We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

9

Prediction of Protein Interaction Sites Using Mimotope Analysis

Jian Huang, Beibei Ru and Ping Dai
School of Life Science and Technology
University of Electronic Science and Technology of China
China

1. Introduction

Biological functions depend on all kinds of interaction networks; life is a miracle of all types of molecular interactions. Among them, proteins interacting with proteins, nucleic acids and small compounds play a central role (Barabasi & Oltvai, 2004; Przulj, 2011; Vidal, et al., 2011). To guide protein engineering studies for better enzymes, antibodies and drugs, structural and functional characterization of protein interaction sites at the residue or atom level is of great help. Experimental approaches such as X-ray diffraction of protein complex can define structural binding sites at the atomic level (Bickerton, et al., 2011; Higurashi, et al., 2009); mutagenesis and binding test are capable of identifying functional binding sites at the residue or group level (Moreira, et al., 2007; Peng, et al., 2011). However, these means are costly, time-consuming and sometimes technically difficult or even impossible. Moreover, they are not always applicable on a large scale. As a result, computer tools for the prediction of protein interaction sites have been increasingly popular for complementing experimental techniques (Fernández-Recio, 2011; Wass, et al., 2011).

The existing methods for the prediction of protein interaction sites can be grouped into three categories based on the main input data used. The first category consists of methods using protein sequence as the only input (Ofran & Rost, 2007; Res, et al., 2005). Methods in the second category such as molecular docking and simulation solely use structure data as input (Kozakov, et al., 2010; Mashiach, et al., 2010). Methods of the third category make use of a mimotope motif or a set of mimotope sequences together with protein sequence or structure as input (Huang, et al., 2011).

In this chapter, we review methods of the third category, focusing on their current statuses, discussing challenges and providing suggestions to advance this field.

2. Mapping protein-protein interaction sites using mimotope analysis

Mimotopes are peptides mimicking protein interaction sites; they are initially acquired from chemical synthesis (Geysen, et al., 1986). High-throughput obtainment of mimotopes has achieved since phage display and other surface display technologies became available (Smith, 1985; Smith & Petrenko, 1997). Taking phage display as an example, random DNA

sequences can be inserted into genes coding for coat proteins of bacteriophage to make combinatorial libraries. As shown in Figure 1, the combinatorial library can be incubated and selected with an immobilized protein, termed as the target. The natural partner of the target is called as the template. Phages without affinity to the target are washed away with buffer. Then, bound phages are eluted with the target, the template or stronger buffer only. The bound phages are further amplified by infecting bacteria to form a secondary library, which is then used for the next round of incubating, washing, eluting and amplifying. After several rounds of such processes which are well known as biopanning, phage clones are picked randomly from the isolation of bound phages and sequenced. The affinities of these phage clones or corresponding peptides to the target are measured by surface plasmon resonance, enzyme-linked immunosorbent assay or other binding assays. The foreign inserts which enable corresponding phage clones to bind the target competitively with a template are considered as mimotopes.

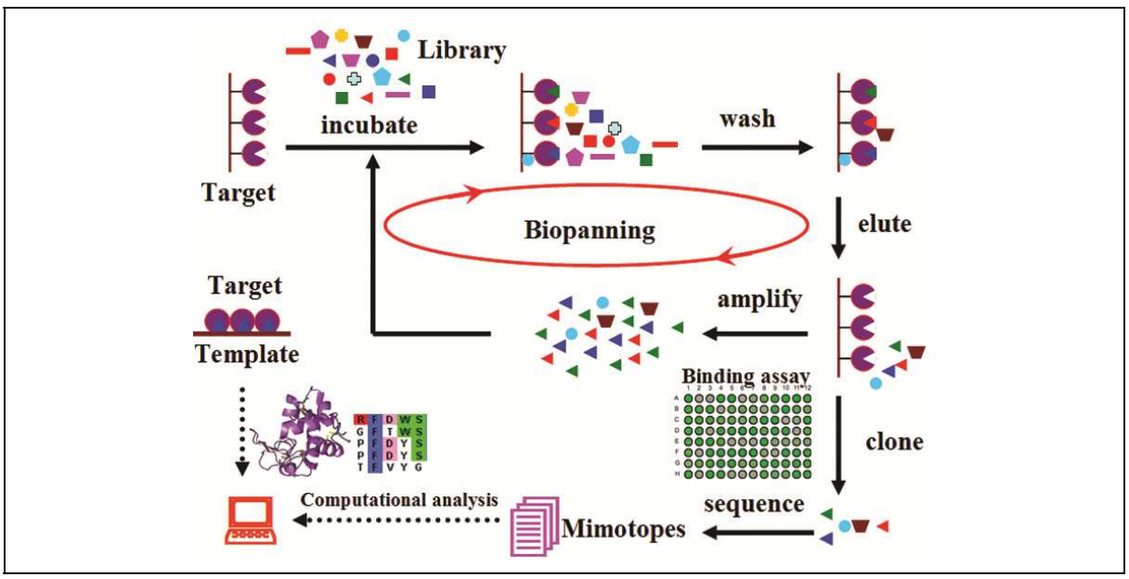


Fig. 1. Schematic view of in vitro phage display and mimotope analysis.

As described above, a set of mimotopes can be readily obtained via phage display. They are capable of binding to the target and blocking the interaction between the target and the template. Therefore, it implies that the information of protein interaction sites is encoded in mimotopes and can be predicted by decoding mimotopes. It is only natural to suppose that the mimotopes are similar to the binding site on the template at the sequential or structural level. Indeed, all approaches to prediction of protein interaction sites based on mimotopes depend on either the sequence or the structure of the template. Thus, the existing methods can be divided into the following two groups:

2.1 Methods based on template sequence

Various methods based on template sequence are summarized in Figure 2. In brief, a set of mimotopes are aligned with the corresponding template to find out the similar region in sequence, which is thought to be at least a part of the target-binding site on the template

protein. Sometimes, sequences of paralogs or orthologs of the template are also aligned to help the identification of the protein interaction site. In some studies, consensus sequences or motifs are derived from the blocks of mimotope alignments. Then, consensus sequences are aligned to the template sequence; motifs are scanned along the template sequence. And the template segments similar to consensus sequences or matching the motifs are considered to be a part of the protein interaction sites. If the template itself is not determined, local alignment search with each mimotope or the consensus sequence against the protein database would help to predict reasonable candidates of template and its binding sites. The template and corresponding interaction sites can also be predicted through pattern search with mimotope motifs against the protein database.

