PROTEINS: Structure, Function, and Bioinformatics 68:294–304 (2007)

Stepwise Prediction of Conformational Discontinuous B-Cell Epitopes Using the Mapitope Algorithm

Erez M. Bublil,[†] Natalia Tarnovitski Freund,[†] Itay Mayrose, Osnat Penn, Anna Roitburd-Berman, Nimrod D. Rubinstein, Tal Pupko, and Jonathan M. Gershoni[*]
Department of Cell Research and Immunology, Tel Aviv University, Tel-Aviv, Israel

ABSTRACT **Mapping the epitope of an antibody is of great interest, since it contributes much to our understanding of the mechanisms of molecular recognition and provides the basis for rational vaccine design. Here we present Mapitope, a computer algorithm for epitope mapping. The algorithm input is a set of affinity isolated peptides obtained by screening phage display peptide-libraries with the antibody of interest. The output is usually 1–3 epitope candidates on the surface of the atomic structure of the antigen. We have systematically tested the performance of Mapitope by assessing the effect of the algorithm parameters on the final prediction. Thus, we have examined the effect of the statistical threshold (*ST*) parameter, relating to the frequency distribution and enrichment of amino acid pairs from the isolated peptides and the *D* (distance) and *E* (exposure) parameters which relate to the physical parameters of the antigen. Two model systems were analyzed in which the antibody of interest had previously been co-crystallized with the antigen and thus the epitope is a given. The Mapitope algorithm successfully predicted the epitopes in both models. Accordingly, we formulated a stepwise paradigm for the prediction of discontinuous conformational epitopes using peptides obtained from screening phage display libraries. We applied this paradigm to successfully predict the epitope of the Trastuzumab antibody on the surface of the Her-2/neu receptor in a third model system. Proteins 2007; 68:294–304.** © 2007 Wiley-Liss, Inc.

Key words: antibody; combinatorial phage display peptide library; computational algorithm; epitope mapping; vaccine

INTRODUCTION

Antibodies bind their corresponding antigens with exquisite specificity and high affinity, and as such are pivotal participants in the immune response against pathogens, allowing effective discrimination between self and nonself. In view of this, antibodies have been exploited as critical components for a variety of applications such as the development of immuno-diagnostics and therapeutics as well as highly effective research tools.[1,2] Being able to backtrack from an antibody of interest to its corresponding epitope is not only an academic challenge but a capability that could be extremely useful for the rational design of "epitope-based" vaccines.[2a]

The most direct means to discover an epitope is to co-crystallize the antibody bound to its antigen. This, however, is rarely achieved, as there are no more than 150 unique antibody:antigen co-crystals described in the Protein Data Bank (PDB) database.[3] Consequently, a number of alternative empirical approaches to epitope mapping have been developed and used with various levels of success. These include, for example, computerized docking, in which the structures of the antibody of interest and its antigen have been solved separately, and docked *in silico* onto the surface of one another.[4–6] Saturating mutagenesis, in which one attempts to correlate the effect of a given mutation on binding, is also often used (e.g. Benjamin and Perdue[7]). Alternatively, a combinatorial approach has been developed in which phage displayed random peptides are screened against an antibody of interest and information about the epitope is derived from the panel of affinity purified peptides obtained.[8–11] The peptides are interpreted to mimic the genuine epitope and as such may be of use in defining it.

In the event that the panel of peptides has a common motif which corresponds well to a linear sequence within the antigen, mapping the epitope becomes self evident. Often, however, the isolated peptides have no obvious similarity to the actual epitope and therefore a major task is to be able to correlate the peptides to the epitope they are taken to represent. In this case, at least two

Abbreviations: AAP, amino acid pair; EGFR, epidermal growth factor receptor; gp120, glycoprotein 120; HEL, hen egg lysozyme; Her-2, human EGF receptor 2; HIV, human immunodeficiency virus; mAb, monoclonal antibody; PDB, protein data bank; SSP, statistically significant pair; *ST*, statistical threshold; TAU, Tel-Aviv University.

Grant sponsors: Israel Science Foundation; Yeshaia Horvitz Association; Israeli Ministry of Science and Technology; Edmond J. Safra program in bioinformatics.

[†]Erez M. Bublil and Natalia Tarnovitski Freund contributed equally to this work.

*Correspondence to: Jonathan M. Gershoni, Ph.D., Department of Cell Research and Immunology, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel. E-mail: gershoni@tauex.tau.ac.il

Received 7 August 2006; Revised 7 December 2006; Accepted 20 December 2006

Published online 10 April 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21387

approaches are available for epitope-mapping. In the first approach, the "best binder" amongst the panel, either directly selected from the random library or otherwise optimized for preferred binding by targeted mutagenesis, is regarded as the ultimate epitope mimetic.[12–14] In such a case, one seeks to find correspondence of this optimal peptide within the three-dimensional (3D) structure of the antigen.[12]

We have proposed a second approach in which one seeks to use the data obtained from a pool of phage display peptides collectively, to discover the genuine epitope on the surface of the antigen (in contrast to singling-out and focusing on a specific mimetic).[15] According to the "peptide panel" approach, each peptide in the panel is assumed to contain elements or fractions of the epitope, and collectively the panel represents the epitope in its entirety. Thus, within limits, increasing the panel's size should concomitantly improve the signal-to-noise ratio.

Practically, we implemented this second approach by developing a computer algorithm originally described by Enshell-Seijffers et al.[15] The algorithm uses a set of peptides obtained from screening phage display peptide-libraries and the 3D structure of the antigen as data input. The output of the algorithm is the predicted epitope.

In this study we provide a detailed description of "Mapitope", an improved version of our previously published algorithm,[15] in which we have incorporated a number of new options and features (see Material and Methods). Moreover, Mapitope is user-friendly and very robust, and its running time was considerably improved compared to the original version. We have systematically studied the performance of Mapitope using two model systems in which the antibody of interest had been co-crystallized along with its antigen and thus the molecular composition of its epitope is a given. Through systematic analyses of each of the parameters of the algorithm, we have formulated a stepwise approach for the prediction of discontinuous conformational epitopes which is provided herewith, and is illustrated in detail for three antibody:antigen co-crystals.

MATERIALS AND METHODS

The Mapitope Algorithm

Mapitope is based on the original notion that the panel of peptides derived from a random peptide library collectively represents the epitope of the monoclonal antibody (mAb) which they bind.[15] The computer program that implements the Mapitope algorithm was written in C++.

