clear how well they perform in large-scale benchmarks. In particular, it has not been established that interface comparison is useful for the detection of biological relationships.

To address these issues, we present a novel method, iAlign, for the structural comparison of protein–protein interfaces. Below, we first introduce scoring functions for measuring similarity between protein–protein interfaces. We then describe the alignment algorithm and statistical models for the estimation of the level of similarity. Large-scale benchmark tests were performed on both docking models and experimental structures. In addition, iAlign is compared with MM-align and I2I-SiteEngine. Finally, we discuss both the advantages and limitations of our approach.

2 METHODS

We adopt a common definition of a protein–protein interface (Janin et al., 2008): an interfacial contact is defined if the two residues from two separate proteins have at least one pair of their respective heavy atoms within 4.5 Å. A contact is, therefore, defined at the residue level. A protein–protein interface is the collection of all residues that have at least one interfacial contact. The length (or size) of an interface is the number of amino acids constituting the interface.

2.1 Similarity measure

In a typical scenario, one compares a query protein–protein interface against a template interface from a library of interfaces. Suppose a query of length $L_{Q\text{is}}$ aligned to a template of length $L_T$. We consider two scoring functions for measuring interface similarity in iAlign. The first is the Template Modeling score (TM-score; Zhang and Skolnick, 2004),

$$\text{TM-score} = \max_{L_Q \in \text{template}} \sum_{i=1}^{\min(L_Q, L_T)} \left( 1 - \frac{d_{i}}{d_{0}} \right)^{\frac{3}{2}}$$  \hspace{1cm} (1)

where $N_i$ is the number of aligned residue pairs, $d_i$ is the distance in Å between the $C_{\alpha}$ atoms from the $i$-th aligned residue pair and the empirical scaling factor $d_{0} \approx 1.24(L_{Q} - 15)^{1/3} \approx 1.8$. The constants in $d_{0}$ were obtained through fitting the distribution of $C_{\alpha}$ distances in random alignments (Zhang and Skolnick, 2004). In order to calculate the distance $d_{i}$, aligned residues are superimposed with the Kabsch algorithm (Kabsch, 1976). The notation max denotes that the TM-score is the maximum of all possible superpositions. A heuristic iterative extension algorithm is employed to calculate the TM-score (Zhang and Skolnick, 2004), similar to the one used for calculating the GDT-score (Zemla, 2003) and MaxSub (Siew et al., 2000). The definition of the TM-score is exactly the same as used in measuring an alignment of individual proteins by TM-align (Zhang and Skolnick, 2005) or of complexes by MM-align (Mukherjee and Zhang, 2009). To avoid confusion, we denote iTM-score for the TM-score of two interfaces compared by iAlign, nTM-score for TM-score of two non-covalent proteins compared by TM-align, and dTM-score for TM-score of two dimetric complexes compared by MM-align.

Note that these TM-scores have different levels of statistical significance at the same numerical value (see below).

The iTM-score only considers geometric distances. We further introduce the Interface Similarity score (IS-score), which not only measures geometric
The algorithm of iAlign is a further development of the original algorithm implemented in TM-align (Zhang and Skolnick, 2005). Although the algorithm is heuristic, as shown below, the algorithm is sufficiently accurate and highly efficient for practical use. Briefly, the algorithm has two major phases: in the first phase, several guessed solutions are generated through gapless alignments or secondary structure comparison. In the second phase, starting from a guessed alignment, dynamic programming is iteratively applied. The best alignment according to either the iTM-score or IS-score (whichever is specified by the user) is retained.