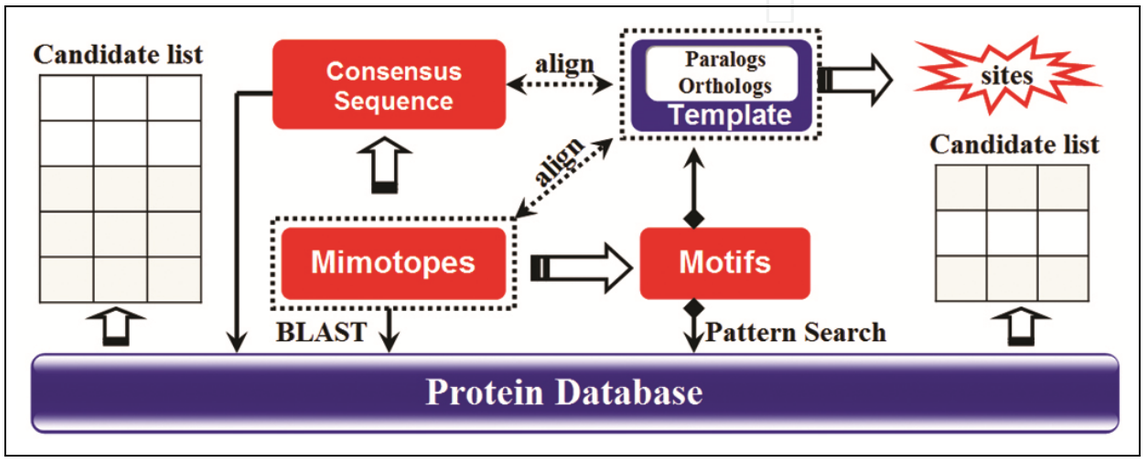


Fig. 2. Flow chart of methods based on template sequence.

As shown in Figure 2, methods based on template sequence involve several steps and tasks such as aligning sequence, inferring consensus sequence or motifs, searching local alignment or motif against the protein database. Among them, sequence alignment is undoubtedly the most important one. Methods based on template sequence can be fulfilled with visual inspection, general-purpose programs and tools specially designed for mimotope analysis.

2.1.1 Manual sequence analysis with visual inspection

Some mimotopes are very similar or even identical to some part of the template sequence every now and then, indicating the segment involving in binding the target protein. In this situation, the protein binding site can be easily depicted through aligning mimotope and template sequence manually by visual inspection. A 6mer random library was screened with the monoclonal antibody GDO5 raised against the Hantaan virus glycoprotein G2. After three rounds of panning, the mimotope obtained had the sequence LEYPWH, which was very similar to the template sequence 94YEYPWH99, implying the site where GDO5 bound (Fack, et al., 1997). The Ph.D.-7 random phage library was panned using the anti-SEB monoclonal antibody ab53981 (Urushibata, et al., 2010). Among the mimotopes obtained, SPDELHK was almost identical to 8PDELHK13 of the staphylococcal enterotoxin B. The ab53981 binding site was thus located. Four anti-HBsAg

monoclonal antibodies, namely H5, H35, H53 and H166 were characterized using phage display (Chen, et al., 1996). Manually aligning 16 H166-binding mimotopes with the HBsAg sequences from subtype adw2, ayw2 and ayw3, Chen et al found that most mimotopes have the CRTC or CKTC subsequences by visual inspection, which were identical to the segments from 121 to 124 of the HBsAg. The epitope recognized by H166 was thus indicated. Nonetheless, sequence alignment programs are necessary when there are a lot of sequences to be aligned or the similarity between the mimotope and template sequence is not obvious.

2.1.2 General-purpose sequence analysis tools

General-purpose tools for sequence alignment, local alignment and pattern search have been widely used in the prediction of protein interaction sites based on mimotope and template sequences. As we described above, Chen et al identified the H166-binding site by visual inspection. However, the software GENEWORK was used in the left three cases (Chen, et al., 1996). Significant matches were found by manual analysis on the dot-matrix diagrams produced by GENEWORK. For example, ARARCEHRSGLSL as one part of an H35-selected mimotope was aligned to 166ASARFSWLSL175 of the HBsAg sequence from subtype ayw3, locating the H35-binding site (Chen, et al., 1996). MDM2-binding peptides were obtained using mRNA display (Shiheido, et al., 2011); the peptides were aligned using the ClustalW program (Larkin, et al., 2007). Compared with the sequence of P53, a similar segment 17-28 was found be the MDM2-P53 interaction site (Shiheido, et al., 2011). A monoclonal antibody against the West Nile virus capsid protein was generated and designated as 6D3 (E. C. Sun, et al., 2011). A 12mer peptide library was screened with 6D3 to produce a set of mimotopes. Alignment revealed a consensus segment KKPGGPG, which was same to the subsequence 3-9 of West Nile virus capsid protein. A monoclonal antibody 2A10G6 was raised against the heat-inactivated dengue virus and used to screen the Ph.D.-12 random phage library. Alignment of mimotopes revealed a consensus FFDRTWP, which corresponded well with 98DRGW101 located at the tip of the fusion loop of E protein of dengue virus (Deng, et al., 2011). In the studies of Sun and Deng, MegAlign software within the Lasergene suite was used to align the orthologs of the template. In the study of Urushibata, the ClustalW program was used to align the paralogs of the template (Urushibata, et al., 2010). These studies showed that orthologs or paralogs were helpful for locating the binding sites.

Unlike the interaction between an antigen and corresponding antibody, a protein may have quite a few partners sometimes or its natural partner may be unknown. In these situations, the sequence alignment between mimotope and template cannot be done directly. However, a local alignment search against the protein sequence database helps to identify candidate templates and binding sites. The AC3 protein of geminiviruses was characterized using phage display (Pasumarthy, et al., 2011). Each AC3-specific peptide sequence obtained was then searched for local alignment against the Arabidopsis non-redundant protein database at NCBI through BLASTP program adjusted for short sequence (Mount, 2007). Proteins from a few metabolic pathways were identified as putative AC3-interacting proteins. For example, YALKHLPESTIP was very similar with 704YALKHIRES712 of the Hua Enhancer 1 (HEN1). Thus HEN1 might interact with the AC3 protein around 704YALKHIRES712

(Pasumarthy, et al., 2011). LipL32, the major outer membrane protein of pathogenic Leptospira, was panned against the Ph.D.-7 phage library. For each mimotope obtained, a BLASTP search against the protein database was performed. Quite a few proteins expressed on the surface of target cells of pathogenic Leptospira were suggested to interact with LipL32. For example, the mimotope HLPPNHT is similar with the sequence PLPPEHT of Collagen XX, indicating LipL32 might bind to Collagen XX at that site (Chaemchuen, et al., 2011). The strategy of mimotope blast against protein databases has also been used to deduce small molecule binding sites and drug targets (Chen, et al., 2006; Takakusagi, et al., 2010; Takami, et al., 2011), infer proteins involving in cell interactions (Kanki, et al., 2011; Zhao, et al., 2010).

Besides tools for local alignment search, pattern search against the protein database has also been used to find possible templates and their target-binding sites. SdrC is important in the interactions between Staphylococcus aureus and its host. However, the host ligand interacting with SdrC was not previously identified. The Ph.D.-12 phage library was screened with SdrC and eight phage clones displayed significantly higher affinity to SdrC. These clones were sequenced, and an alignment revealed the consensus sequence HHIHHFH. It was then used for a pattern search against the human protein database allowing for zero, one and two residue mismatches. The results showed that human neurexin 1β, 2β, 3β and a T-type voltage-dependent calcium channel might be the host ligands interacting with SdrC. Among them, the subsequence 10-16 of human neurexin 1β was identical to the consensus sequence, which implied SdrC might bind to human neurexin 1β at the site 10HHIHHFH16 (Perosa, et al., 2010). Autoantibodies against centromere associated protein A (CENP-A) were purified from sera of eight systemic sclerosis patients with the immunodominant epitope of CENP-A (Ap17–30). These antibodies were used to screen a phage library. The binding phage clones were sequenced, and the inserted peptides were aligned with MULTALIN (Corpet, 1988) to derive antigenic motifs. Human proteins containing such motifs were searched in the SwissProt Protein Sequence Database using the ScanProsite tool. Taking the PTPxxGPxxR motif as an example, 20PTPTPGPSRR29 of human CENP-A was certainly found. However, 53PTPAPGPGRR62 of human Forkhead box protein E3 (FOXE3) also matched the motif, indicating those autoantibodies could interact with FOXE3 at the site around 53-62. Indeed, the peptide 53-62 of FOXE3 was confirmed to behave similarly in binding and inhibition assays with anti-Ap17–30 IgG (Barbu, et al., 2010).