The underlying hypothesis of Mapitope is that the simplest meaningful fragment of an epitope is an "amino acid pair" (AAP) of residues that lie within the footprint of the epitope. These AAPs are proposed to be related to one another on the surface of the antigen such that a cluster of pairs is defined which constitutes a major feature of the epitope-footprint, that is, the predicted epitope is in essence a cluster of connected AAPs. The AAPs of the epitope need not be consecutive tandem residues of the antigen, but often are the result of juxtaposition of distant residues brought together through folding of the polypeptide chain. Accordingly, we defined *D* as the distance parameter. The question is however, what distance and between which atoms of the AAP should be taken to define a legitimate pair? Based on an extensive analysis of antibody:antigen co-crystal data, it is evident that the majority of antibody:antigen contacts involve side-chain residues of the antigen rather than backbone atoms (to be published elsewhere, and see Refs. 15 and 16). Thus, *D* should reflect the distance between side-chains. This distance, however, would be different for each possible AAP and even more possibilities exist when one considers the variety of specific atoms associated with each side chain. Thus, to reduce the measurement to a single unambiguous value, *D* has been defined as the distance between carbon alphas of any given pair (albeit this is only loosely related to the actual distance of the interacting moieties themselves). A legitimate AAP can be considered as such when the distance between its two corresponding carbon alphas is less than a given *D* value. The most effective default value for *D* has been determined empirically to be between 9 and 10 Å. AAPs of the epitope are simulated by tandem residues of the peptides, affinity selected from a random peptide-library. Thus, each peptide is assumed to contain one or more epitope-relevant AAPs which are the basis for mAb recognition of that peptide. To identify the most meaningful AAPs present in the panel of peptides, the peptides are deconvoluted into pairs. For example, a peptide of the sequence ABCDE... would be written as the series of pairs: AB BC CD DE and so forth. All AAPs derived from a panel of peptides are then pooled and the frequency of each type is calculated and determined whether its representation in the pool is higher than its random expectation (based on the theoretical amino acid frequencies of the phage library). Semi random AAPs are also taken into account in situations when the random peptides contain, for example, constant cysteine residues introduced to impose disulfide constrained looped configurations (for more detail see Ref. 15). Thus, for each pair, the number of standard deviations above its expected random occurrence is computed. Only pairs for which this number exceeds the "Statistical Threshold" parameter, *ST*, are considered as statistically significant pairs (SSPs). The algorithm seeks these pairs for a selected *D* value on the surface of the antigen and attempts to link them into clusters. Mapitope considers exposed residues only. Thus, a third parameter, *E*, is the surface accessibility threshold. The accessibility of each amino acid is calculated using the "Surface Racer" software,[17] which has been assimilated in the algorithm software.

Depending on the size and quality of the peptide panel, one can overlook specific residues of the epitope that are contained within segments of the predicted clusters. Consider for example the 17b epitope (see Results) which includes a continuous segment, residues V200 to C205. The predicted epitope contains the same continuous segment (residues V200 to K207), but fails to spot the residues Q203 and A204. In such a case, we would

include the missed contact-residues as part of the final prediction as it would close the gap connecting two short predicted segments, generating a continuous element of the predicted epitope. Therefore, the final step of the algorithm is to "fill-in" gaps within predicted discontinuous segments. A fourth parameter, *I*, defines the maximum gap between two residues to be connected.

As contacts between the mAb and the antigen are mostly through functional moieties of the R-groups,[15,16] conserved residues were consolidated into 13 functional subgroups of amino acids and were given single-letter notations:

B = R, K; **J** = E, D; **O** = S, T; **U** = L, V, I; **X** = Q, N; **Z** = W, F;
A = A; **C** = C; **G** = G; **H** = H; **M** = M;**P** = P; **Y** = Y;

In summary, a mAb is used to screen a random peptide library to generate a panel of affinity selected peptides. These peptides are deconvoluted into AAPs and the most statistically significant pairs (SSPs) are identified. These are then mapped on the surface of the atomic structure of the antigen and the most elaborate and diverse clusters are identified. These are regarded as the predicted epitope candidates.

In comparison with the original algorithm[15] some of the improvements in the upgraded version, Mapitope, are as follows:

- Easy access to the parameters of the algorithm. Mapitope allows the user to alter parameters such as *D*, *ST*, *E*, and *I* easily, and thus to test the effect of each of the parameters on the prediction outcome.
- Gaps created between predicted and nonpredicted amino acids within the same strands can be automatically filled in by the algorithm according to the user definition of what constitutes a minimal gap.
- Improved predictive speed. The running time of any of the analyses presented here was within 10–20 s (using a Pentium 4, 2.4 GHz computer with 512 MB of RAM).
- The algorithm can now automatically locate the Q-point (see later) in which, for example, two clusters merge, thus defining the upper most limit of the predicted cluster.
- Improved output files. The output of Mapitope has been updated to include all AAPs, AAP types, and total amino acids predicted to participate in each cluster. Each phage displayed peptide is now marked for its contribution to each cluster, and the output is also written as a Rasmol script which allows one to easily view the clusters on the surface of the antigen using the RasTop or Rasmol programs.[18] Clusters are ranked and color-coded from the most likely to least likely and the top five epitope predictions are given as output.

Phage Displayed Peptides

In this study, we used three model systems to investigate the performance of Mapitope. Phage displayed peptides used as input were obtained through screening of phage display libraries with the antibody of interest. The peptides for each mAb are provided herewith:

mAb 17b peptides[15]

1. CLHIRVNETAYRVC	2. CEFFQQHMLRVPRC
3. CNMKLKLREMTQRC	4. CMTRPTSLTQLTGC
5. CMVRPSNWDALTRC	6. CDFLREHGMKNPRC
7. CRSRPTNMTTLRDC	8. CAAYNATRGTVSAC
9. CQLLHTWEDKMRKC	10. CRNGELWLRRPGLC
11. CSGLRNETFLRC	

mAb 13b5 peptides[15]

1. CAHFPPRSQMIADC	2. CAHFAPGTAMYSDC
3. CRQFPHSSSMYTDC	4. CRESRAALERGWWC
5. CEARTHNEARRRRC	6. CAAARSTGETSAHY
7. CYYRMGANYTVGEC	8. CSVSPLYAADDPLC
9. CTQMHEMDPNFPPC	10. CVTALGPNYTGQEC
11. VWRCNWF	12. CVVFLDVSEAFRDC
13. CADVMGPLVTAAEC	14. CADVMGPLVTAGEC
15. AASWNGR	16. CYVQQPWWVLEREC

Trastuzumab peptides ("Riemer" peptides[19])

1. CQMWAPQWGPDC	2. CKLYWADGELTC
3. CKLYWADGEFTC	4. CVDYHYEGTITC
5. CVDYHYEGAITC	

Novel trastuzumab peptides ("TAU" peptides)

Trastuzumab was used to screen the FMC12C phage display library as previously described[8] and the following peptides were isolated and validated for antibody binding.

1. CGWASGMADGSSNC	2. CAGWKTGEADGSSC
3. CGWTSGKSDGSASC	4. CTSLVADLDHLSSC
5. CPNIGELSHYDPFC	6. CAAWHTGKAEGNGC
7. CAGRWEHGTAEGDC	8. CTLWVLGLADGTRC

RESULTS

The first step in any Mapitope prediction is to conduct a preliminary analysis of a panel of peptides applying default parameters (D = 9Å, *ST = 3, and E* = 5%). This generates as a first approximation a collection of candidate epitopes, that is, clusters. The question to be asked, therefore, is whether by manipulation of the defining-parameters described earlier, one can identify that cluster which most likely represents the genuine epitope? Subsequently, the boundaries of the best cluster should be determined.