Four initial alignments are generated during the first phase: (i) the first initial alignment is the gapless alignment that gives the best iTM/IS-score of two interfaces. (ii) the second initial alignment is the best secondary structure match. The match is obtained through dynamic programming with a scoring matrix whose elements are 1 for residues with identical secondary structure and 0 otherwise. A gap penalty of $-1$ is used. (iii) The third initial alignment is obtained by superimposing fragments of interfaces, similar to ideas suggested in Fv-TM-align (Pandit and Skolnick, 2008) and MM-align (Mukhopadhyay and Zhang, 2009). Let $L_{0AB} = \min(L_A, L_B)$. The interfaces are partitioned into fragments with a length of $\min(L_{0AB}/20, 5)$. Superposition is performed for all fragment pairs with at least one pair of residues having identical secondary structure. Corresponding to each fragment superposition, a scoring matrix [Equation (4)] for dynamic programming is calculated and applied to align the two full-length interfaces. The global alignment with the highest iTM-score is the third initial alignment. (iv) After the first three initializations and dynamic programming iterations, the best alignment gives a distance matrix with elements $1/(1 + d_{ij}/d_{0ij})^2$. The elements of the distance matrix and the secondary structure matching matrix are summed with weight 0.5, leading to a new scoring matrix for dynamic programming, which generates the fourth initial alignment.

In the second phase, the above four initial alignments are subjected to dynamic programming iterations for which the scoring matrix is defined as

$$S_j = \begin{cases} (1 + d_{ij}/d_{0ij})^{-2} & \text{for TM-score} \\ (d_{ij} + 3/2)/d_{ij}^{-2} & \text{for IS-score} \end{cases}$$

Here, $d_{ij}$ is the distance between the $i$-th residue of one structure and the $j$-th residue of the other structure, and $d_{0ij} = \min(d_{ai}/a_j + 3/2)/2$, where $a_i$ and $b_j$ are the numbers of interfacial contacts of the $i$-th and $j$-th residues, respectively, and $c_{ij}$ is the number of pairs of overlapped interfacial contacts at the same position (see Supplementary Fig. S1). A pair of interfacial contacts overlaps if the residues forming these contacts are aligned in the two pairs of chains. The scaling factor $s$ $\equiv$ $1.18 + 0.35 L_{ij}^1$ is introduced to make the IS-score length independent (see below).

The constants in $a_i$ were obtained by fitting the distribution of raw scores of unrelated interfaces. Both the iTM/IS-score range from 0 to 1 and identical structures give the maximum score of one.

### 2.2 Alignment algorithm

The algorithm of iAlign is a further development of the original algorithm implemented in TM-align (Zhang and Skolnick, 2005). Although the algorithm is heuristic, as shown below, the algorithm is sufficiently accurate and highly efficient for practical use. Briefly, the algorithm has two major phases: in the first phase, several guessed solutions are generated through gapless alignments or secondary structure comparison. In the second phase, starting from a guessed alignment, dynamic programming is iteratively applied. The best alignment according to either the iTM-score or IS-score (whichever is specified by the user) is retained.

Four initial alignments are generated during the first phase: (i) the first initial alignment is the gapless alignment that gives the best iTM/IS-score of two interfaces. (ii) the second initial alignment is the best secondary structure match. The match is obtained through dynamic programming with a scoring matrix whose elements are 1 for residues with identical secondary structure and 0 otherwise. A gap penalty of $-1$ is used. (iii) The third initial alignment is obtained by superimposing fragments of interfaces, similar to ideas suggested in Fv-TM-align (Pandit and Skolnick, 2008) and MM-align (Mukhopadhyay and Zhang, 2009). Let $L_{0AB} = \min(L_A, L_B)$. The interfaces are partitioned into fragments with a length of $\min(L_{0AB}/20, 5)$. Superposition is performed for all fragment pairs with at least one pair of residues having identical secondary structure. Corresponding to each fragment superposition, a scoring matrix [Equation (4)] for dynamic programming is calculated and applied to align the two full-length interfaces. The global alignment with the highest iTM-score is the third initial alignment. (iv) After the first three initializations and dynamic programming iterations, the best alignment gives a distance matrix with elements $1/(1 + d_{ij}/d_{0ij})^2$. The elements of the distance matrix and the secondary structure matching matrix are summed with weight 0.5, leading to a new scoring matrix for dynamic programming, which generates the fourth initial alignment.

