*Published online 3 November 2011* *Nucleic Acids Research, 2012, Vol. 40, Database issue* ***D271–D277***
*doi:10.1093/nar/gkr922*

# MimoDB 2.0: a mimotope database and beyond

**Jian Huang*, Beibei Ru, Ping Zhu, Fulei Nie, Jun Yang, Xuyang Wang, Ping Dai, Hao Lin, Feng-Biao Guo and Nini Rao**

Key Laboratory for Neuroinformation of Ministry of Education, School of Life Science and Technology, University of Electronic Science and Technology of China, No.4, 2nd Section, North Jianshe Road, Chengdu, Sichuan 610054, China

Received August 24, 2011; Revised October 7, 2011; Accepted October 8, 2011

## ABSTRACT

**Mimotopes are peptides with affinities to given targets. They are readily obtained through biopanning against combinatorial peptide libraries constructed by phage display and other display technologies such as mRNA display, ribosome display, bacterial display and yeast display. Mimotopes have been used to infer the protein interaction sites and networks; they are also ideal candidates for developing new diagnostics, therapeutics and vaccines. However, such valuable peptides are not collected in the central data resources such as UniProt and NCBI GenPept due to their 'unnatural' short sequences. The MimoDB database is an information portal to biopanning results of random libraries. In version 2.0, it has 15 633 peptides collected from 849 papers and grouped into 1818 sets. Besides the core data on panning experiments and their results, broad background information on target, template, library and structure is included. An accompanied benchmark has also been compiled for bioinformaticians to develop and evaluate their new models, algorithms and programs. In addition, the MimoDB database provides tools for simple and advanced searches, structure visualization, BLAST and alignment view on the fly. The experimental biologists can easily use the database as a virtual control to exclude possible target-unrelated peptides. The MimoDB database is freely available at http://immunet.cn/mimodb.**

## INTRODUCTION

The term mimotope was first coined by Mario Geysen in 1986 (1). It was originally used to describe peptides mimicking epitope. Before long, this concept was extended to refer peptide mimic of all types of binding sites. Mimotopes can be readily obtained from random peptide libraries through biopanning with all kinds of substances ranging from metal ions to drugs, nucleic acids to proteins, cells to organs. Usually, the substance used to screen combinatorial peptide library is termed target. The natural partner of target is called template. As the mimic of binding site, mimotope analysis has been widely used in mapping epitopes, identifying drug target and inferring protein interaction networks (2–4). Furthermore, mimotope has also shown its potential in the development of new diagnostics, therapeutics and vaccines (5–8). In addition, special affinities mediated by mimotopes to various semiconductors and other materials have shown very encouraging promise in new material and new energy studies (9,10). Gathering information on mimotopes into a special database therefore deserves.

When the concept of mimotope formed more than 25 years ago, it was not easy to get them. The construction of combinatorial peptide libraries was not only the starting point but also the rate-limiting step to acquire mimotopes. However, the situation soon changed when biological libraries came into this field. Unlike chemical libraries, biological libraries are constructed at the nucleic acid level. Peptides are translated by biological systems rather than chemical synthesis. George Smith introduced the earliest biological library, i.e. phage-displayed random peptide library (11). Since then, combinatorial peptide libraries constructed by other display technologies such as mRNA display, ribosome display, bacterial display and yeast display have emerged one after another. All these biological libraries have made it cheap, efficient and convenient to obtain mimotopes. High-throughput screening of combinatorial peptide libraries has led the amount of peptide sequences from biopanning results increasing quickly. However, these peptide sequences were not collected in the central data resources such as UniProt (12) and NCBI GenPept (13) due to their 'unnatural' short sequences. Scattered in the full-text papers, it was hard to access and utilize information on these import

*To whom correspondence should be addressed. Tel: +86 28 83208232; Fax: +86 28 83208238; Email: hj@uestc.edu.cn

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

peptides. Hence, it is urgent to have a special database in this area.

In 2000, a special database called ASPD was constructed by Valuev *et al*. (14). The ASPD database is short for Artificial Selected Proteins/Peptides Database, which focuses on peptides and proteins acquired via *in vitro* evolution, mainly by phage display. It has 4345 sequences grouped as 195 sets of mimotopes curated from 112 papers. Regretfully, however, this database has not been updated since August 2001. Thus, a more comprehensive, more frequently updated database for mimotope and closely related scientific community is still needed.