The following is a systematic analysis of the impact of the parameters *ST*, *D*, *E*, and *I* on Mapitope predictions. Two model systems of antibodies, whose epitopes were determined through co-crystallization, were used. The first model is that of the 17b human mAb directed against the surface envelope protein of HIV, gp120.[20] The second model is that of the 13b5 human mAb directed against p24 of HIV.[21] On the basis of these anal-

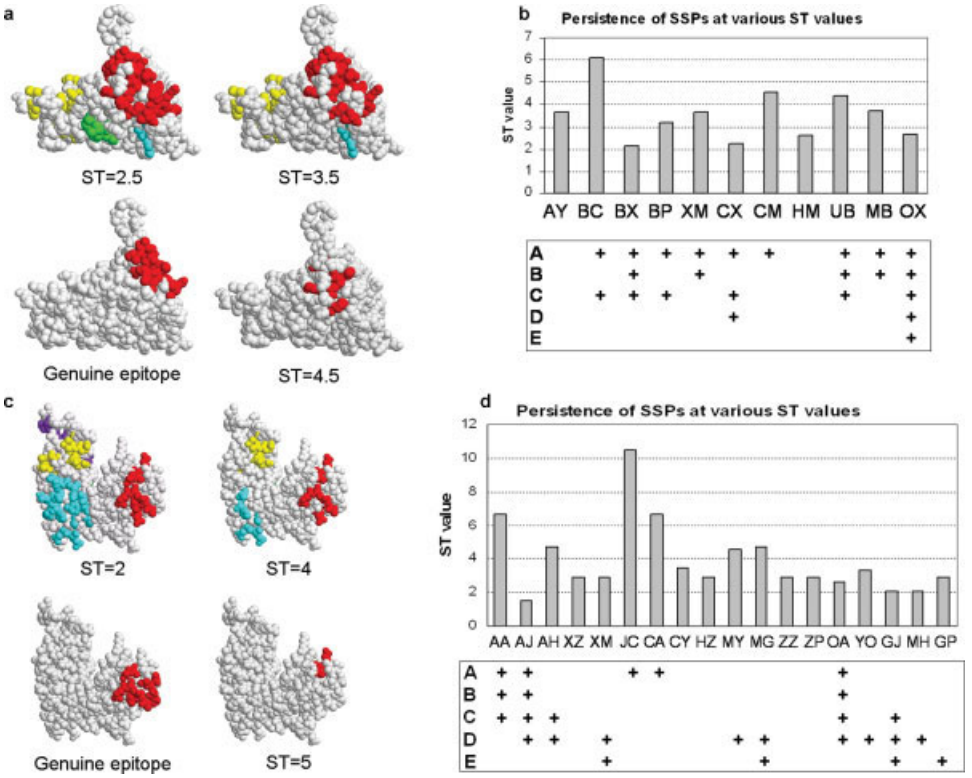

Fig. 1. Impact of the *ST* parameter on the prediction of the 17b and 13b5 epitopes. The predicted epitopes at different *ST* values for 17b (**a**) and for 13b5 (**c**) are compared with the genuine epitopes. Histograms **(b)** and **(d)** show the highest *ST* value where each SSP still contributes to the predictions (for more details see text). The lower panels show the distribution of the SSPs for each cluster. Cluster A is red, Cluster B yellow, Cluster C green, Cluster D cyan and Cluster E purple. Note that Cluster E and Cluster C for the 17b and 13b5 predictions, respectively, are not visible.

yses we have formulated a paradigm for epitope prediction, which has been used to predict the epitope for a third model antibody: Trastuzumab, which was previously co-crystallized with its corresponding receptor Her-2/neu.[22] For each parameter tested, the other parameters were fixed to their default values.

Impact of the *ST* Parameter

As described in Materials and Methods, following deconvolution of the peptides into a pool of AAPs, Mapitope identifies those pairs which are statistically enriched (SSPs) and uses them to predict the epitope on the surface of the antigen. Among the AAPs that exist in the pool of peptides, those that are critical for antibody-recognition are expected to be particularly enriched. Consequently, we evaluated the ability to focus our prediction on the most meaningful clusters by performing the analyses using only SSPs of the highest *ST* values. This was intended to reduce the number of potential epitope candidates and improve their quality.

When limiting the prediction, that is, using only SSPs of highest *ST* values, clusters are considered valid so long as they contain two or more pair types. Figure 1(a) shows the predicted clusters on the surface of gp120, using the 17b peptide database. Cluster A (red) overlaps the genuine epitope, while all the other predicted clusters (Clusters B–E) are in reality false positives. When performing predictions at increasing *ST* values, false positive clusters are gradually lost and the cluster corresponding to the genuine epitope stands out as most persistent as is illustrated for $ST = 4.5$. Figure 1(b) is a histogram describing the persistence of SSPs incorporated in the predictions made at different *ST* values. The bottom panel of Figure 1(b) presents the distribution of the SSPs in the various clusters. Cluster A contains nine pair types of which three are of an $ST \geq 4$ (BC, CM, and UB). Three additional pair types are of an $ST \geq 3$ (BP, XM, and MB). Only three pair types (BX, CX, and OX) are of $ST < 3$. Considering the diversities and the *ST* values of the pairs, one appreciates that Cluster A is the most elaborate one, making it the "strongest" prediction.

A second example is given in Figure 1(c) demonstrating the predicted clusters using the 13b5 peptide database on the surface of p24. As with 17b, increased *ST* values focus the prediction on Cluster A, which overlaps the genuine epitope. Figure 1(d) shows that three pair types utilized by this cluster have an $ST \geq 6$ (AA, JC, CA). Interestingly, whereas Cluster D (cyan) initially contains a higher number of SSPs (nine pair types), it does not persist throughout the analyses as we increase the *ST* parameter above five (of the nine pair types five

have $ST < 3$), illustrating that although Cluster D is derived from a higher number of pairs at the default $ST = 3$, the majority are of lesser statistical significance, making this cluster less favorable.

We conclude that predictions based on pairs of relatively high *ST* values have a better chance of overlapping the genuine epitope. Therefore, we would recommend that one should collect as many peptides as possible, thus ensuring more meaningful statistics, increasing the signal to background ratio, thereby making *bona fide* SSPs more visible.

Impact of the *D* Parameter

Whereas the statistical analysis and definition of SSPs are based on the peptide-database, the distance parameter (*D*) deals with the physical nature of the antigen. The predictive algorithm seeks to find discontinuous pairs on and within the antigen that are functionally represented by tandem residues in the peptides (for more details see Materials and Methods and Ref. 15). Therefore, the definition of what constitutes a legitimate pair is critical for Mapitope prediction. We have defined the parameter *D* as the distance between two carbon alphas (Mapitope also permits defining *D* as the distance between any two atoms of a given AAP).

Figure 2 shows the effect of *D* on the number of amino acids included in a given predicted cluster. In both examples (17b and 13b5), as *D* increases there is a point in which a sharp increase in the number of amino acids participating in a given cluster is apparent (at $D = 10.5$ Å and $D = 17$ Å for 17b and 13b5, respectively). The *D* value just preceding this point is termed the "Q-point" (the point of "quantum increase"). The sharp increase just beyond the Q-point is typically the result of merging two or more clusters. It seems that above Q one looses resolution and the ability to distinguish between meaningful clusters. In principle, predictions should be made at *D* values that are less than the Q-point. For most cases *D*-default < Q and as such, predictions at the default distances are satisfactory. Some improvement can be achieved in predictions by increasing the *D* value higher than the default, yet still less than Q, and in this way include additional residues to the Cluster in question. In the situations where the default-*D* is greater than an empirically detected Q the final prediction should be conducted at Q so to avoid inclusion of irrelevant amino acids contributed by the other merging clusters.[23]

Impact of the *E* Parameter

For an antibody to bind its epitope, contact residues must be accessible to the surface of the antigen, at least to some degree. Mapitope allows the user to employ a variety of accessibility thresholds (defined by the *E* parameter) for amino acids to be considered as components of a predicted cluster. Setting the *E* value to 5%, only those amino acids that are at least 5% exposed, will be considered.

Conducting the analyses on both 17b and 13b5 peptides using *E* values ranging between 5 and 20% had little effect on the prediction of either epitope (data not shown). We thus set a default *E* value of 5%.