2.1.3 Specially designed sequence analysis tools

Even very recently, general-purpose tools for sequence alignment, local alignment and pattern search remain popular in the study of mapping protein interaction sites based on mimotopes. One reason for this is that these tools are freely, stably and conveniently available. However, these general-purpose tools have their limits. For example, most of them are not good at aligning a very short sequence (mimotope) to quite a long sequence (template). Furthermore, they are less efficient to deduce conformational binding sites, which are made of segments far away in primary sequence but close on the surface of template structure. Specially designed tools are thus needed for sequence analyses of mimotopes and templates.

To delineate conformational binding sites on protein, the program FINDMAP was proposed (Mumey, et al., 2003). FINDMAP allowed any permutations (e.g. inversion) of the mimotope sequence to align its template sequence. Furthermore, gaps even large gaps were permitted in both mimotope and template sequences. Such alignment was proven to be NP-complete and a branch-and-bound algorithm was used to solve the problem in practice. As FINDMAP could deal with only one mimotope each time, an improved version called EPIMAP was introduced later. It was capable of aligning each mimotope to the template, producing a set of top-scoring alignments, selecting the most mutually compatible alignments and filtering out spurious alignments (Mumey, et al., 2006). MimAlign was a meta-method. It combined results from four multiple sequence alignments of the template and its mimotopes (Moreau, et al., 2006). In the RELIC suite, there were quite a few tools specially designed for analysis on mimotopes (Mandava, et al., 2004). For example, MOTIF1 and MOTIF2 were designed to identify weak sequence motifs within short peptide sequence; MATCH, FASTAcon and FASTAskan were designed for optimal sequence alignments between mimotopes and its template. Although the RELIC suite focused on the interaction between small molecule and protein, its sequence tools were often used in the analysis of protein-protein interaction sites. For instance, MMACHC-binding peptides were aligned to MMADHC with tools in RELIC and five MMACHC-binding sites on the protein MMADHC were predicted (Plesa, et al., 2011). Mouse monoclonal antibodies against the predominant VSGs LiTat 1.3 and LiTat 1.5 of T.b. gambiense were used to screen PhD.-12 and Ph.D.-C7C phage libraries. Epitopes were identified by sequence alignment performed manually and with RELIC suite (Van Nieuwenhove, et al., 2011). For example, ALLPFKDHLPYP selected with the monoclonal antibody H12H3 against VSG LiTat 1.5 was aligned to 269AQAVYKDHDPDQ280 of VSG LiTat 1.5. The following experiment did show that the binding of H12H3 to synthetic ALLPFKDHLPYP was inhibited by human African trypanosomiasis sera. Regretfully, all the special tools described here are now hard or impossible to access.

2.1.4 Methods based on template sequence: challenges and suggestions

Methods based on template sequence are of their advantages. For example, they can be used in any condition because no structural information is required during prediction. Even if the template sequence is not given, local alignment or pattern search against protein databases may fulfil the task of inferring possible templates and protein interaction sites. However, to evaluate the results of sequence alignment, local alignment search and pattern search is still a great challenge.

Two formulae have been proposed to compute the frequency of finding similar sequences in two random sequences with different lengths (Chen, et al., 1996). One formula is for a single sequence match; another is for nearby matches within a pair of two sequences. This was a good attempt to evaluate if a continuous or discontinuous match was significant or just by chance. Chen et al assumed that 20 different residues were with equal probability at each position of the two sequences. However, it is not true in real case. To be more reasonable, we suggest using the residue frequency of the corresponding phage library for mimotopes and the actual frequency for template with long sequence. For short or unknown template, use the amino acids frequency of SwissProt.

Although the BLAST program has its statistical means to evaluate a match, they are not fit for short peptides such as mimotopes. In the study of Pasumarthy et al, a lot of matches were found. Among them, the mimotope FPKAFHHHKIY was found to be similar with 317HKIY310 of the Retinoblastoma like protein (pRBR) with an E-value of 1250 and pRBR was known to be an AC3-interacting protein. The mimotope DAMIMKKHWHRF was found to be similar with 164MIMK167 of the Geminivirus Rep interacting kinase 1 (GRIK1) with an E-value of 517 and GRIK1 did interact with the AC1 protein. Thus, they used the E-value 1250 as the threshold to filter the blast results. The candidate list was further shortened with one of the following conditions: (1) at least two hits from the same or different peptides; (2) with E-value less than 517 (Pasumarthy, et al., 2011). In another study, only a tri-peptide or longer sequence match was considered (Kanki, et al., 2011). It seems that the evaluation of sequence matches found by sequence alignment, local alignment search and pattern search are rather arbitrary. As the standard is different case by case, the results from these tools are more like a kind of indication rather than a formal prediction. The results can be confirmed only when more background information is available. Results from sequence alignment, local alignment search and pattern search are same in nature: similarity matches between mimotope and a protein sequence. Thus, a general statistics model or method that evaluates the similarity match reasonably is urgently needed.

As described previously, methods based on template sequence have succeeded in many cases. However, it is more frequent that mimotopes show little similarities to the template, especially when the interaction sites are conformational. Thus methods based on template sequence often fail too. TSOL18 is a host-protective oncosphere antigen of Taenia solium, which is a cestode parasite causing cysticercosis in humans and pigs. The Ph.D.-12 phage library was screened with the anti-TSOL18 monoclonal antibody 17E1. The mimotopes were aligned to the TSOL18 protein sequence using ClustalW software. No significant match was found (Guo, et al., 2010). Intact oocytes surrounded by canine zona pellucida proteins were used to identify peptide sequences from phage display libraries that could recognize and bind to zona pellucida proteins (Samoylova, et al., 2010). The selection of a 12mer library resulted in identification of four sequences with the common NNXXPIL motif discovered by the MOTIF2 program in the RELIC suite. Among them, NNQSPILKLSIH was synthesized and immunized in dogs. The anti-NNQSPILKLSIH antibodies did bind to the acrosomal region of the canine sperm cell. However, BLAST search did not result in identification of homologies to known sperm proteins or other mammalian proteins. Thus, to predict protein interaction sites that are discontinuous using only sequences of mimotope and template is a great challenge. Though the FINDMAP program is a good attempt on this, it is still far from satisfactory. As the entry number of the PDB database increases exponentially, more and more protein structures become available to be used in the prediction of protein interaction sites based on mimotope analysis (Rose, et al., 2011).

2.2 Methods based on template structure

When sequence similarities are not found, it is very likely that mimotopes resemble a special region on the surface of template rather than a linear segment of template sequence. The

prediction of protein interaction sites based on mimotope sequences and corresponding template structure is actually to identify and evaluate surface regions on the template that are similar to mimotopes.