In 2010, we released the MimoDB database version 1.0, which had 10 716 peptides grouped into 1229 sets (15). These peptides were extracted from biopanning results of phage-displayed random peptide libraries reported in 571 papers. Since the MimoDB database came online, it has been updated four times to the current version 2.0. The data entries have increased substantially; an accompanied benchmark has been created; the web interface has been improved; several new data-mining tools have been developed. In this article, we describe how we construct the MimoDB database version 2.0, its new features and our data-mining works and new findings.

## MATERIALS AND METHODS

### Data collection and organization

The MimoDB database version 2.0 collects peptides selected from combinatorial peptide libraries. The peptides are required to be from random libraries and within 3–40 amino acids long. Peptides selected from mutation libraries and cDNA libraries, e.g. antibody phage display libraries, are excluded. The data collecting strategy and data inclusion standards of the MimoDB database version 2.0 are similar with those of previous version (15). However, significant difference also exists. While the MimoDB database version 1.0 only collects phage display data, the version 2.0 also gathers data from other surface display technologies such as mRNA display, ribosome display, bacterial display and yeast display. In the MimoDB database version 1.0, all papers processed were indexed in PubMed. However, some peer-reviewed papers in other reference databases are also processed in version 2.0. These papers are tracked through 'cited' and 'cited by' search.

The data in the MimoDB database is organized in an experiment-centered style (15). A set of peptides rather than each individual sequence are grouped as an entry if they come from an independent experiment with identical method, condition and parameters. Not only sequences, but also its appearing times (an integer) or frequency (a decimal between zero and one) is recorded. Such information is put in the brackets right behind each sequence if available. Other experiment information, e.g. the panning method, the round of biopanning, the brief experiment process, etc. is also stored. Besides experiment-dependent data extracted directly from the published paper, a lot of background information on target, template, structure and library are taken from closely related papers and external databases such as Uniprot, GenBank, PDB and PDBsum (12,13,16,17).

### Database design and implementation

The infrastructures of the MimoDB database version 1.0 and 2.0 are basically the same (15). In brief, the database has five main tables for mimotope set, target, template, library and complex structure, respectively, in which the table for mimotope set is undoubtedly the core. This design corresponds with the data organization style described previously. There are also two joint tables, which help the database produce dynamic new fields and contents on the front-end.

The MySQL relational database management system is used to store and manage the data. The web interface for entry browse and advanced search are coded in PHP with the support of PEAR packages. The interactive structure viewer for target–template complex or target–mimotope complex is implemented with JmolApplet and PHP. According to the feedback of users, several revisions to the web interface of the MimoDB database have also been made in the version 2.0 to improve user experience.

### New data-mining tools and MimoDB data analyses

Three new data-mining tools are implemented as CGI program with Perl. The first one is for batched peptide search. The second one with the internal name 'MimoBlast' is a tool doing blast optimized for short peptides against the MimoDB database. And the last one is an on-the-fly alignment viewer for blast results. These new tools and other programs such as SAROTUP (18) were used to analyze all peptides in the MimoDB database version 2.0.

### Benchmark construction and preliminary benchmarking

The advanced search tool of the MimoDB database was used to find all mimotope sets that have 3D structure of target–template complex. Corresponding mimotope sets and related background information were then manually grouped, checked and compiled into a benchmark for programs that predict protein interaction site based on mimotopes. A primary evaluation was done with the benchmark on three tools in this filed, i.e. Mapitope, EpiSearch and MimoPro (19–21).

## RESULTS

### Database content and web interface

As shown in Table 1, the content of the MimoDB database version 2.0 has substantially increased compared with version 1.0. Not only entry numbers, but also content in many data fields have been improved. For example, in the MimoDB database version 2.0, the NCBI taxonomy ID has been added to each species which the target or template belongs to. The EC number has also been added and linked to EXPASY ENZYME if the target or template is an enzyme and its Enzyme Commission number is available.