Filling-In the Gaps

Filling-in the gaps within segments of the predicted clusters leads to inclusion of relevant residues as well as irrelevant ones. We set the default value of the *I* parameter to three residues, which has been set empirically. Accordingly, the final prediction of the 17b epitope (Fig. 3) would be the peptide segments L116-P124, V200-K207, P417-I439 of gp120, which includes the four beta strands of the genuine epitope. Figure 3 also illustrates the comparison of the genuine 13b5 epitope with its corresponding predicted cluster.

Summary of 17b and 13b5 Predictions

The co-crystal of the 17b mAb bound to the CD4/gp120 complex[20] reveals that the amino acid residues of gp120 that contact the antibody reside on four anti-parallel beta strands. These strands comprise three discontinuous segments of gp120 brought together by protein folding (Fig. 3). Out of the 16 amino acids comprising the genuine epitope, only two (Q203, A204) are not predicted by Mapitope. Of the 31 predicted residues, 14 coincide precisely with the genuine contact residues of the epitope, whereas 17 residues were falsely predicted (according to the co-crystal structure). Figure 3 also provides a spacefill summary of the prediction versus the genuine epitope. The Mapitope preliminary analysis predicted four clusters of which Cluster A was correctly identified as most likely to coincide with the genuine epitope of 17b. Moreover, when using the fill-in option the two missed residues (Q203, A204) are included in the final prediction. Similar comparisons between the genuine epitope of the 13b5 mAb and its Mapitope prediction are given in Figure 3.

A Paradigm for Epitope Mapping

In view of the results described earlier, we have attempted to formulate a paradigm designed to determine unknown epitopes, provided that a model of the atomic structure of the antigen is available and that a panel of mAb-specific peptides has been produced. The following is a description of the steps involved, which are now fully automated features in the program.

Step (1): Identifying potential clusters

For this, an initial analysis is carried out using the default values $ST = 3$, $D = 9$ Å, $E = 5\%$, and $I = 0$. This provides a number of clusters representing potential epitope candidates. Clusters are defined as containing at least two pair types, and all clusters complying with this description are considered valid candidates (intuitively, clusters containing a larger number of pair types should be better candidates).

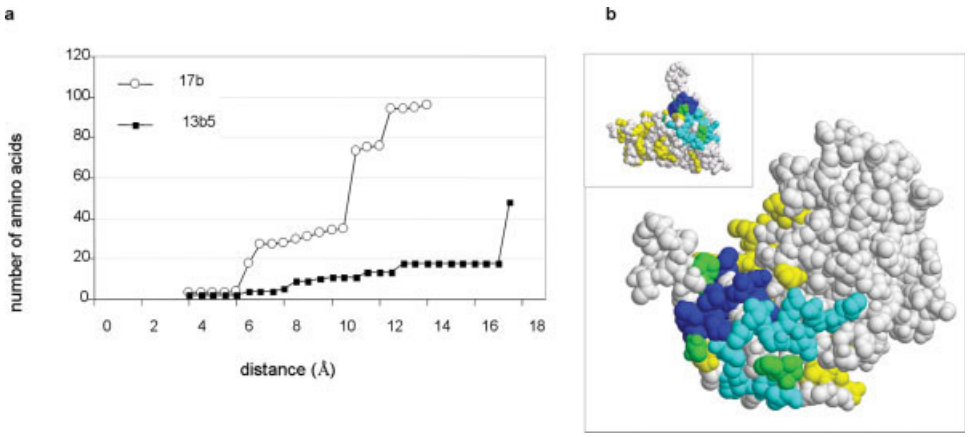

Fig. 2. Impact of the *D* parameter on the 17b and 13b5 epitope predictions. The 17b and 13b5 epitopes were predicted using increasing values of *D*. **(a)** The number of amino acids associated with the predicted clusters corresponding to the genuine epitope (see Fig. 1(a,c) for mAbs 17b and 13b5, respectively) as a function of ever increasing *D* values. Q-point exists at $D = 10.5$ Å for 17b and at $D = 17$ Å for 13b5. (**b**) Amino acids incorporated into the cluster corresponding to the 17b genuine epitope, using different *D* values, are represented on the spacefill model of gp120. Amino acids predicted using $D = 6$ Å are colored blue, amino acids added using $D = 8$ Å are cyan, $D = 10$ Å green, and $D = 12$ Å yellow.

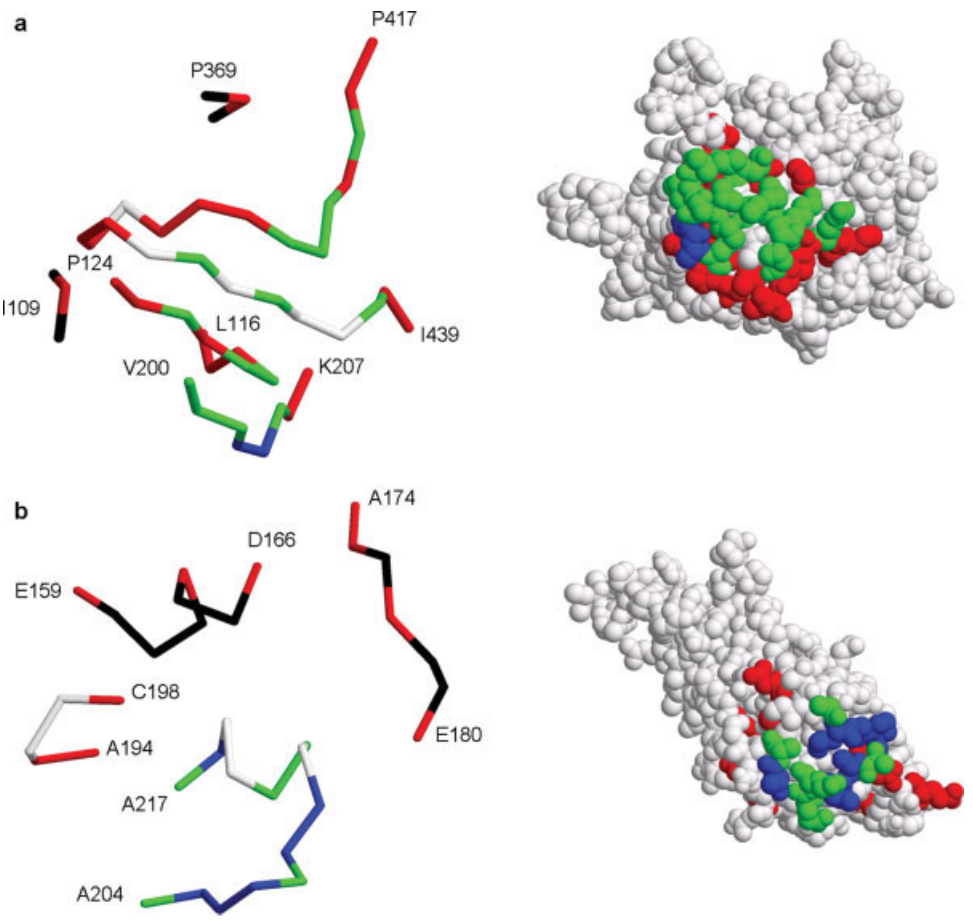

Fig. 3. Final prediction and comparison of the predicted 17b and 13b5 epitopes with the genuine epitopes. (**a**) Left: Backbone representation of the strands comprising the genuine and predicted 17b epitope (at Q-point). Hits are colored green, over-predicted amino acids are in red and missed amino acids are in blue. Amino acids used to fill-in the predicted ones are colored light grey (except for Q203-A204 which are filled-in but are colored blue as they are part of the genuine 17b epitope). Filled-in amino acids in outlier strands (i.e. added at Q-point but are not predicted at the $D = 9$ Å default) are colored black. Right: Same as on the left, but presented on the spacefilling model of gp120 and no fill-in is included. (**b**) Left: Backbone representation of the strands comprising the genuine and predicted 13b5 epitope (at Q-point). Coloring is the same as in (a). Right: Same as on the left, but presented on the spacefilling model of p24 and no fill-in is included.