In the second phase, the above four initial alignments are subjected to dynamic programming iterations for which the scoring matrix is defined as

$$S_j = \begin{cases} (1 + d_{ij}/d_{0ij})^{-2} & \text{for TM-score} \\ (d_{ij} + 3/2)/d_{ij}^{-2} & \text{for IS-score} \end{cases}$$

Here, $d_{ij}$ is the distance between the $i$-th residue of one structure and the $j$-th residue of the other structure, and $d_{0ij} = \min(d_{ai}/a_j + 3/2)/2$, where $a_i$ and $b_j$ are the numbers of interfacial contacts of the $i$-th and $j$-th residues, respectively, and $c_{ij}$ is the number of pairs of overlapped interfacial contacts at the same position (see Supplementary Fig. S1). A pair of interfacial contacts overlaps if the residues forming these contacts are aligned in the two pairs of chains. The scaling factor $s$ $\equiv$ $1.18 + 0.35 L_{ij}^1$ is introduced to make the IS-score length independent (see below).

The constants in $a_i$ were obtained by fitting the distribution of raw scores of unrelated interfaces. Both the iTM/IS-score range from 0 to 1 and identical structures give the maximum score of one.

### 2.3 Statistical significance

The statistical significance of iTM/IS-scores is estimated through comparing about 1.8 million random interface pairs (Section 2.4). Figure 2A shows the means of both iTM and IS-scores of random interfaces of similar lengths. For a given length, we consider all random pairs whose lengths are between 95% and 105% of the length. The raw iTM-score is calculated using a fixed value of $d_{ij}$ at 4 Å, and the raw IS-score is calculated using $d_{ij}/d_{0ij}$. Without applying proper scaling factors, the raw iTM/IS-scores decrease exponentially as the length of interfaces increase. In contrast, proper scaling yields approximately length-independent iTM/IS-scores. The means of iTM/IS-scores for random interfaces of similar lengths are 0.206/0.156, respectively.

As shown in Figure 2B, the scores from the random background (RB) follow Gumbel distributions [Equation (5)], also known as type I extreme value distributions. These distributions are over maximum values and, therefore, are suitable to our cases since the iTM/IS-score are the maxima of many structural alignments. The statistical models allow us to calculate the $P$-values of scores, as proposed previously (Levit and Gerstein, 1998). A list of $P$-values and corresponding iTM/IS-scores is given in Table 1. For example, an iTM-score of 0.310/0.21 indicates a similarity at a significant
Statistical significance of the scores for interfaces alignments

\[ P\text{-value} = 5 \times 10^{-1} \times 10^{-2} \times 1 \times 10^{-3} \times 1 \times 10^{-4} \times 1 \times 10^{-5} \times 1 \times 10^{-6} \times 1 \times 10^{-10} \]

\[ \text{iTM-score} = 0.270 \quad 0.311 \quad 0.368 \quad 0.426 \quad 0.484 \quad 0.542 \quad 0.773 \]

\[ \text{IS-score} = 0.191 \quad 0.214 \quad 0.247 \quad 0.279 \quad 0.311 \quad 0.343 \quad 0.473 \]

The three datasets used in the study were derived from the M-TASSER formula. Finally, the maximum likelihood estimates with the EVD package in R by linear fitting to the location and scale parameters, which were obtained through maximum likelihood estimates with the EVD package in R (http://www.r-project.org/). In this procedure, alignments of each of the interfaces of different lengths, due to score normalization, were therefore discarded for this analysis, although many have significant iTM/IS-scores according to iAlign.

\[ P(\text{value of 0.01. One may use these scores to quickly estimate statistical significance.}}

A more accurate estimation of statistical significance is achieved through modeling the distributions of scores at specific lengths. Since the scores are asymmetric for interfaces of different lengths, due to score normalization, we calculate the \( P\)-value that corresponds to the higher score, namely, the score normalized by the smaller interface. This gives a single \( P\)-value for a pair of interfaces. Supplementary Figure S2 shows the observed and modeled distributions of scores at various lengths. Each distribution is modeled by the Gumbel distribution, where 

\[ P(\text{value}) \sim \exp(-e^{\text{exp}(z)}) \]

where \( z = (\mu - \mu) / \sigma \). The variable \( z \) denotes the iAlign score; \( \mu \) and \( \sigma \) are the location and the scale parameters, respectively. These parameters are estimated through linear regression fits

\[ \mu = \alpha \ln(L_q) + c \ln(L_f) \]

\[ \sigma = \exp(\ln(L_q) + \ln(L_f)) \]