**Table 1.** Comparison of database content between the MimoDB database version 2.0 and 1.0

| | MimoDB version 2.0 | MimoDB version 1.0 |
|---|---|---|
| All peptides | 15 633 | 10 716 |
| Unique peptides | 14 083 | 9805 |
| Mimotope sets | 1818 | 1229 |
| References | 849 | 571 |
| Targets | 1110 | 775 |
| Templates | 360 | 257 |
| Complex structures | 206 | 58 |
| Libraries | 340 | 250 |

According to the feedback from users, the pager system of browse interface has been revised. Users can now go to any page of the summary table or any entry just by inputting the page or entry number and pressing the 'Go' button. All embedded secondary menus in the MimoDB database version 1.0 are replaced by a conspicuous tabular menu bar on the top. Now it is hard to get confused when surfing the web interface of version 2.0 since it has been revised to be more self-evident. Thus, better user experience can be expected with the MimoDB database version 2.0. As required by users, a downloaded page is also added into the web interface. From the version 2.0 on, all data in the MimoDB database can be freely downloaded as EXCEL files or XML files from corresponding compressed archive.

All web interfaces of the MimoDB database version 2.0 have been tested with various browsers such as the Internet Explorer, Mozilla Firefox, Google Chrome, Opera and Safari on Windows, Linux or Mac OS platforms. Although appearances may be a little different, all the tools work normally on all tested browsers and platforms. In all the browsers we tested, the Mozilla Firefox and Internet Explorer give the best user experience. Hence, we recommend users surf the database with one of the two browsers.

### Batched peptide search tool

Peptide sequences and other related information in the MimoDB database can be used as a comprehensive control for biopanning. Experimental scientists can search their peptides against all peptides in the MimoDB database to verify if each peptide has been reported by other groups with different targets. 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 acquired may be due to other common factors in the biopanning systems rather than by the target. Therefore, such peptide may be noise rather than signal users need. In the MimoDB database version 1.0, the work can be done with 'Advanced Search' tool. However, only one peptide can be searched each time, and target name is not explicitly shown in the result table. In the MimoDB database version 2.0, a batched peptide search tool is implemented as a CGI program with Perl. Now, all peptides in FASTA format or as raw sequences from a biopanning experiment can be input or uploaded all at once for search. The target name and mimoset in which the peptide found will be reported in a table if exact matches exist.

### MimoDB blast tool

While the batched peptide search tool can only find identical peptides, the blast tool for the MimoDB database can further find out very similar peptides. As the chance of panning out highly similar peptides from a very large library with various targets is still small, experimental biologists can further utilize the 'MimoBlast' tool to exclude possible target-unrelated peptides (TUP). With the blast results, experimental biologists can be more confident if their sequences are true binders. Guided by the BLAST documentations from NCBI, the MimoDB blast tool is specially optimized for short nearly exact matches by default. Briefly, the expect value cutoff is set to 20 000, word size 2, scoring matrix PAM30, composition-based statistics off and the filters of low-complexity regions off. The sequence input requirement for this tool is same to the batched peptide search tool. The blast result file can be downloaded by users. For each peptide used to blast the database, a separated result file is produced and can be read online. To further facilitate the analysis on the blast result at glance, we also build an on-the-fly alignment viewer, which is powered by JavaScript and Perl codes. Moving mouse cursor over a similar peptide found by blast tool in the result table, a small viewer window will pop up with sequence alignment and related information. Moving mouse cursor to next peptide, the alignment will switch to that peptide automatically. However, the alignment viewer tool functions normally only if pop-up windows from the site of the MimoDB database are allowed.

### New findings from mining MimoDB data with new tools

In the biopanning result, there are not only mimotopes, but also all kinds of target-unrelated peptides (22–25). One category of TUP is called selection-related TUP (22). Although they cannot bind the target site, they can react with contaminants or other components of the screening system and then sneak into the biopanning results. Another category of TUP is called propagation-related TUP (23–25). They creep into and even dominate the output of biopanning because they grow faster in host cells. As researchers are often annoyed by target-unrelated peptides, we have developed SAROTUP, a data-cleaning program based on known TUP motifs. SAROTUP has shown its power in filtering noise and thus improving the performance of computational tools for epitope prediction (18). However, a lot of target-unrelated peptides bear no known motifs. As experimental researchers have used the MimoDB database as control to identify TUP from their panning results, we analyze the data in the MimoDB database version 2.0 against itself with the new tools we developed. The results show that there are 600 peptide sequences appear two or more times. Further analyses indicate that some of them are repeated because they are panned with the same target. Some are panned with different but closely related targets and have a common template, e.g.