Step (2): Focusing on the clusters most likely to represent the genuine epitope

Once clusters are identified in Step (1), *ST* values are increased to identify those with the highest perseverance. Those clusters that continue to be predicted at the highest *ST* values are considered as the most likely candidates to represent the genuine epitope. All the clusters that persist at this stage are considered as valid epitope candidates.

Step (3): Avoiding prediction of merged clusters

Clusters identified in Step (2) are then analyzed individually maintaining $ST = 3$, while altering D values between 4 Å and the default-D. AAPs within the affinity selected peptides are tandem by definition; therefore the minimum distance taken between two carbon alphas in the epitope is that of a peptide bond, about 4 Å.[24] In those instances where Q < default-D the prediction is performed at the Q-point as opposed to all other situations where default-D should be used. To ensure that the cluster in question is not the product of the merged combination of two or more clusters, the process of determining the Q-point is now also performed automatically by the algorithm.

Step (4): Refining the prediction

Steps (1)–(3) identify residues based on the statistical parameters of the peptides (*ST* values) or the physical parameters of the antigen (D and E values). Step (4) fills in gaps, including residues which were not predicted in the previous steps. For this, every segment predicted by the algorithm is included in its entirety provided that the first and last residue of this segment is predicted and gaps to be filled in are no greater than the I value (typically $I = 3$). The fill-in function (I) is now also an integral part of the algorithm.

The output of Mapitope is a number of clusters ranked such that the cluster considered most likely to represent the genuine epitope is that which includes the maximal number of pair types, followed by the highest number of pairs, and finally the total number of residues.

Mapitope Prediction of the Trastuzumab Epitope on the Surface of Her-2/neu

To illustrate the application of the stepwise protocol described earlier, we used Mapitope for prediction of the Trastuzumab (Herceptin®) epitope on the Her-2/neu receptor protein. The Her-2/neu receptor belongs to the epidermal growth factor receptors family (EGFR), and its abnormal activation is correlated with several types of cancer.[25,26] Trastuzumab was shown to be highly efficacious in the treatment of these cancers.[25,27] The Trastuzumab Fab was co-crystallized along with the extracellular domain of the human Her-2/neu receptor[22] and the epitope was found to incorporate 3 segments: amino acids P557-Q561, D570-F573, and K593-P603. Riemer et al.[19] have used Trastuzumab to screen phage display peptide libraries and isolated a panel of five peptides (see Materials and Methods). The following is a description of a Mapitope analysis of the "Riemer" peptides and an additional panel of peptides derived in the course of this study (designated "TAU" peptides).

Step (1)

Preliminary analysis of the "Riemer" peptides was conducted using the default parameters of $ST = 3$, $D = 9$ Å, $E = 5\%$, and $I = 0$. As seen in the left panel of Figure 4, this prediction generated five clusters (Clusters A–E), of which Cluster A and Cluster B ranked equally as each utilizes seven pair types.

Step (2)

To focus on the cluster most likely to correspond to the genuine epitope, all the clusters were analyzed by elevating the *ST* values (Fig. 4). Three clusters–Clusters A, B, and E continued to exist at $ST > 5$. A slight preference for Cluster A does however, become apparent at the highest *ST* value where the clusters still exist ($ST = 7.5$). Here Cluster A contained three pair types, as opposed to Clusters B and E which contained only two pair types each. In view of the above, all three clusters would be considered valid candidates, with a marginal advantage to Cluster A.

To further test this conclusion (i.e. that Cluster A ranks above other predicted clusters), we decided to screen Trastuzumab against our phage display peptide library so as to increase the total number of peptides and also provide a second independent data set. We isolated eight peptides that bound the antibody well. These "TAU" peptides showed neither common linear homology with Her-2/neu, nor with the five peptides previously isolated by Riemer et al. Nonetheless, as illustrated in the right panel of Figure 4, in the preliminary prediction using default parameters, the previously identified Clusters A and B were predicted anew, and a novel "TAU" Cluster C is also apparent. Note that "TAU" Cluster C is in essence an expansion of "Riemer's" Cluster C which has also merged with the persistent Cluster E (see Fig. 4) Step 2 in the analysis of the "TAU" peptides was preformed, increasing the *ST* values. Here too, three candidates, Clusters A, B, and C, all continued to exist even at exceptionally high *ST* values (7.66). As shown in Figure 4, Clusters A and B are predicted at high *ST* values using both panels of peptides and contain the largest number of SSPs. In view of the above, Clusters A and B would be considered as valid epitope candidates with a slight advantage to Cluster A. Combining the two data sets, that is, a total of 13 peptides, did not improve the analyses (data not shown). Given the fact that a co-crystal exists, we were able to compare our prediction to the solved structure and found that Cluster A coincides with the genuine epitope (Fig. 4, small insert). This illustrates the power of the algorithm to propose a limited number of solutions (1–2 epitope candidates) with high probability of including the *bona fide* epitope.

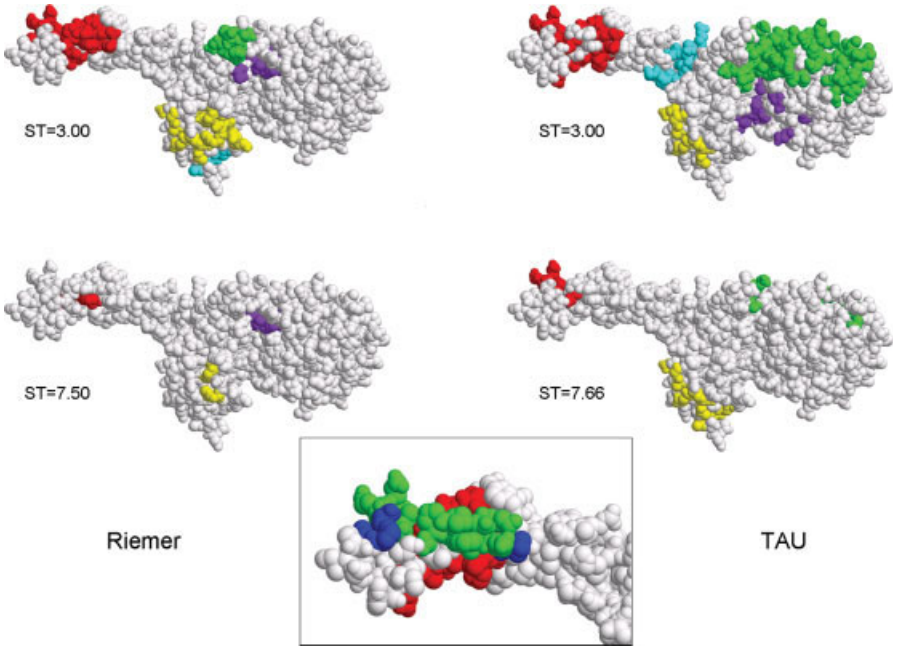

Fig. 4. Mapitope prediction of the Trastuzumab epitope. Predictions of the Trastuzumab epitope based on the peptides published by Riemer et al. ("Riemer") and on the TAU peptides as indicated. The top predictions are made using the default *ST* value ($ST = 3$) and the lower predictions at the highest *ST* value in which clusters are still obtained. The red cluster in both instances corresponds to the genuine Trastuzumab epitope and is designated Cluster A. Cluster B is yellow, Cluster C green, Cluster D cyan, and Cluster E purple. Comparison between the Mapitope prediction and the genuine epitope is given in the insert. Hits are colored green, over-predicted amino acids are in red and missed amino acids are in blue.