2.2.1 Algorithms, programs and web servers

In 1995, Pizzi et al described the first method that predicted discontinuous antibody binding site based on mimotopes and the antigen structure (Pizzi, et al., 1995). Since then, quite a few algorithms, programs and web servers have been published by different teams around the world. All these methods can be divided into four groups. The first one is the motif-based group, which align a motif or consensus sequence to template structure. This group includes 3DEX (Schreiber, et al., 2005), MIMOX (Huang, et al., 2006) and the MimCons section of MIMOP program (Moreau, et al., 2006). The second group includes Mapitope (Bublil, et al., 2006; Bublil, et al., 2007; Enshell-Seijffers, et al., 2003; Tarnovitski, et al., 2006) and its derivatives (Denisov, et al., 2009; Denisova, et al., 2008; Denisova, et al., 2009; Denisova, et al., 2010). It can be called the pairs-based group because amino acid pairs on the template surface are considered to be simulated by amino acid pairs in the mimotope sequence. The third one is the patch-based group, which evaluates similarities between surface patches on template and mimotopes. SiteLight (Halperin, et al., 2003) and EpiSearch (Negi & Braun, 2009) belong to this group. The fourth is the graph-based group, which aligns a set of query peptides to a graph representing the template surface. Pepsurf (Mayrose, Shlomi, et al., 2007) and Pep-3D-Search (Huang, et al., 2008) belong to this group. To improve the performances of existing programs, hybrid methods such as MimoPro (Chen, et al., 2011) and meta-servers such as Pepitope (Mayrose, Penn, et al., 2007) were also proposed.

As tools mentioned above have been reviewed in detail recently (Huang, et al., 2011), here we only introduce LocaPep, a tool proposed very recently (Pacios, et al., 2011). For each mimotope, this program firstly scans the template surface to select seeds. Then it searches residues adjacent to each seed to form a cluster. For each residue in a cluster, its total score is the weighted sum of the area, exposure, contacts and distance score. At last, the final consensus cluster is calculated to form the binding site predicted. LocaPep is written with Fortran90 independent of any specific library and runs in command line mode. Its source code, manual and binaries are available at http://atenea.montes.upm.es.

2.2.2 Benchmarking tools of the trade

As described above, quite a few methods based on template structure are available for the phage display community to predict protein interaction sites. All these methods have succeeded in some case studies. These test cases were either compiled from published papers or from special databases such as the ASPD database (Valuev, et al., 2002) and the MimoDB database (Huang, et al., 2012; Ru, et al., 2010). However, no systematic evaluations were done when these methods were published. This is due to a relative lack of the type of data where the target-template complex is solved and the relevant mimotope data is available simultaneously.

As the protein structure and mimotope data increase rapidly (Huang, et al., 2012; Rose, et al., 2011), now it becomes possible to make benchmarks for the trade to evaluate its tools at a

larger scale. Sun et al compiled a benchmark from the PDB database (Sun, et al., 2011) and the MimoDB database. It included 47 test cases in which 18 cases were with structures of the antigen-antibody complexes and 29 cases had structures of other protein-protein complexes. They further kept only one test case for each complex with the same template, which made a representative dataset with 30 test cases. Five popular tools, i.e. Mapitope, PepSurf, Pepitope, EpiSearch and Pep-3D-Search, were evaluated with the benchmark and the representative dataset. The results showed that performances of these tools were better than random predictions. However, their overall performances were still not satisfactory. Most tools were good at some cases but failed with other cases.

Our group has also compiled a benchmark called MimoBench (Huang, et al., 2012). It can be freely accessed from http://immunet.cn/mimodb/mimobench.php. Currently, MimoBench has 23, 23 and 27 sets of data for antibody–antigen complex, receptor–ligand complex and other protein-protein complex respectively. Using this benchmark, we have performed a preliminary evaluation on Mapitope, Episearch and MimoPro by their default parameters. Our results showed that performances of these tools were poor in many cases. However, they made quite accurate predictions in some cases. Taking the AUC value 0.8 as a cutoff, the three benchmarked tools succeeded in overlapping but different cases, which suggested that these tools complemented each other. Thus, it is recommended to use several tools together in the prediction of protein–protein interaction sites based on mimotopes.

2.2.3 Methods based on template structure: Challenges and suggestions

Methods based on template structure are capable of predicting the conformational sites of protein-protein interactions. However, the existing tools are not robust enough. Sun et al reported that many test cases in their benchmark dataset could not be applied to the five tools they evaluated due to software limitations (Sun, et al., 2011). We met the same problem when we compared Mapitope, Episearch and MimoPro using MimoBench. For example, four test cases were excluded from benchmarking because these tools did not work on the template with two or more chains. Another 10 cases were dropped because MimoPro returned no results for unknown reason (Huang, et al., 2012). Hence, tools in the future should be more robust. Furthermore, they should also be more convenient to access. It is hoped that web sites of these tools are stable and easy to access. No login is required. Thus, they can be utilized more conveniently whether they are standalone tools or web servers.

As described in the previous section, performances of the existing tools based on template structure are poor in many cases. To improve their performances is one of the greatest challenges in this field. We have suggested that the poor performance might partly due to information loss and noise inclusion during the experimental and computational process (Huang, et al., 2009). Considering the two points in mind, the accuracy of deciphering protein interaction sites using mimotopes might be improved. We will discuss on this issue in the following section.

2.3 Data cleaning tools

Due to the limitation of experiments, the biopanning results are noisy. They are usually a mixture of mimotopes (desired signal) and target-unrelated peptides (unwanted noise).

Target-unrelated peptides (TUPs) can be divided into two categories. One is called selection-related TUP. They appear in the biopanning results because they are selected by contaminants or other components of the screening system rather than the target (Menendez & Scott, 2005; Vodnik, et al., 2011). Propagation-related TUP makes another category (Brammer, et al., 2008; Derda, et al., 2011; Thomas, et al., 2010). They sneak into the output of biopanning because they have a higher infection rate or faster secretion rate. Phages with growth advantage can be not only noise but also decrease the library diversity and lead to a loss of useful mimotopes. Simulations and experiments showed that subtle differences in growth rate yielded drastic differences in clone abundances after rounds of amplifications (Derda, et al., 2011). Thus, propagation-related TUP may even dominate the biopanning results. As TUPs are peptides unrelated to the target, they undoubtedly interfere with the prediction of protein interaction sites based on mimotopes if a TUP is taken as a mimotope. Changing experimental conditions and improving experimental methods can decrease TUPs. For example, increasing the stringency of panning may reduce TUPs; subtractive procedures may decrease selection-related TUPs; amplification in isolated compartment can mitigate the growth advantage of propagation-related TUPs (Derda, et al., 2010). However, TUPs cannot be eradicated experimentally. To exclude TUPs from the biopanning with computational tools has become an alternative and more convenient choice.

2.3.1 Data cleaning tool based on information theory

Based on the information theory, the program INFO in the RELIC suite (Mandava, et al., 2004) calculates information content for each peptide of the panning result. Two input files are required. The first one is a text file with a minimum of 50 peptide sequences from clones randomly selected from a naive library. The second file is the query of users, one or more peptide sequences selected from that same library. INFO first uses AAFREQ to calculate the amino acid frequency distributions at each position of the inserted peptide sequences from the parent library. The probability of random occurrence of any peptide can be calculated by multiplying the probability of each amino acid occurring at each position. The natural logarithm of the probability of a peptide multiplied by -1 is defined as its information content. Obviously, if the query peptide has very high information content, it is less possible to appear in the panning result. If it does occur in the result, it is more likely to be the mimotope selected by specific binding to the target. On the contrary, when a peptide with very low information content is observed in the result, it is less confident of taking it as a mimotope because it may be a propagation-related TUP. The INFO program was also integrated in other tools in the RELIC suite such as MATCH, HETEROalign and FASTAskan. However, all these tools have regretfully been inaccessible for about one year, which makes the RELIC suite now a real relic.