**Table 2.** Peptides seen in five or more mimotope sets

| Peptide | Target numbers | Mimoset numbers | SAROTUP[a] | Known or suspected to be TUP before |
|---|---|---|---|---|
| SVSVGMKPSPRP | 42 | 49 | + | Yes |
| LLADTTHHRPWT | 13 | 15 | − | No, new finding |
| HAIYPRH | 13 | 19 | + | Yes |
| LPLTPLP | 11 | 17 | + | Yes |
| KSLSRHDHIHHH | 9 | 14 | + | Yes |
| TMGFTAPRFPHY | 7 | 7 | − | No, new finding |
| SILPYPY | 6 | 8 | − | No, new finding |
| APWHLSSQYSRT | 6 | 7 | + | Yes |
| FHENWPS | 6 | 7 | + | Yes |
| HWGMWSY | 5 | 6 | + | Yes |
| KLWVIPQ | 5 | 6 | + | Yes |
| SAHGTSTGVPWP | 5 | 5 | − | No, new finding |
| HLPTSSLFDTTH | 4 | 6 | + | No, new finding |
| GETRAPL | 4 | 5 | − | No, new finding |

[a]In this column, '+' means known TUP motif is found by SAROTUP, '−' is on the contrary.

**Table 3.** Possible TUP similar with SVSVGMKPSPRP taken from MimoDB blast results

| Peptide | Target numbers | Mimoset numbers | Expect value |
|---|---|---|---|
| SVSVGMNPSPRP | 1 | 1 | 0.14 |
| SVSVGLKPSPRP | 1 | 1 | 0.15 |
| SVSVGMKPSHRP | 1 | 1 | 0.16 |
| SVSVGMKPRPRP | 2 | 2 | 0.17 |
| SVSVGMKPSPRK | 1 | 1 | 0.17 |
| SVSVGKKPSPRP | 1 | 1 | 0.24 |
| SVSGGMKPSPRP | 1 | 1 | 0.26 |
| SVSVGMLPSPRP | 1 | 1 | 0.32 |
| YVYVGMKPSPRP | 1 | 1 | 0.32 |

different monoclonal antibodies against the same antigen. All these types of repeats (456 peptides in total) are considered reasonable and excluded from the following study. The left 144 peptides are scanned by SAROTUP and 35 peptides are found with known TUP motifs. Then we focus on peptides seen in five or more mimotope sets. As shown in Table 2, we suspected with enough confidence that the following six peptides, i.e. LLADTTHHRPWT, TMGFTAPRFPHY, SILPYPY, SAHGTSTGVPWP, HLPTSSLFDTTH and GETRAPL are target-unrelated peptides, which was not reported before. Notably, all the peptides in the Table 2 come from Ph.D.-12 or Ph.D.-7 phage library.

We also blast the sequences in the Table 2 against the MimoDB database 2.0. The results show that there are quite a few sequences which are highly similar with the sequences in the Table 2. Let us take the notorious peptide SVSVGMKPSPRP as an example. As shown in Table 3, the expect value of these short peptide alignments are not exceeded 0.32, indicating these peptides are very likely to be target-unrelated peptides, although most of them are just appear once in one mimotope set and panned out with one target. As these possible TUPs cannot be detected by SAROTUP or batched peptide search tool, the power of the blast tool is shown. Furthermore, new TUP motif might be derived from the sequence block in Table 3, which can make SAROTUP more competent when new motifs are added in its new version.