Step (3)

According to the paradigm, each cluster would be analyzed at ever increasing D values, starting from 4 Å and up to 9Å, confirming that no Q-point exists below the default D (in our case both Clusters A and B are analyzed).

Step (4)

The last step of the paradigm is filling-in and interconnecting all specifically predicted amino acids using the default parameters, $ST = 3$ and $D = 9$ Å. As shown in Figure 4, once this assignment is performed, the final prediction comprises all three loops known to encompass the epitope, except six amino acids (P557, K593-P595, N602-P603).

DISCUSSION

The ultimate golden standard for definitive determination of an epitope of a given mAb is co-crystallization of the antibody bound to its antigen. However, this is a challenging task as reflected by the fact that as of December 2006 only 250 crystals of mAb:antigen complexes have been described, many of which are redundant or contain peptides rather than intact antigens. Therefore, alternative methods are required when one is interested in localizing the epitope of a specific mAb for which there is no co-crystal. In the event that the structure of the antigen has been solved, one can resort to a number of mapping strategies, of which the Mapitope algorithm has been found to be extremely useful.

Phage display peptide libraries have been extensively used for the mapping of linear epitopes, where the homology between the peptides and the antigen is self evident.[28–30] However, a much more challenging task is to use the peptides for the prediction of discontinuous conformational epitopes, where no such homology exists. The use of a panel of affinity isolated phage displayed peptides as a database for epitope mapping is central to Mapitope. The algorithm is based on two fundamental ideas; (i) the hypothesis that AAPs can serve as elementary recognition cues for antibody binding leading to those that are specifically enriched (i.e. SSPs) and (ii) the existence of clusters which are in essence interconnected SSPs on the surface of the antigen.

The critical role played by SSPs in peptide recognition has recently been demonstrated by Bublil et al.[23] in which two serine residues were present in a given affinity purified peptide (CLWSDLLSQYTKPC). The Serine in position 4 is a member of two SSPs (WS and SD), whereas the AAPs including serine 8 (LS and SQ) were not enriched. Mutation of serine 4 to alanine abolished the mAb's binding totally, whereas altering serine 8 to alanine had no effect at all.

In this study, we have systematically examined three mAb:antigen co-crystals[20–22] as positive control models

so as to evaluate the impact of the parameters *ST*, *D*, *E*, and *I* on epitope prediction fidelity and offer a stepwise paradigm for the application of Mapitope in the future.

Our analyses show that the most important parameter in terms of identifying the correct cluster (that which overlaps with the genuine epitope) is *ST*. As we discovered, increasing the stringency of the analyses by using the most enriched SSPs, tends to weed-out false positive clusters that seem to be based on less meaningful pairs. This can be achieved by increasing the number of peptides in a given panel, which tends to accentuate the *bona fide* enriched pairs and balance-out the fortuitous fluctuations in frequency of the irrelevant pairs to background.

One possible concern, however, as with any epitope mapping algorithm is that one or more peptides may reflect non-specific binding (i.e. noise). The amount of such noisy peptides could be highly variable. It depends on the experimental procedures used during panning of the phage library, such as the stringency of selection and the use of multiple rounds of panning and enrichment. To test the sensitivity of Mapitope to the quality of the input data, we simulated a scenario in which random irrelevant peptides were included in the original panel of peptide sequences. The augmented peptide list was given as input to Mapitope and its success in predicting the true epitope was evaluated. In these simulations we tested the inclusion of 1–20 peptides to the 17b dataset and repeated this process 10 times. As expected, as the number of random peptides increased, the quality of prediction deteriorated. Interestingly, including a small number of random peptides (up to four) did not impair the prediction markedly. In fact, in some scenarios the prediction was superior to the original one owing to the fact that only the most significant pairs, that is, SSPs became more accentuated.

The impact of *D* is primarily on the number of segments and residues to be included in a given cluster. So long as a Q-point does not develop before the default-*D* value, running predictions at default-*D* between 9 and 10 Å consistently gives reasonable results as compared to the genuine epitopes determined by co-crystallization. In the event, however, that a Q-point does develop below the default-*D* then predictions should be made at the Q-point. A case in point is the recent analysis of the epitope of the HIV-neutralizing b12 mAb.[23] In the initial analysis the default $D = 9$ Å was used and Cluster B was identified as the preferred prediction. However, upon further investigation, it became clear that Cluster B developed a Q-point at 8.5 Å, and thus final prediction was conducted at this Q-point. Performing a prediction for mAb b12 at $D = 9$ Å would have wrongly included extraneous strands and irrelevant amino acids as has since been confirmed by comparing our Mapitope predictions to a recent publication.[32]

Given the availability of a solved or modeled structure for an antigen, Mapitope prediction of an epitope of interest only requires the production of a panel of mAb affinity isolated peptides. This is a rather straightforward task that for the most part is easily achievable. How then does Mapitope analysis compare with other methods for prediction of discontinuous conformational epitopes?

The problem of epitope mapping has been studied extensively.[33] The most common method for epitope mapping is point directed mutagenesis, which alters amino acids on the surface of the antigen that are suspected to participate in binding of the antibody. Epitope relevant residues are identified as those that when mutated significantly reduce binding. However, it was previously shown that often, mutations of what is ultimately a *bona fide* contact residue may have no effect on binding at all. For example, Dall'acqua et al.[34] have shown that mutations in 8 contact positions in the Hen Egg Lysozyme (HEL) antigen had little or no effect on binding of the D1.3 mAb (similar results have been obtained for the growth hormone binding to its receptor[35]). More specifically, Dall'acqua et al. have analyzed the mutation of HEL residue D18, a critical D1.3 contact residue. Altering the aspartic acid to alanine, results in the loss of one hydrogen bond and seven van der Waals contacts. Nevertheless, as demonstrated by surface plasmon resonance detection using a BIAcore biosensor, the affinity of the mAb for the mutated antigen is nearly identical to that of the wild type antigen. Subsequently, it was found that the loss of contacts could be compensated by the stable inclusion of additional water molecules at the interface and by local rearrangement in solvent structure. Another example of misleading results of antigen mutagenesis is found in Rizzuto et al.[36] Here it is shown that the mutations L122S, K432A, V200S, and Q203L, all of which are contact residues of the 17b mAb on the gp120 antigen, had no effect on 17b binding to gp120 wtΔ (gp120 lacking the V1 and V2 variable loops and NH2-terminus). The opposite situation also exists, that is, mutagenesis of residues that are clearly not directly part of the mAb epitope yet have a profound inhibitory effect on antibody binding. Numerous residues in HIV's gp120, have been mutated that interfere with mAb b12 binding to gp120, and certainly cannot all be genuine physical components of the mAb's epitope. For example, residue I213 has been reported to inhibit b12 binding by 80%, yet it is situated on the "backside" of gp120.[37] These examples show that mutagenesis should be interpreted with caution. In the event that a region of the antigen is already suspected to be an epitope, targeted mutagenesis may provide insights as to the involvement of one residue or other.