2.3.2 Data cleaning tool based on TUP motif

We have developed a free web server called SAROTUP, which is short for scanner and reporter of target-unrelated peptides (Huang, et al., 2010). It can be used to scan and exclude possible target-unrelated peptides from biopanning result. SAROTUP is based on known TUP motifs and sequences. In the current version, a set of 26 TUP motifs and 27 known TUP

sequences are collected from literature and compiled into the program. Among them, nine sequences are known or highly suspected to be propagation-related TUPs; the left 42 motifs or sequences are for selection-related TUPs, including 14 for albumin binders, six for unrelated antibody binders, five for immunoglobulin Fc region binders, five for streptavidin binders, five for plastic binders, four for bivalent metal ion binders and one for biotin binders, protein A binders and lipid A binders respectively. We had tested SAROTUP before the MimoDB database was constructed. The results showed that: (1) TUPs were often seen and taken as mimotopes; (2) epitope prediction based on mimotopes was greatly interfered if TUPs were used in the analysis; (3) SAROTUP improved performances of epitope mapping based on mimotopes through cleaning the input data; (4) SAROTUP also helped to explain experiment results. However, as a tool based on pattern search, SAROTUP cannot deal with TUPs without known motifs.

2.3.3 Data cleaning tool based on database search

The problem mentioned above was partly solved when the MimoDB database became available (Huang, et al., 2012; Ru, et al., 2010). With a lot of biopanning results and relevant background information collected, this database can be used as a virtual and comprehensive control for experimental biologists. In the MimoDB database version 2.0, a batched peptide search tool can be used for a set of peptides to search against the database. If a peptide has been reported by different groups with different targets, it may be a TUP rather than a mimotope. This is because the chance of obtaining an identical peptide from a library having millions or billions of different peptides with a completely different target is extremely small. If this happens, the peptide obtained may be due to some common factors in the biopanning systems rather than by the target. The MimoBlast tool of the MimoDB database can further find out those peptides not identical but highly similar to the query peptides. Such peptides may also be TUPs. New TUP motifs can be derived from analyzing the result of MimoBlast. With these tools, we studied the peptides in the MimoDB database and claimed confidently that GETRAPL, SILPYPY, LLADTTHHRPWT, TMGFTAPRFPHY, SAHGTSTGVPWP and HLPTSSLFDTTH are TUPs which were not reported before (Huang, et al., 2012).

2.3.4 Data cleaning tools: Challenges and suggestions

Although the data cleaning tools described in this section complement each other, none of them are real classifiers but rather reminders. Without a solid statistical estimation, they can only tell users that a peptide in the result may be a TUP rather than a mimotope. However, as the entries in the MimoDB database increases rapidly, it is now practical to construct various TUP predictors based on machine learning methods. Secondly, the data cleaning procedure was ignored by most existing tools for the prediction of protein interaction sites based on mimotopes. This situation should be changed in the future.

3. Conclusion

Identification of the protein interaction site is very important for basic and applied research. Computational analysis on mimotopes obtained from phage display or other

surface display experiments is a relatively cheap, convenient and efficient strategy to locate a protein interaction site at the segment or residue level. Although used mostly in epitope prediction, this strategy can also be used to other types of protein interaction sites. Insights can be gained by methods based on template sequence, which find sequence similarities between mimotopes and template through sequence alignment, local alignment search and pattern search. Conformational sites can also be mapped by methods based on template structure. However, performances of all existing methods are not satisfactory enough. This is at least partly due to TUPs that crept into the biopanning result. Several tools are available to detect TUPs based on information theory, known TUP motifs or special database. With the rapid accumulation of experimental data and improvement of methods, an evidence-based virtual phage display platform is expected to be established and the performance of predicting protein interaction sites based on mimotopes will substantially be increased.

4. Acknowledgment

This work was supported by the National Natural Science Foundation of China (grant 61071177) and the Scientific Research Foundation of UESTC for Youth (grant JX0769).

5. References

Barabasi, A. L. & Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. Nat Rev Genet, Vol. 5, No. 2, (Feb), pp. 101-113, ISSN 1471-0056

Barbu, E. M.; Ganesh, V. K.; Gurusiddappa, S.; Mackenzie, R. C.; Foster, T. J.; Sudhof, T. C. & Hook, M. (2010). beta-Neurexin is a ligand for the Staphylococcus aureus MSCRAMM SdrC. PLoS Pathog, Vol. 6, No. 1, (Jan), p. e1000726, ISSN 1553-7374

Bickerton, G. R.; Higueruelo, A. P. & Blundell, T. L. (2011). Comprehensive, atomic-level characterization of structurally characterized protein-protein interactions: the PICCOLO database. BMC Bioinformatics, Vol. 12, p. 313, ISSN 1471-2105

Brammer, L. A.; Bolduc, B.; Kass, J. L.; Felice, K. M.; Noren, C. J. & Hall, M. F. (2008). A target-unrelated peptide in an M13 phage display library traced to an advantageous mutation in the gene II ribosome-binding site. Anal Biochem, Vol. 373, No. 1, (Feb 1), pp. 88-98, ISSN 0003-2697

Bublil, E. M.; Yeger-Azuz, S. & Gershoni, J. M. (2006). Computational prediction of the cross-reactive neutralizing epitope corresponding to the monoclonal antibody b12 specific for HIV-1 gp120. FASEB J, Vol. 20, No. 11, (Sep), pp. 1762-1774, ISSN 1530-6860

Bublil, E. M.; Freund, N. T.; Mayrose, I.; Penn, O.; Roitburd-Berman, A.; Rubinstein, N. D.; Pupko, T. & Gershoni, J. M. (2007). Stepwise prediction of conformational discontinuous B-cell epitopes using the Mapitope algorithm. Proteins, Vol. 68, No. 1, (Jul 1), pp. 294-304, ISSN 1097-0134

Chaemchuen, S.; Rungpragayphan, S.; Poovorawan, Y. & Patarakul, K. (2011). Identification of candidate host proteins that interact with LipL32, the major outer membrane protein of pathogenic Leptospira, by random phage display peptide library. Vet Microbiol, Vol. 153, No. 1-2, (Nov 21), pp. 178-185, ISSN 1873-2542

Chen, T.; Cui, J.; Liang, Y.; Xin, X.; Owen Young, D.; Chen, C. & Shen, P. (2006). Identification of human liver mitochondrial aldehyde dehydrogenase as a potential target for microcystin-LR. Toxicology, Vol. 220, No. 1, (Mar 1), pp. 71-80, ISSN 0300-483X

Chen, W. H.; Sun, P. P.; Lu, Y.; Guo, W. W.; Huang, Y. X. & Ma, Z. Q. (2011). MimoPro: a more efficient Web-based tool for epitope prediction using phage display libraries. BMC Bioinformatics, Vol. 12, p. 199, ISSN 1471-2105