## MimoBench and preliminary benchmarking results

Mapping protein interaction sites based on mimotopes is a challenging task for theoretical biologists (26). Quite a few methods and tools such as SiteLight, 3DEX, MIMOP, MIMOX, Mapitope, Pepsurf, Pepitope, Pep-3D-Search, Episearch, MimoPro and LocaPep (19–21,27–32) etc. have been developed in recent years. These tools have been tested in a few case studies, however systematical evaluations are absent due to short of benchmarks. In fact, our data has been taken to benchmark available tools right after the MimoDB database version 1.0 published (21,33). From MimoDB 2.0 on, an accompanied benchmark is compiled and scheduled to be update with the database. The benchmark is called MimoBench and can be freely accessed at http://immunet.cn/mimodb/mimobench.php. At present, MimoBench is mainly for development and evaluation of tools that predict protein–protein interaction sites based on mimotopes. It has 23, 23 and 27 sets of data for antibody–antigen complex, receptor–ligand complex and other protein-protein complex, respectively. Using MimoBench, we have performed a preliminary evaluation on Mapitope, Episearch and MimoPro by their default parameters. As all the tools tested could not manage template with multiple chains, four data sets were excluded from benchmarking. For unknown reason, MimoPro returned no results to another 10 data sets. Thus, the three tools were compared on the left 59 sets of data. For each case, the area under the curve (AUC) of each tool is computed as the arithmetic average value of its specificity and sensitivity. As shown in Figure 1, it seems that all the tools perform better with the antibody–antigen cases, and worse with receptor–ligand cases. In many cases, performances of these tools are far from satisfactory, indicating there is still enough space for us to improve these tools. Taken the AUC value 0.8 as a cutoff, the three tools succeed in overlapping but different cases (Figure 2). Therefore, it is hard to say which tool is better but