Another method developed to decipher protein:protein interaction is computational docking (for a comprehensive review see Halperin et al.[38]). The absolute prerequisite for this procedure is that the atomic structures for the antigen and the antibody be available. Given that, computational docking goes on to predict the relative position and orientation of the binding partners to one another. Computational docking has been tested extensively with the D1.3:lysozyme complex as a model system.[39–41] Most docking approaches are able to produce a Fab:lysozyme complex, however, the genuine epitope is rarely ranked with the top docking scores.

It seems that conditions that are favorable for docking are tightly fitting ligands and a well defined binding site. This can be satisfied best with small molecules that bind to topographically complementary binding surfaces such as in the case of enzyme:inhibitor complexes.[4,42] However, in the event of examining antibody:antigen interactions, the problem is much more complicated. Often the epitope recognized by an antibody has a broad footprint and may seem featureless. For example, because of the large size of the Hemagglutinin complexed to the BH151 mAb, none of the investigators that attempted to dock the two were able to reproduce the experimental data, that is, the relative position and orientation of the two proteins as determined from the co-crystal.[42] Even in the event of a successful docking of an antibody to its epitope within a given antigen, often one finds 3–5 solutions.[4]

The use of phage displayed peptides for epitope mapping is not novel (a comprehensive study of epitope-predictive algorithms was published by Mayrose et al.,[31] see also Refs. 14, 19, 43, and 44). However, the specific case reported by Riemer et al.[19] in which phage displayed peptides were used to map the epitope of Trastuzumab is particularly relevant to the current analyses. The Riemer algorithm was able to predict one segment (P595–P603) out of the three comprising the Trastuzumab epitope, in contrast to the Mapitope prediction, which successfully predicted all three loops of the epitope (Fig. 4).

The predictions presented in this study are based on cysteine-looped peptides from phage display libraries. Do constrained looped peptides provide a better database for Mapitope predictions as compared to linear peptides? In view of the fact that Mapitope seeks AAPs in the peptides that simulate pairs of amino acids on the antigen's surface, one would assume that the antibody must "see" both the R groups of the AAP on the same plane. In fully extended linear peptides the R groups of tandem residues face opposite planes forming a classical beta strand. On the other hand, imposing a constrained loop in a relatively short peptide (5–15 residues long) one tends to prevent the beta strand orientation. Thus, it would appear that for the most effective presentation of tandem R groups of AAPs, the looped peptide libraries are preferable.

Predictive computational algorithms or empirical methodologies typically generate a number of possible solutions which then must be further analyzed to ultimately discover the correct one that corresponds to the genuine epitope being sought. In this respect, Mapitope is no different and as such generates up to five candidate clusters that need to be further examined. The strength of Mapitope is that it can reduce an enormously large number of possible epitope candidates to an experimentally manageable few. This can be accomplished with relative ease and with reasonable chances for success. The question that arises is "can we, from our predictions thus far, gain new insights about general properties and amino acid compositions of the antibody:antigen interface that may enable us to improve the predictions?" Clearly, for this one would require a comprehensive set of Mapitope analyses. To date, in addition to the three examples presented here, three additional mAbs have been analyzed with Mapitope. In the case of the Bo2C11 mAb that binds factor VIII, we used the published panel of 27 phage displayed peptides and compared the Mapitope predictions with the co-crystal published by Spiegel et al.[45] and scored very well (unpublished data). In addition to the anti-HIV mAb b12 epitope prediction mentioned earlier,[23] we have recently published another prediction of the 80R mAb epitope, on the SARS corona virus spike protein.[46] The subsequent crystallization of 80R with its antigen,[47] reveals that here as well, Mapitope prediction overlaps with the genuine epitope. Nonetheless, six independent Mapitope analyses are still too few to draw any general conclusion regarding the characterization of epitope amino acid composition. In view of this we have taken an alternative approach, in which we are systematically analyzing all existing antibody:antigen co-crystal structures available in the Protein Data Bank. Out of some 250 co-crystals, about 65 have been identified as unique and include intact antigens or substantial fragments thereof. A variety of parameters are being analyzed such as amino acid prevalence at the binding interface, involvement of side-chain as opposed to backbone contacts, the existence of secondary structures and others. Clearly as we learn more about the characteristics of the binding surface, we will be able to incorporate this information into the Mapitope algorithm to enhance the prediction of correct contacts and filter out false positives.

The upgrade of our original algorithm[15] to the Mapitope format makes the predictions quick and, following the paradigm presented here, leads to more accurate predictions. Our continuing effort to upgrade and improve computational epitope predictions will provide simple practical means to discover conformational discontinuous epitopes of medically relevant antibodies and thus contribute to the rational design of novel drugs and vaccines.

ACKNOWLEDGMENTS

E.M.B. and N.T.F are the recipients of the Jakov, Mirianna and Jorge Saia Doctoral Prize. This study is part of the PhD thesis of E.M.B. O.P and N.D.R are fellows of the Edmond J. Safra. Program in Bioinformatics at Tel-Aviv Univesity.

REFERENCES

1. Tetin SY, Stroupe SD. Antibodies in diagnostic applications. Curr Pharm Biotechnol 2004;5:9–16.
2. Berger M, Shankar V, Vafai A. Therapeutic applications of monoclonal antibodies. Am J Med Sci 2002;324:14–30.
2a. Gershoni JM, Roitburd-Berman A, Siman-Tov DD, Tarnovitski Freund N, Weiss Y. Epitope mapping. The first step in developing epitope-based vaccines. BioDrugs 2007;21(2):1.
3. Sussman JL, Lin D, Jiang J, Manning NO, Prilusky J, Ritter O, Abola EE. Protein Data Bank (PDB): database of three-dimensional structural information of biological macromolecules. Acta Crystallogr D Biol Crystallogr 1998;54 (Part 6, Part 1):1078–1084.