The parameters \( \alpha \) to \( f \), given in Supplementary Table S1, were obtained by linear fitting to the location and scale parameters, which were obtained through maximum likelihood estimates with the EVD package in R (http://www.r-project.org/). Finally, the \( P\)-value is calculated using the formula

\[ P\text{-value} = 1 - \exp(-e^{\text{exp}(z)}) \]  

2.4 Datasets

The three datasets used in the study were derived from the M-TASSER template library (Chen and Skolnick, 2008). The library consists of 1838 dimeric protein–protein complexes, non-redundant at 35% sequence identity. Since coiled-coil complexes are trivially similar, we removed the 48 such complexes from the library, resulting in 1790 complexes. Since coiled–coil complexes are trivially similar, we removed the 48 such complexes from the library, resulting in 1790 complexes. For each pair of complexes, the structural similarity of individual proteins that form the complexes was assessed with TM-align, which reports mTM-score between individual proteins. There are four combinations of individual proteins from two dimers. We discarded the pair of complexes if the maximum mTM-score among these four combinations is higher than 0.35. Note that for iTM-score a value of 0.35 suggests that two protein structures are dissimilar (Zhang et al., 2006). The protein–protein interfaces of the remaining complex pairs are the RB chosen for estimating the statistical significance described above.

2.4.1 Random background RB is a set of \( \sim 1.77 \) million pairs of interfaces curated from all against–all comparisons of 1790 complexes. For each pair of complexes, the structural similarity of individual proteins that form the complexes was assessed with TM-align, which reports mTM-score between individual proteins. There are four combinations of individual proteins from two dimers. We discarded the pair of complexes if the maximum mTM-score among these four combinations is higher than 0.35. Note that for iTM-score a value of 0.35 suggests that two protein structures are dissimilar (Zhang et al., 2006). The protein–protein interfaces of the remaining complex pairs are the RB chosen for estimating the statistical significance described above.

2.4.2 Dimer1517 From the template library, we collected 1517 dimers that have SCOP assignments (version 1.75). The set consists of 327 heterodimers and 1190 homodimers. We further examine all against–all pairs among Dimer1517. According to SCOP, protein domains within the same superfamily are biologically related (Murzin et al., 1995). Two complexes are considered to have related protein–protein interfaces, if (i) at least one pair of their interacting domains has the same SCOP superfamily assignments, and (ii) the two protein–protein interfaces share strong structural similarity. To avoid self-testing iAlign, we designed an alternative procedure to assess interface similarity for complexes that have one pair of domains from the same SCOP superfamily. In this procedure, alignments of each of the monomers in one dimer to the monomers in the other dimer were obtained with TM-align. From these alignments, one can count the number of overlapped contacts and calculate the contact overlap ratio, which is the count divided by the smaller number of total interfacial contacts between the two complexes. A minimum contact overlap ratio of 0.3 is required for assigning related interfaces. In these related pairs, the best mTM-score is usually higher than 0.5 between individual proteins from these complexes. On the other hand, two protein–protein complexes have a pair of unrelated protein–protein interfaces, if (i) none of their interacting domains has the same SCOP superfamily assignments, or (ii) the contact overlap ratio is zero and the fraction of aligned interface residues over the shorter length of the two interfaces is \( < 0.15 \). This cutoff was conservatively chosen to tolerate a scenario where a small fraction of interfacial residues are aligned by chance. The second condition incorporates complexes with similar global structures but dissimilar interaction modes. In total, 1128 pairs of biologically related and \( \sim 1.15 \) million pairs of unrelated interfaces were classified. The remaining 3167 interface pairs are ambiguous cases that cannot be confidently classified by analyzing the global structural alignments of individual proteins, and they were therefore discarded for this analysis, although many have significant iTM/IS-scores according to iAlign.

2.4.3 Dimer597 A subset (597) of Dimer1230 is curated by limiting the lengths of individual proteins within 200 amino acids. The subset contains 373 related pairs and 176 875 unrelated pairs.

3 RESULTS

We first tested iAlign on docking models and obtained encouraging results (Supplementary Material). Below, we describe the result of detecting evolutionarily related, structurally similar protein–protein interfaces, and comparison with MM-align and I2I-SiteEngine.