Chen, Y. C.; Delbrook, K.; Dealwis, C.; Mimms, L.; Mushahwar, I. K. & Mandecki, W. (1996). Discontinuous epitopes of hepatitis B surface antigen derived from a filamentous phage peptide library. Proc Natl Acad Sci U S A, Vol. 93, No. 5, (Mar 5), pp. 1997-2001, ISSN 0027-8424

Corpet, F. (1988). Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res, Vol. 16, No. 22, (Nov 25), pp. 10881-10890, ISSN 0305-1048

Deng, Y. Q.; Dai, J. X.; Ji, G. H.; Jiang, T.; Wang, H. J.; Yang, H. O.; Tan, W. L.; Liu, R.; Yu, M.; Ge, B. X.; Zhu, Q. Y.; Qin, E. D.; Guo, Y. J. & Qin, C. F. (2011). A broadly flavivirus cross-neutralizing monoclonal antibody that recognizes a novel epitope within the fusion loop of E protein. PLoS One, Vol. 6, No. 1, p. e16059, ISSN 1932-6203

Denisov, D. A.; Denisova, G. F.; Lelic, A.; Loeb, M. B. & Bramson, J. L. (2009). Deciphering epitope specificities within polyserum using affinity selection of random peptides and a novel algorithm based on pattern recognition theory. Mol Immunol, Vol. 46, No. 3, (Jan), pp. 429-436, ISSN 0161-5890

Denisova, G.; Denisov, D.; Evelegh, C.; Weissgram, M.; Beck, J.; Foley, S. R. & Bramson, J. L. (2009). Characterizing complex polysera produced by antigen-specific immunization through the use of affinity-selected mimotopes. PLoS One, Vol. 4, No. 4, p. e5309, ISSN 1932-6203

Denisova, G. F.; Denisov, D. A.; Yeung, J.; Loeb, M. B.; Diamond, M. S. & Bramson, J. L. (2008). A novel computer algorithm improves antibody epitope prediction using affinity-selected mimotopes: a case study using monoclonal antibodies against the West Nile virus E protein. Mol Immunol, Vol. 46, No. 1, (Nov), pp. 125-134, ISSN 0161-5890

Denisova, G. F.; Denisov, D. A. & Bramson, J. L. (2010). Applying bioinformatics for antibody epitope prediction using affinity-selected mimotopes - relevance for vaccine design. Immunome Res, Vol. 6 Suppl 2, p. S6, ISSN 1745-7580

Derda, R.; Tang, S. K. & Whitesides, G. M. (2010). Uniform amplification of phage with different growth characteristics in individual compartments consisting of monodisperse droplets. Angew Chem Int Ed Engl, Vol. 49, No. 31, (Jul 19), pp. 5301-5304, ISSN 1521-3773

Derda, R.; Tang, S. K.; Li, S. C.; Ng, S.; Matochko, W. & Jafari, M. R. (2011). Diversity of phage-displayed libraries of peptides during panning and amplification. Molecules, Vol. 16, No. 2, pp. 1776-1803, ISSN 1420-3049

Enshell-Seijffers, D.; Denisov, D.; Groisman, B.; Smelyanski, L.; Meyuhas, R.; Gross, G.; Denisova, G. & Gershoni, J. M. (2003). The mapping and reconstitution of a conformational discontinuous B-cell epitope of HIV-1. J Mol Biol, Vol. 334, No. 1, (Nov 14), pp. 87-101, ISSN 0022-2836

Fack, F.; Hugle-Dorr, B.; Song, D.; Queitsch, I.; Petersen, G. & Bautz, E. K. (1997). Epitope mapping by phage display: random versus gene-fragment libraries. J Immunol Methods, Vol. 206, No. 1-2, (Aug 7), pp. 43-52, ISSN 0022-1759

Fernández-Recio, J. (2011). Prediction of protein binding sites and hot spots. WIREs Computational Molecular Science, Vol. 1, No. 5, pp. 680-698, ISSN 1759-0884

Geysen, H. M.; Rodda, S. J. & Mason, T. J. (1986). A priori delineation of a peptide which mimics a discontinuous antigenic determinant. Mol Immunol, Vol. 23, No. 7, (Jul), pp. 709-715, ISSN 0161-5890

Guo, A.; Cai, X.; Jia, W.; Liu, B.; Zhang, S.; Wang, P.; Yan, H. & Luo, X. (2010). Mapping of Taenia solium TSOL18 antigenic epitopes by phage display library. Parasitol Res, Vol. 106, No. 5, (Apr), pp. 1151-1157, ISSN 1432-1955

Halperin, I.; Wolfson, H. & Nussinov, R. (2003). SiteLight: binding-site prediction using phage display libraries. Protein Sci, Vol. 12, No. 7, (Jul), pp. 1344-1359, ISSN 0961-8368

Higurashi, M.; Ishida, T. & Kinoshita, K. (2009). PiSite: a database of protein interaction sites using multiple binding states in the PDB. Nucleic Acids Res, Vol. 37, No. Database issue, (Jan), pp. D360-364, ISSN 1362-4962

Huang, J.; Gutteridge, A.; Honda, W. & Kanehisa, M. (2006). MIMOX: a web tool for phage display based epitope mapping. BMC Bioinformatics, Vol. 7, p. 451, ISSN 1471-2105

Huang, J.; Xia, M.; Lin, H. & Guo, F.B. Information loss and noise inclusion risk in mimotope based epitope mapping, The 3rd International Conference on Bioinformatics and Biomedical Engineering(iCBBE2009), pp. 1-3, Beijing, 2009

Huang, J.; Ru, B.; Li, S.; Lin, H. & Guo, F. B. (2010). SAROTUP: scanner and reporter of target-unrelated peptides. J Biomed Biotechnol, Vol. 2010, p. 101932, ISSN 1110-7251

Huang, J.; Ru, B. & Dai, P. (2011). Bioinformatics resources and tools for phage display. Molecules, Vol. 16, No. 1, pp. 694-709, ISSN 1420-3049

Huang, J.; Ru, B.; Zhu, P.; Nie, F.; Yang, J.; Wang, X.; Dai, P.; Lin, H.; Guo, F. B. & Rao, N. (2012). MimoDB 2.0: a mimotope database and beyond. Nucleic Acids Res, Vol. 40, No. D1, pp. D271-D277, ISSN 1362-4962

Huang, Y. X.; Bao, Y. L.; Guo, S. Y.; Wang, Y.; Zhou, C. G. & Li, Y. X. (2008). Pep-3D-Search: a method for B-cell epitope prediction based on mimotope analysis. BMC Bioinformatics, Vol. 9, p. 538, ISSN 1471-2105

Kanki, S.; Jaalouk, D. E.; Lee, S.; Yu, A. Y.; Gannon, J. & Lee, R. T. (2011). Identification of targeting peptides for ischemic myocardium by in vivo phage display. J Mol Cell Cardiol, Vol. 50, No. 5, (May), pp. 841-848, ISSN 1095-8584

Kozakov, D.; Hall, D. R.; Beglov, D.; Brenke, R.; Comeau, S. R.; Shen, Y.; Li, K.; Zheng, J.; Vakili, P.; Paschalidis, I. & Vajda, S. (2010). Achieving reliability and high accuracy in automated protein docking: ClusPro, PIPER, SDU, and stability analysis in CAPRI rounds 13-19. Proteins, Vol. 78, No. 15, (Nov 15), pp. 3124-3130, ISSN 1097-0134