4. Sotriffer CA, Flader W, Winger RH, Rode BM, Liedl KR, Varga JM. Automated docking of ligands to antibodies: methods and applications. Methods 2000;20:280–291.
5. Totrov M, Abagyan R. Detailed ab initio prediction of lysozyme-antibody complex with 1.6 A accuracy. Nat Struct Biol 1994;1:259–263.
6. Saphire EO, Parren PW, Pantophlet R, Zwick MB, Morris GM, Rudd PM, Dwek RA, Stanfield RL, Burton DR, Wilson IA. Crystal structure of a neutralizing human IGG against HIV-1: a template for vaccine design. Science 2001;293:1155–1159.
7. Benjamin DC, Perdue SS. Site-directed mutagenesis in epitope mapping. Methods 1996;9:508–515.
8. Enshell-Seijffers D, Smelyanski L, Gershoni JM. The rational design of a 'type 88' genetically stable peptide display vector in the filamentous bacteriophage fd. Nucleic Acids Res 2001;29:E50–50.
9. Cortese R, Felici F, Galfre G, Luzzago A, Monaci P, Nicosia A. Epitope discovery using peptide libraries displayed on phage. Trends Biotechnol 1994;12:262–267.
10. Cwirla SE, Peters EA, Barrett RW, Dower WJ. Peptides on phage: a vast library of peptides for identifying ligands. Proc Natl Acad Sci USA 1990;87:6378–6382.
11. Scott JK, Smith GP. Searching for peptide ligands with an epitope library. Science 1990;249:386–390.
12. Zwick MB, Bonnycastle LL, Menendez A, Irving MB, Barbas CF, III, Parren PW, Burton DR, Scott JK. Identification and characterization of a peptide that specifically binds the human, broadly neutralizing anti-human immunodeficiency virus type 1 antibody b12. J Virol 2001;75:6692–6699.
13. Fleming TJ, Sachdeva M, Delic M, Beltzer J, Wescott CR, Devlin M, Lander RC, Nixon AE, Roschke V, Hilbert DM, Sexton DJ. Discovery of high-affinity peptide binders to BLyS by phage display. J Mol Recognit 2005;18:94–102.
14. Schreiber A, Humbert M, Benz A, Dietrich U. 3D-Epitope-Explorer (3DEX): localization of conformational epitopes within three-dimensional structures of proteins. J Comput Chem 2005;26:879–887.
15. Enshell-Seijffers D, Denisov D, Groisman B, Smelyanski L, Meyuhas R, Gross G, Denisova G, Gershoni JM. The mapping and reconstitution of a conformational discontinuous B-cell epitope of HIV-1. J Mol Biol 2003;334:87–101.
16. Lo Conte L, Chothia C, Janin J. The atomic structure of protein-protein recognition sites. J Mol Biol 1999;285:2177–2198.
17. Tsodikov OV, Record MT, Jr, Sergeev YV. Novel computer program for fast exact calculation of accessible and molecular surface areas and average surface curvature. J Comput Chem 2002;23:600–609.
18. Sayle RA, Milner-White EJ. RASMOL: biomolecular graphics for all. Trends Biochem Sci 1995;20:374.
19. Riemer AB, Kraml G, Scheiner O, Zielinski CC, Jensen-Jarolim E. Matching of trastuzumab (Herceptin) epitope mimics onto the surface of Her-2/neu–a new method of epitope definition. Mol Immunol 2005;42:1121–1124.
20. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA. Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature 1998;393:648–659.
21. Berthet-Colominas C, Monaco S, Novelli A, Sibai G, Mallet F, Cusack S. Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. EMBO J 1999;18:1124–1136.
22. Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, Jr, Leahy DJ. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature 2003;421:756–760.
23. Bublil EM, Yeger-Azuz S, Gershoni JM. Computational prediction of the cross-reactive neutralizing epitope corresponding to the monoclonal antibody b12 specific for HIV-1 gp120. FASEB J 2006;20:1762–1774.
24. Nissen P, Hansen J, Ban N, Moore PB, Steitz TA. The structural basis of ribosome activity in peptide bond synthesis. Science 2000;289:920–930.
25. Yarden Y. Biology of HER2 and its importance in breast cancer. Oncology 2001;61 (Suppl 2):1–13.
26. Slamon DJ, Godolphin W, Jones LA, Holt JA, Wong SG, Keith DE, Levin WJ, Stuart SG, Udove J, Ullrich A, Press MF. Studies of the HER-2/neu proto-oncogene in human breast and ovarian cancer. Science 1989;244:707–712.
27. McKeage K, Perry CM. Trastuzumab: a review of its use in the treatment of metastatic breast cancer overexpressing HER2. Drugs 2002;62:209–243.
28. Denisova G, Stern B, Raviv D, Zwickel J, Smorodinsky NI, Gershoni JM. Humoral immune response to immunocomplexed HIV envelope glycoprotein 120. AIDS Res Hum Retroviruses 1996; 12:901–909.
29. Oleksiewicz MB, Botner A, Toft P, Normann P, Storgaard T. Epitope mapping porcine reproductive and respiratory syndrome virus by phage display: the nsp2 fragment of the replicase polyprotein contains a cluster of B-cell epitopes. J Virol 2001;75: 3277–3290.
30. Stern B, Denisova G, Buyaner D, Raviv D, Gershoni JM. Helical epitopes determined by low-stringency antibody screening of a combinatorial peptide library. FASEB J 1997;11:147–153.
31. Mayrose I, Shlomi T, Rubinstein ND, Gershoni JM, Ruppin E, Sharan R, Pupko T. A graph-based algorithm for epitope mapping using combinatorial phage-display libraries. Nucleic Acid Research 2007;35(1):69–78.
32. Zhou T, Xu L, Dey B, Hessell AJ, Van Ryk D, Xiang SH, Yang X, Zhang MY, Zwick MB, Arthos J, Burton DR, Dimitrov DS, Sodroski J, Wyatt R, Nabel GJ, Kwong PD. Structural definition of a conserved neutralization epitope on HIV-1 gp120. Nature 2007;445(7129):732–737.
33. Van Regenmortel MH, Pellequer JL. Predicting antigenic determinants in proteins: looking for unidimensional solutions to a three-dimensional problem? Pept Res 1994;7:224–228.
34. Dall'Acqua W, Goldman ER, Lin W, Teng C, Tsuchiya D, Li H, Ysern X, Braden BC, Li Y, Smith-Gill SJ, Mariuzza RA. A mutational analysis of binding interactions in an antigen-antibody protein-protein complex. Biochemistry 1998;37:7981–7991.
35. Cunningham BC, Wells JA. Comparison of a structural and a functional epitope. J Mol Biol 1993;234:554–563.
36. Rizzuto CD, Wyatt R, Hernandez-Ramos N, Sun Y, Kwong PD, Hendrickson WA, Sodroski J. A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. Science 1998;280:1949–1953.
37. Pantophlet R, Ollmann Saphire E, Poignard P, Parren PW, Wilson IA, Burton DR. Fine mapping of the interaction of neutralizing and nonneutralizing monoclonal antibodies with the CD4 binding site of human immunodeficiency virus type 1 gp120. J Virol 2003;77:642–658.
38. Halperin I, Ma B, Wolfson H, Nussinov R. Principles of docking: an overview of search algorithms and a guide to scoring functions. Proteins 2002;47:409–443.
39. Cherfils J, Duquerroy S, Janin J. Protein–protein recognition analyzed by docking simulation. Proteins 1991;11:271–280.
40. Walls PH, Sternberg MJ. New algorithm to model protein–protein recognition based on surface complementarity. Applications to antibody-antigen docking. J Mol Biol 1992;228:277–297.
41. Helmer-Citterich M, Tramontano A. PUZZLE: a new method for automated protein docking based on surface shape complementarity. J Mol Biol 1994;235:1021–1031.
42. Dixon JS. Evaluation of the CASP2 docking section. Proteins 1997;Suppl 1:198–204.
43. Pellequer JL, Westhof E. PREDITOP: a program for antigenicity prediction. J Mol Graph 1993;11:191–202, 204–210.
44. Halperin I, Wolfson H, Nussinov R. SiteLight: binding-site prediction using phage display libraries. Protein Sci 2003;12:1344–1359.
45. Spiegel PC, Jr, Jacquemin M, Saint-Remy JM, Stoddard BL, Pratt KP. Structure of a factor VIII C2 domain-immunoglobulin G4kappa Fab complex: identification of an inhibitory antibody epitope on the surface of factor VIII. Blood 2001;98:13–19.
46. Tarnovitski N, Matthews LJ, Sui J, Gershoni JM, Marasco WA. Mapping a neutralizing epitope on the SARS coronavirus spike protein: computational prediction based on affinity-selected peptides. J Mol Biol 2006;359:190–201.
47. Hwang WC, Lin Y, Santelli E, Sui J, Jaroszewski L, Stec B, Farzan M, Marasco WA, Liddington RC. Structural basis of neutralization by a human anti-SARS spike protein antibody, 80R. J Biol Chem 2006;281:34610–34616.