3.1 Predicting biologically related interfaces

We benchmark the performance of iAlign in detecting biologically related protein–protein interfaces from experimental structures. The SCOP superfamily classification and structural similarity are used to decide whether two complexes share biologically related interfaces (Section 2.4.2). Using either the iTM-score or the IS-score as the similarity measure, iAlign estimates a significant \( P < 0.01 \) for all related protein–protein interfaces, except for one case (Fig. 3A). On the other hand, about 0.8/1.8% of unrelated pairs have a significant iTMIS-score with an estimated \( P < 0.01 \). The cumulative fraction of unrelated interface pairs according to their SCOP classification is the observed, or ‘ideal’, \( P\)-value. Ideally, the observed \( P\)-value should match the \( P\)-value estimated by iAlign from Equation (5). In the regime above \( 1 \times 10^{-5}, \) the \( P\)-values according to our statistical model are in good agreement with the observed \( P\)-values. In particular, corresponding to estimated \( P\)-values of 0.01, 0.001 and 0.0001 for the iTM-score, the observed \( P\)-values are 0.009, 0.0009 and 0.0001, respectively. This excellent agreement is a bit surprising, because we did not use the SCOP classification as the RB for deriving our statistical models. In the regime below \( 1 \times 10^{-5}, \) however, there is some separation between the estimated and the observed \( P\)-values. This is largely attributed to the existence of structurally similar yet evolutionarily unrelated pairs, such as four-helix bundles. Overall, the result shows that iAlign identifies significant similarity between related interfaces, and that the estimates of statistical significance are reasonable. The results are consistent between homodimers and...
employed, interface alignments by iAlign always provide better score and coverage, respectively. Moreover, regardless of the similarity measure employed, interface alignments by iAlign always provide better discrimination than global alignments of complex structures by MM-align, with the difference being most notable in the highly confident regime relevant to practical applications. At 0.01 EPQ, for example, dTM-score produces 0.60 coverage, which is considerably lower than 0.79/0.76 coverage by P-value of iTM/IS-score calculated with iAlign. Heavily weighting the interface region gives MM-align slightly better performance, but the improvement is small, for instance, ~4% increase in coverage at 0.01 EPQ.

The difference between interface alignment and global alignment of complexes is not surprising because interface similarity is not a priori equivalent to global similarity, though the two are correlated. As shown in Supplementary Figure S6A, more than half of complex pairs with significant interface similarity at a P < 0.01 have an insignificant dTM-score < 0.4. Nevertheless, highly significant interface similarity often leads to highly significant global similarity. The number of complexes with dissimilar global structures drops dramatically when a significant P-value threshold < 1 × 10^{-5} is employed, though there are exceptions with one example provided in Supplementary Figure S7A. On the other hand, as we mentioned in Section 1, a high dTM-score does not guarantee interface similarity. For example, among dimer pairs with dTM-score > 0.6, ~15% of them have dissimilar interfaces with an insignificant iAlign P < 0.01 (Supplementary Fig. S6B). These are complexes sharing very similar global folds but dissimilar interaction modes. Such false positives are significantly reduced through interface similarity evaluation, which is the main reason for the better performance of iAlign than MM-align. One example is shown in Supplementary Figure S7B.

### 3.2 Comparison with I2I-SiteEngine

The performance of iAlign is compared with a previously published interface alignment method in I2I-SiteEngine (Shulman-Peleg et al., 2004). The I2I-SiteEngine demands relatively large computing resources. To reduce computing costs, we selected a subset (Dimer597) of Dimer1517 by limiting the length of the individual protein chains to 200 residues. I2I-SiteEngine reports two scores for each of top 10 alignments, the Match-score (M-score) and the Total-score (T-score), the former normalized by the Total-score of query compared with itself. For a pair of interfaces, we conducted two runs with each interface as the query, and then selected the best M-score or T-score among all top alignments.

As shown in Figure 4A, the coverage-precision (also known as recall-precision) curve evaluates the accuracy of two methods. The coverage is the same as above, and the precision is defined as the fraction of true positives (both pairs of complexes have evolutionarily related, structurally similar interfaces) among all positives predicted. Clearly, iAlign has substantially higher accuracy than the I2I-SiteEngine. At 80% precision, iAlign can identify 85% of true positives according to the P-value of the IS-score, whereas I2I-SiteEngine can only identify 21/47% of true positives with the M/T-score, respectively. At a higher precision of 90%, iAlign has a coverage value of 75%, about five/two times the coverage of 15/39% by the M/T-score.