Larkin, M. A.; Blackshields, G.; Brown, N. P.; Chenna, R.; McGettigan, P. A.; McWilliam, H.; Valentin, F.; Wallace, I. M.; Wilm, A.; Lopez, R.; Thompson, J. D.; Gibson, T. J. &

Higgins, D. G. (2007). Clustal W and Clustal X version 2.0. Bioinformatics, Vol. 23, No. 21, (Nov 1), pp. 2947-2948, ISSN 1367-4811

Mandava, S.; Makowski, L.; Devarapalli, S.; Uzubell, J. & Rodi, D. J. (2004). RELIC--a bioinformatics server for combinatorial peptide analysis and identification of protein-ligand interaction sites. Proteomics, Vol. 4, No. 5, (May), pp. 1439-1460, ISSN 1615-9853

Mashiach, E.; Schneidman-Duhovny, D.; Peri, A.; Shavit, Y.; Nussinov, R. & Wolfson, H. J. (2010). An integrated suite of fast docking algorithms. Proteins, Vol. 78, No. 15, (Nov 15), pp. 3197-3204, ISSN 1097-0134

Mayrose, I.; Penn, O.; Erez, E.; Rubinstein, N. D.; Shlomi, T.; Freund, N. T.; Bublil, E. M.; Ruppin, E.; Sharan, R.; Gershoni, J. M.; Martz, E. & Pupko, T. (2007). Pepitope: epitope mapping from affinity-selected peptides. Bioinformatics, Vol. 23, No. 23, (Dec 1), pp. 3244-3246, ISSN 1367-4803

Mayrose, I.; Shlomi, T.; Rubinstein, N. D.; Gershoni, J. M.; Ruppin, E.; Sharan, R. & Pupko, T. (2007). Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. Nucleic Acids Res, Vol. 35, No. 1, pp. 69-78, ISSN 0305-1048

Menendez, A. & Scott, J. K. (2005). The nature of target-unrelated peptides recovered in the screening of phage-displayed random peptide libraries with antibodies. Anal Biochem, Vol. 336, No. 2, (Jan 15), pp. 145-157, ISSN 0003-2697

Moreau, V.; Granier, C.; Villard, S.; Laune, D. & Molina, F. (2006). Discontinuous epitope prediction based on mimotope analysis. Bioinformatics, Vol. 22, No. 9, (May 1), pp. 1088-1095, ISSN 1367-4803

Moreira, I. S.; Fernandes, P. A. & Ramos, M. J. (2007). Hot spots--a review of the protein-protein interface determinant amino-acid residues. Proteins, Vol. 68, No. 4, (Sep 1), pp. 803-812, ISSN 1097-0134

Mount, D. W. (2007). Using the Basic Local Alignment Search Tool (BLAST). CSH Protoc, Vol. 2007, p. pdb top17, ISSN 1559-6095

Mumey, B.; Ohler, N.; Angel, T.; Jesaitis, A. & Dratz, E. (2006). Filtering Epitope Alignments to Improve Protein Surface Prediction, In: Frontiers of High Performance Computing and Networking – ISPA 2006 Workshops, G. Min, B. Di Martino, L. Yang, M. Guo and G. Ruenger, (Eds.), 648-657, Springer,

Mumey, B. M.; Bailey, B. W.; Kirkpatrick, B.; Jesaitis, A. J.; Angel, T. & Dratz, E. A. (2003). A new method for mapping discontinuous antibody epitopes to reveal structural features of proteins. J Comput Biol, Vol. 10, No. 3-4, pp. 555-567, ISSN 1066-5277

Negi, S. S. & Braun, W. (2009). Automated detection of conformational epitopes using phage display Peptide sequences. Bioinform Biol Insights, Vol. 3, pp. 71-81, ISSN 1177-9322

Ofran, Y. & Rost, B. (2007). ISIS: interaction sites identified from sequence. Bioinformatics, Vol. 23, No. 2, (Jan 15), pp. e13-16, ISSN 1367-4811

Pacios, L. F.; Tordesillas, L.; Palacin, A.; Sanchez-Monge, R.; Salcedo, G. & Diaz-Perales, A. (2011). LocaPep: localization of epitopes on protein surfaces using peptides from phage display libraries. J Chem Inf Model, Vol. 51, No. 6, (Jun 27), pp. 1465-1473, ISSN 1549-960X

Pasumarthy, K. K.; Mukherjee, S. K. & Choudhury, N. R. (2011). The presence of tomato leaf curl Kerala virus AC3 protein enhances viral DNA replication and modulates virus induced gene-silencing mechanism in tomato plants. Virol J, Vol. 8, p. 178, ISSN 1743-422X

Peng, L.; Oganesyan, V.; Damschroder, M. M.; Wu, H. & Dall'acqua, W. F. (2011). Structural and Functional Characterization of an Agonistic Anti-Human EphA2 Monoclonal Antibody. J Mol Biol, Vol. 413, No. 2, (Oct 21), pp. 390-405, ISSN 1089-8638

Perosa, F.; Vicenti, C.; Racanelli, V.; Leone, P.; Valentini, G. & Dammacco, F. (2010). The immunodominant epitope of centromere-associated protein A displays homology with the transcription factor forkhead box E3 (FOXE3). Clin Immunol, Vol. 137, No. 1, (Oct), pp. 60-73, ISSN 1521-7035

Pizzi, E.; Cortese, R. & Tramontano, A. (1995). Mapping epitopes on protein surfaces. Biopolymers, Vol. 36, No. 5, (Nov), pp. 675-680, ISSN 0006-3525

Plesa, M.; Kim, J.; Paquette, S. G.; Gagnon, H.; Ng-Thow-Hing, C.; Gibbs, B. F.; Hancock, M. A.; Rosenblatt, D. S. & Coulton, J. W. (2011). Interaction between MMACHC and MMADHC, two human proteins participating in intracellular vitamin B metabolism. Mol Genet Metab, Vol. 102, No. 2, (Feb), pp. 139-148, ISSN 1096-7206

Przulj, N. (2011). Protein-protein interactions: making sense of networks via graph-theoretic modeling. Bioessays, Vol. 33, No. 2, (Feb), pp. 115-123, ISSN 1521-1878

Res, I.; Mihalek, I. & Lichtarge, O. (2005). An evolution based classifier for prediction of protein interfaces without using protein structures. Bioinformatics, Vol. 21, No. 10, (May 15), pp. 2496-2501, ISSN 1367-4803

Rose, P. W.; Beran, B.; Bi, C.; Bluhm, W. F.; Dimitropoulos, D.; Goodsell, D. S.; Prlic, A.; Quesada, M.; Quinn, G. B.; Westbrook, J. D.; Young, J.; Yukich, B.; Zardecki, C.; Berman, H. M. & Bourne, P. E. (2011). The RCSB Protein Data Bank: redesigned web site and web services. Nucleic Acids Res, Vol. 39, No. Database issue, (Jan), pp. D392-401, ISSN 1362-4962

Ru, B.; Huang, J.; Dai, P.; Li, S.; Xia, Z.; Ding, H.; Lin, H.; Guo, F. & Wang, X. (2010). MimoDB: a New Repository for Mimotope Data Derived from Phage Display Technology. Molecules, Vol. 15, No. 11, pp. 8279-8288, ISSN 1420-3049