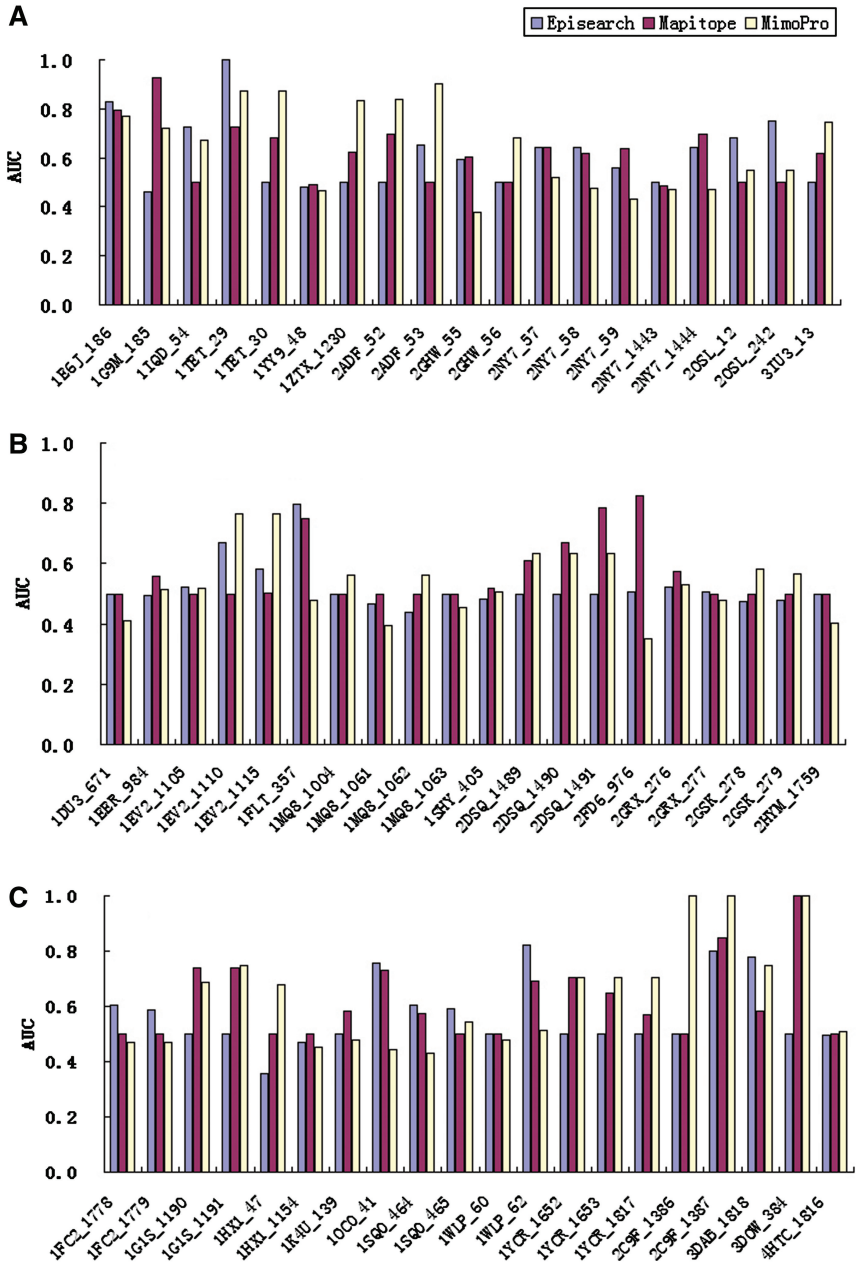


**Figure 1.** Benchmark Mapitope, Episearch and MimoPro with MimoBench. The string under the *X*-axis is the case tested. Each case has the format: PDB ID_Mimotope Set ID, where the left part is the PDB code of corresponding target–template structure, the right part is entry ID of the mimotope set. (**A**) Antibody–antigen group, (**B**) receptor–ligand group and (**C**) other protein–protein interaction group.

rather all the tools complement each other. Hence, it is recommended to use several tools together in predicting the protein–protein interaction site based on mimotopes. However, it is still a tough task to decide which result should be adopted when the results share no overlapping.

## DISCUSSION

### Closely related databases

In many biological fields, there are usually quite a few similar databases compete and complement each other. However, the biopanning field is an exception. As far as we know, there are only three databases available that have some peptides from biopanning results. One is the ASPD database mentioned previously (14). This flat file-based database has ceased updating since 2001. The second database is called PepBank, which collects all types of peptide data rather than focusing on peptides from biopanning results (34). In 2007, the short peptide sequences of the ASPD database was incorporated into PepBank. Nevertheless, the major data of the PepBank database comes from a text-mining program that extracts peptide sequence from MEDLINE abstracts. The third database is our MimoDB database. As a manually curated database focusing on panning results

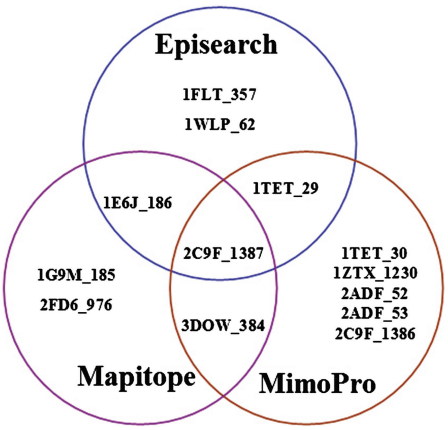


**Figure 2.** Success cases of Mapitope, Episearch and MimoPro with AUC above 0.8. The cases are written in the same format described in Figure 1.

of combinatorial peptide library, it has shown its value for both experimental biologists and computational biologists.

### Future development

Mimotopes can be regarded as kind of peptide aptamers. The latter is typically comprised of a variable peptide region of 8 to 20 amino acids in length displayed by a scaffold protein (35). The major difference is that the so called 'peptide aptamers' are selected by a yeast-two-hybrid system and an interaction trap system rather than affinity selection *in vitro* or *in vivo*. However, considering their similarities, these classical peptide aptamers should be stored in the MimoDB database in the future. Furthermore, the MimoDB database will also be extended to the results from chemical libraries. In addition, the MimoDB database should consider collecting peptides before biopanning, where the target is empty and the round of panning is zero (36). Correspondingly, new tools to analyze the peptides before and after panning will also be part of our future work. We expect the MimoDB database will serve the related scientific community better.

## ACKNOWLEDGEMENTS

The authors are grateful to the editor and the anonymous reviewers for their valuable suggestions and comments, which has led to the improvement of this article.

## FUNDING

The National Natural Science Foundation of China (30600138, 61071177); Scientific Research Foundation of UESTC for Youth (JX0769). Funding for open access charge: National Natural Science Foundation of China (61071177).

*Conflict of interest statement*. None declared.

## REFERENCES

1. Geysen,H.M., Rodda,S.J. and Mason,T.J. (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Mol. Immunol.*, **23**, 709–715.
2. Smith,G.P. and Petrenko,V.A. (1997) Phage display. *Chem. Rev.*, **97**, 391–410.
3. Tong,A.H., Drees,B., Nardelli,G., Bader,G.D., Brannetti,B., Castagnoli,L., Evangelista,M., Ferracuti,S., Nelson,B., Paoluzi,S. *et al.* (2002) A combined experimental and computational strategy to define protein interaction networks for peptide recognition modules. *Science*, **295**, 321–324.
4. Thom,G., Cockroft,A.C., Buchanan,A.G., Candotti,C.J., Cohen,E.S., Lowne,D., Monk,P., Shorrock-Hart,C.P., Jermutus,L. and Minter,R.R. (2006) Probing a protein-protein interaction by in vitro evolution. *Proc. Natl Acad. Sci. USA*, **103**, 7619–7624.
5. Riemer,A.B. and Jensen-Jarolim,E. (2007) Mimotope vaccines: epitope mimics induce anti-cancer antibodies. *Immunol. Lett.*, **113**, 1–5.
6. Hsiung,P.L., Hardy,J., Friedland,S., Soetikno,R., Du,C.B., Wu,A.P., Sahbaie,P., Crawford,J.M., Lowe,A.W., Contag,C.H. *et al.* (2008) Detection of colonic dysplasia in vivo using a targeted heptapeptide and confocal microendoscopy. *Nat. Med.*, **14**, 454–458.
7. Knittelfelder,R., Riemer,A.B. and Jensen-Jarolim,E. (2009) Mimotope vaccination - from allergy to cancer. *Expert Opin. Biol. Ther.*, **9**, 493–506.
8. Macdougall,I.C., Rossert,J., Casadevall,N., Stead,R.B., Duliege,A.M., Froissart,M. and Eckardt,K.U. (2009) A peptide-based erythropoietin-receptor agonist for pure red-cell aplasia. *N. Engl. J. Med.*, **361**, 1848–1855.
9. Lee,Y.J., Yi,H., Kim,W.J., Kang,K., Yun,D.S., Strano,M.S., Ceder,G. and Belcher,A.M. (2009) Fabricating genetically engineered high-power lithium-ion batteries using multiple virus genes. *Science*, **324**, 1051–1055.
10. Nam,Y.S., Magyar,A.P., Lee,D., Kim,J.W., Yun,D.S., Park,H., Pollom,T.S. Jr, Weitz,D.A. and Belcher,A.M. (2010) Biologically templated photocatalytic nanostructures for sustained light-driven water oxidation. *Nat. Nanotechnol.*, **5**, 340–344.
11. Smith,G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315–1317.
12. Consortium,U. (2011) Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.*, **39**, D214–D219.
13. Benson,D.A., Karsch-Mizrachi,I., Lipman,D.J., Ostell,J. and Sayers,E.W. (2010) GenBank. *Nucleic Acids Res.*, **38**, D46–D51.
14. Valuev,V.P., Afonnikov,D.A., Ponomarenko,M.P., Milanesi,L. and Kolchanov,N.A. (2002) ASPD (Artificially Selected Proteins/Peptides Database): a database of proteins and peptides evolved in vitro. *Nucleic Acids Res.*, **30**, 200–202.
15. Ru,B., Huang,J., Dai,P., Li,S., Xia,Z., Ding,H., Lin,H., Guo,F. and Wang,X. (2010) MimoDB: a new repository for mimotope data derived from phage display technology. *Molecules*, **15**, 8279–8288.
16. Laskowski,R.A. (2009) PDBsum new things. *Nucleic Acids Res.*, **37**, D355–D359.
17. 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. *et al.* (2011) The RCSB Protein Data Bank: redesigned web site and web services. *Nucleic Acids Res.*, **39**, D392–D401.
18. Huang,J., Ru,B., Li,S., Lin,H. and Guo,F.B. (2010) SAROTUP: scanner and reporter of target-unrelated peptides. *J. Biomed. Biotechnol.*, **2010**, 101932.
19. Mayrose,I., Penn,O., Erez,E., Rubinstein,N.D., Shlomi,T., Freund,N.T., Bublil,E.M., Ruppin,E., Sharan,R., Gershoni,J.M. *et al.* (2007) Pepitope: epitope mapping from affinity-selected peptides. *Bioinformatics*, **23**, 3244–3246.
20. Negi,S.S. and Braun,W. (2009) Automated detection of conformational epitopes using phage display Peptide sequences. *Bioinform. Biol. Insights*, **3**, 71–81.
21. Chen,W.H., Sun,P.P., Lu,Y., Guo,W.W., Huang,Y.X. and Ma,Z.Q. (2011) MimoPro: a more efficient Web-based tool for epitope prediction using phage display libraries. *BMC Bioinformatics*, **12**, 199.

22. Menendez,A. and Scott,J.K. (2005) The nature of target-unrelated peptides recovered in the screening of phage-displayed random peptide libraries with antibodies. *Anal. Biochem.*, **336**, 145–157.
23. Brammer,L.A., Bolduc,B., Kass,J.L., Felice,K.M., Noren,C.J. and 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.*, **373**, 88–98.
24. Thomas,W.D., Golomb,M. and Smith,G.P. (2010) Corruption of phage display libraries by target-unrelated clones: diagnosis and countermeasures. *Anal. Biochem.*, **407**, 237–240.
25. Vodnik,M., Zager,U., Strukelj,B. and Lunder,M. (2011) Phage display: selecting straws instead of a needle from a haystack. *Molecules*, **16**, 790–817.
26. Huang,J., Ru,B. and Dai,P. (2011) Bioinformatics resources and tools for phage display. *Molecules*, **16**, 694–709.
27. Halperin,I., Wolfson,H. and Nussinov,R. (2003) SiteLight: binding-site prediction using phage display libraries. *Protein Sci.*, **12**, 1344–1359.
28. Schreiber,A., Humbert,M., Benz,A. and Dietrich,U. (2005) 3D-Epitope-Explorer (3DEX): localization of conformational epitopes within three-dimensional structures of proteins. *J. Comput. Chem.*, **26**, 879–887.
29. Huang,J., Gutteridge,A., Honda,W. and Kanehisa,M. (2006) MIMOX: a web tool for phage display based epitope mapping. *BMC Bioinformatics*, **7**, 451.
30. Mayrose,I., Shlomi,T., Rubinstein,N.D., Gershoni,J.M., Ruppin,E., Sharan,R. and Pupko,T. (2007) Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. *Nucleic Acids Res.*, **35**, 69–78.
31. Huang,Y.X., Bao,Y.L., Guo,S.Y., Wang,Y., Zhou,C.G. and Li,Y.X. (2008) Pep-3D-Search: a method for B-cell epitope prediction based on mimotope analysis. *BMC Bioinformatics*, **9**, 538.
32. Pacios,L.F., Tordesillas,L., Palacin,A., Sanchez-Monge,R., Salcedo,G. and Diaz-Perales,A. (2011) LocaPep: localization of epitopes on protein surfaces using peptides from phage display libraries. *J. Chem. Inf. Model.*, **51**, 1465–1473.
33. Sun,P., Chen,W., Huang,Y., Wang,H., Ma,Z. and Lv,Y. (2011) Epitope prediction based on random peptide library screening: benchmark dataset and prediction tools evaluation. *Molecules*, **16**, 4971–4993.
34. Shtatland,T., Guettler,D., Kossodo,M., Pivovarov,M. and Weissleder,R. (2007) PepBank–a database of peptides based on sequence text mining and public peptide data sources. *BMC Bioinformatics*, **8**, 280.
35. Borghouts,C., Kunz,C. and Groner,B. (2008) Peptide aptamer libraries. *Comb. Chem. High Throughput Screen.*, **11**, 135–145.
36. Derda,R., Tang,S.K., Li,S.C., Ng,S., Matochko,W. and Jafari,M.R. (2011) Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules*, **16**, 1776–1803.