Regarding the requisite computer time, iAlign is about two orders of magnitude faster than I2I-SiteEngine (Fig. 4B). We collected the total computing time statistics for iAlign, and statistics of total and essential computing time for I2I-SiteEngine. Note that the total computing time does not include the time for generating input tiles, which one needs to construct only once. The essential computing
We show that the interfacial TM-score (iTM-score) can be applied to I2I-SiteEngine. Medians are represented by thick horizontal black lines. The middle plot represents the runtime of essential computing as reported by I2I-SiteEngine, and the plots represent the overall runtime of iAlign and I2I-SiteEngine, and the costs of iAlign (blue) and I2I-SiteEngine (green). The left and right box precision curves from tests on the set Dimer597. (2264)

For measuring the similarity between monomeric protein structures, this interesting phenomenon unfortunately creates a challenge for predicting biological relationships from protein structure. In this study, we demonstrate that comparison of protein–protein interfaces may be utilized for differentiating biological relationships. This works because protein–protein interactions are important for their stability and/or function, leading to the conservation of a specific interaction mode during the course of evolution. However, we emphasize that interface comparison is not a replacement for, but rather is complementary to, structural comparison of individual proteins. This point can be illustrated with one example. Suppose protein A and its homolog A′ interacts with two different proteins B and C, respectively, and the protein–protein interfaces of complexes A/B and A′/C are dissimilar. An interface alignment of A/B and A′/C gives an insignificant score, whereas an alignment between A and A′ yields a significant score. While the score of interface similarity is still informative, indicating that A/B and A′/C exhibit different interaction modes, the homologous relationship between A and A′ is detected only through the alignment of individual protein structures. Therefore, one expects that a combination of both types of structural comparisons should provide a more comprehensive description of protein–protein relationships than using either comparison metric alone. Toward this direction, considerable efforts have been invested recently in classifying protein–protein interfaces (Kim et al., 2006; Mintz et al., 2005; Shoemaker et al., 2006; Tuncbag et al., 2008).

Since two biologically unrelated protein complexes may display similar interfaces (Tsai et al., 1996), the question arises in how similar two protein–protein interfaces have to be for a reliable prediction on their biological relationship. With the P-value estimation provided by iAlign, one can quickly set a suitable P-value.
threshold. When scanning a large library of 10,000 interfaces, which is the same order of magnitude as the size of the current PDB, one expects to see one hit from unrelated interfaces at a P-value of $1 \times 10^{-10}$. In order to reduce the number of false positives to $<0.1$, one needs to set a highly significant P-value threshold of $<1 \times 10^{-5}$ for predicting biologically related interfaces from a library the size of the PDB.

In principle, protein–protein interfaces can be geometrically similar without following a particular sequence order. While iAlign provides an option for permitting non-sequential alignment, we found that the sequential alignments yield much fewer hits across SCOP superfamilies than non-sequential alignments (Gao and Skolnick, unpublished data). Non-sequential alignments detect both biologically related and unrelated pairs of complexes that have structurally similar interfaces. Restriction to sequential alignments significantly increases the likelihood that an evolutionary relationship (e.g. as assessed by SCOP) is detected.

Despite many efforts toward understanding the nature of protein–protein interactions, numerous issues remain unresolved. Some important questions include: What is the repertoire of protein–protein interactions modes that Nature employs? How complete is the structural space of protein–protein interaction modes in the current PDB? How can we use structural information to accurately predict protein-protein interactions? How can we design a protein–protein interface for a desired function? We expect that iAlign will be a useful tool for addressing these outstanding questions and for developing applications involving structural comparison of protein–protein interfaces.

ACKNOWLEDGEMENTS

We thank Dr Shashi B. Pandit for stimulating discussions and Dr Bartosz Iliowski for computing support.

Funding: National Institutes of Health (grant GM-48835).

Conflict of Interest: none declared.

REFERENCES