Samoylova, T. I.; Cox, N. R.; Cochran, A. M.; Samoylov, A. M.; Griffin, B. & Baker, H. J. (2010). ZP-binding peptides identified via phage display stimulate production of sperm antibodies in dogs. Anim Reprod Sci, Vol. 120, No. 1-4, (Jul), pp. 151-157, ISSN 1873-2232

Schreiber, A.; Humbert, M.; Benz, A. & Dietrich, U. (2005). 3D-Epitope-Explorer (3DEX): localization of conformational epitopes within three-dimensional structures of proteins. J Comput Chem, Vol. 26, No. 9, (Jul 15), pp. 879-887, ISSN 0192-8651

Shiheido, H.; Takashima, H.; Doi, N. & Yanagawa, H. (2011). mRNA display selection of an optimized MDM2-binding peptide that potently inhibits MDM2-p53 interaction. PLoS One, Vol. 6, No. 3, p. e17898, ISSN 1932-6203

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science, Vol. 228, No. 4705, (Jun 14), pp. 1315-1317, ISSN 0036-8075

Smith, G. P. & Petrenko, V. A. (1997). Phage Display. Chem Rev, Vol. 97, No. 2, (Apr 1), pp. 391-410, ISSN 1520-6890

Sun, E. C.; Zhao, J.; Yang, T.; Liu, N. H.; Geng, H. W.; Qin, Y. L.; Wang, L. F.; Bu, Z. G.; Yang, Y. H.; Lunt, R. A. & Wu, D. L. (2011). Identification of a conserved JEV serocomplex B-cell epitope by screening a phage-display peptide library with a mAb generated against West Nile virus capsid protein. Virol J, Vol. 8, p. 100, ISSN 1743-422X

Sun, P.; Chen, W.; Huang, Y.; Wang, H.; Ma, Z. & Lv, Y. (2011). Epitope prediction based on random peptide library screening: benchmark dataset and prediction tools evaluation. Molecules, Vol. 16, No. 6, pp. 4971-4993, ISSN 1420-3049

Takakusagi, Y.; Takakusagi, K.; Sugawara, F. & Sakaguchi, K. (2010). Use of phage display technology for the determination of the targets for small-molecule therapeutics. Expert Opinion on Drug Discovery, Vol. 5, No. 4, pp. 361-389, ISSN 1746-0441

Takami, M.; Takakusagi, Y.; Kuramochi, K.; Tsukuda, S.; Aoki, S.; Morohashi, K.; Ohta, K.; Kobayashi, S.; Sakaguchi, K. & Sugawara, F. (2011). A screening of a library of T7 phage-displayed peptide identifies E2F-4 as an etoposide-binding protein. Molecules, Vol. 16, No. 5, pp. 4278-4294, ISSN 1420-3049

Tarnovitski, N.; Matthews, L. J.; Sui, J.; Gershoni, J. M. & Marasco, W. A. (2006). Mapping a neutralizing epitope on the SARS coronavirus spike protein: computational prediction based on affinity-selected peptides. J Mol Biol, Vol. 359, No. 1, (May 26), pp. 190-201, ISSN 0022-2836

Thomas, W. D.; Golomb, M. & Smith, G. P. (2010). Corruption of phage display libraries by target-unrelated clones: diagnosis and countermeasures. Anal Biochem, Vol. 407, No. 2, (Dec 15), pp. 237-240, ISSN 1096-0309

Urushibata, Y.; Itoh, K.; Ohshima, M. & Seto, Y. (2010). Generation of Fab fragment-like molecular recognition proteins against staphylococcal enterotoxin B by phage display technology. Clin Vaccine Immunol, Vol. 17, No. 11, (Nov), pp. 1708-1717, ISSN 1556-679X

Valuev, V. P.; Afonnikov, D. A.; Ponomarenko, M. P.; Milanesi, L. & Kolchanov, N. A. (2002). ASPD (Artificially Selected Proteins/Peptides Database): a database of proteins and peptides evolved in vitro. Nucleic Acids Res, Vol. 30, No. 1, (Jan 1), pp. 200-202, ISSN 1362-4962

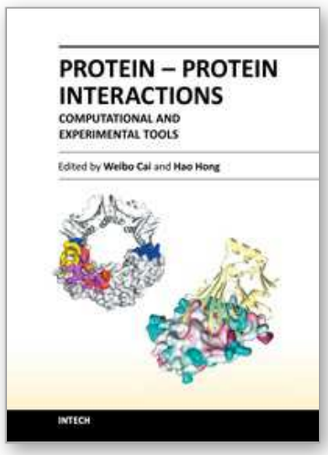
Van Nieuwenhove, L. C.; Roge, S.; Balharbi, F.; Dieltjens, T.; Laurent, T.; Guisez, Y.; Buscher, P. & Lejon, V. (2011). Identification of peptide mimotopes of Trypanosoma brucei gambiense variant surface glycoproteins. PLoS Negl Trop Dis, Vol. 5, No. 6, (Jun), p. e1189, ISSN 1935-2735

Vidal, M.; Cusick, M. E. & Barabasi, A. L. (2011). Interactome networks and human disease. Cell, Vol. 144, No. 6, (Mar 18), pp. 986-998, ISSN 1097-4172

Vodnik, M.; Zager, U.; Strukelj, B. & Lunder, M. (2011). Phage display: selecting straws instead of a needle from a haystack. Molecules, Vol. 16, No. 1, pp. 790-817, ISSN 1420-3049

Wass, M. N.; David, A. & Sternberg, M. J. (2011). Challenges for the prediction of macromolecular interactions. Curr Opin Struct Biol, Vol. 21, No. 3, (Jun), pp. 382-390, ISSN 1879-033X

Zhao, S.; Zhao, W. & Ma, L. (2010). Novel peptide ligands that bind specifically to mouse embryonic stem cells. Peptides, Vol. 31, No. 11, (Nov), pp. 2027-2034, ISSN 1873-5169



Protein-Protein Interactions - Computational and Experimental Tools

Edited by Dr. Weibo Cai

ISBN 978-953-51-0397-4
Hard cover, 472 pages
Publisher InTech
Published online 30, March, 2012
Published in print edition March, 2012

Proteins are indispensable players in virtually all biological events. The functions of proteins are coordinated through intricate regulatory networks of transient protein-protein interactions (PPIs). To predict and/or study PPIs, a wide variety of techniques have been developed over the last several decades. Many in vitro and in vivo assays have been implemented to explore the mechanism of these ubiquitous interactions. However, despite significant advances in these experimental approaches, many limitations exist such as false-positives/false-negatives, difficulty in obtaining crystal structures of proteins, challenges in the detection of transient PPI, among others. To overcome these limitations, many computational approaches have been developed which are becoming increasingly widely used to facilitate the investigation of PPIs. This book has gathered an ensemble of experts in the field, in 22 chapters, which have been broadly categorized into Computational Approaches, Experimental Approaches, and Others.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Jian Huang, Beibei Ru and Ping Dai (2012). Prediction of Protein Interaction Sites Using Mimotope Analysis, Protein-Protein Interactions - Computational and Experimental Tools, Dr. Weibo Cai (Ed.), ISBN: 978-953-51-0397-4, InTech, Available from: http://www.intechopen.com/books/protein-protein-interactions-computational-and-experimental-tools/prediction-of-protein-interaction-sites-using-mimotope-analysis

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen