0273-2289/94/4702-03—0119$04.00

Antibody-Combining Sites

Extending the Natural Limits

DAVID M. WEBSTER, JAN PEDERSEN, DAVID STAUNTON, ALISON JONES, AND ANTHONY R. REES*

Department of Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK

ABSTRACT

The antibody repertoire is very large with at least 10^9 different antibody specificities, yet there are currently only 800 variable-region sequences known and < 23 Fab structures deposited with the Brookhaven Protein Data Bank. To engineer the antibody-combining site rationally, we need to define the rules that govern antibody structure. To understand the process of antibody-antigen recognition, we need not only to predict complementary determining regions accurately, but to simulate accurately the interaction of antibody with antigen. We have made progress in the modeling of antibody-combining sites and in the simulation of antibody complex formation. The combination of these approaches will allow us to extend the natural limits of antibody-combining sites in a more rational manner.

Index Entries: Antibody; antigenic epitope; antibody-combining site; complementarity determining region; molecular docking; peptide mimetics; metal binding sites.

INTRODUCTION

Antibodies, because of their vast repertoire of specificities, offer the possibility of an unlimited resource of binding species. This repertoire has now been extended to include catalytic antibodies, produced either by application of rational methods, such as transition-state analog immunization, or by identification of naturally occurring antibodies that exhibit

*Author to whom all correspondence and reprint requests should be addressed.

catalytic activity. To exploit this resource fully requires methods that can adequately tap the available repertoire. This is often not achievable because the monoclonal antibodies produced, either by conventional hybridoma technology or by phage display, lack the subtlety of specificity, or affinity of binding, required. In this situation, protein engineering by site-directed or random methods of mutagenesis can lead to improvements. Our own antibody work has focused on two areas. First, we have developed ways to model the variable regions of antibodies with high accuracy, using the models to guide protein engineering experiments. The introduction of specific, high-affinity binding sites for cofactors, such as transition metal ions, is one example of the application of these methods. Second, we are developing new docking algorithms to identify macromolecular complexes rapidly. We hope to combine such algorithms with combining-site design to generate a new generation of nonpeptide mimetics that may one day replace antibodies.

ANTIBODY MODELING

The antibody-variable domain consists of a β-barrel formed from two antiparallel β-sheets, one from the V_L chain and the other from the V_H chain. Antigen recognition is localized to six regions known as the hypervariable or complementarity determining regions (CDRs) (*1*). The CDRs are the most variable regions of the Fv domain in both sequence and structure as shown in Fig. 1. Since the Fv framework is considered to be well conserved between antibody structures (*2–4*), it is modeled first and then the more difficult CDR loops. The modeling protocol for both framework and CDRs is shown in Fig. 2, and is encoded in the computer program AbM (*5*).

Modeling the Framework

Although the Fv framework region is considered to be well conserved between antibody structures, variability in the packing of β-sheets and strands does occur. The orientation of the β-sheets inclined to one another may vary by as much as 30° ± 18° (*2,3*) and the strands at the domain interface are inclined at approx 50° (*4*). An analysis of 12 antibody structures has indicated that strands 1–6 of the variable region are highly conserved, whereas strands 7 and 8 are variable (*6*). It is interesting to note that these latter two strands interconnect CDR H3 of the heavy chain, which is the most variable CDR in both sequence and structure. It should also be noted that small differences in V_L/V_H orientation and packing, particularly where the orientation affects the CDR takeoff trajectories, can result in large errors in positioning the CDRs on the framework. The following protocol attempts to minimize these differences.

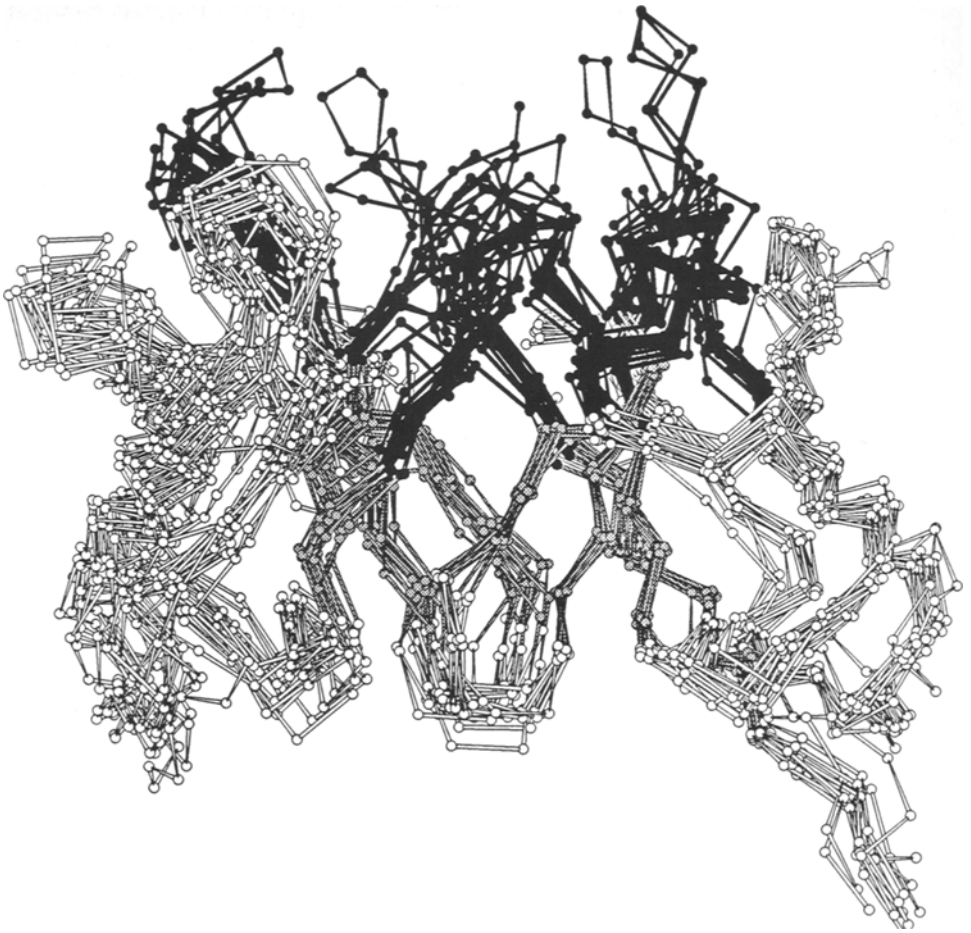

Fig. 1. Superimposition of antibody Fv region C-α backbones. The framework regions depicted in white show a high degree of structural similarity when compared with the CDR regions depicted in black. The CDRs form the antibody-combining site, the site of antigen recognition.

- The V_L and V_H domains are selected from a data base of antibody structures using sequence homology as the criteria for selection;
- If the V_L and V_H domains come from different antibodies, the backbones are least-squared fitted to a mean β-barrel based on strands 1–6;
- Strands 7–8 are taken from its antibody structure and fitted to the putative β-barrel; and
- Side chains are replaced using a maximum overlap procedure in which side-chain templates are fitted onto the backbone atoms adjusting torsion angles to those of the parent. Those atoms that do not have parent atoms are constructed by the conformational methods described in the following section.

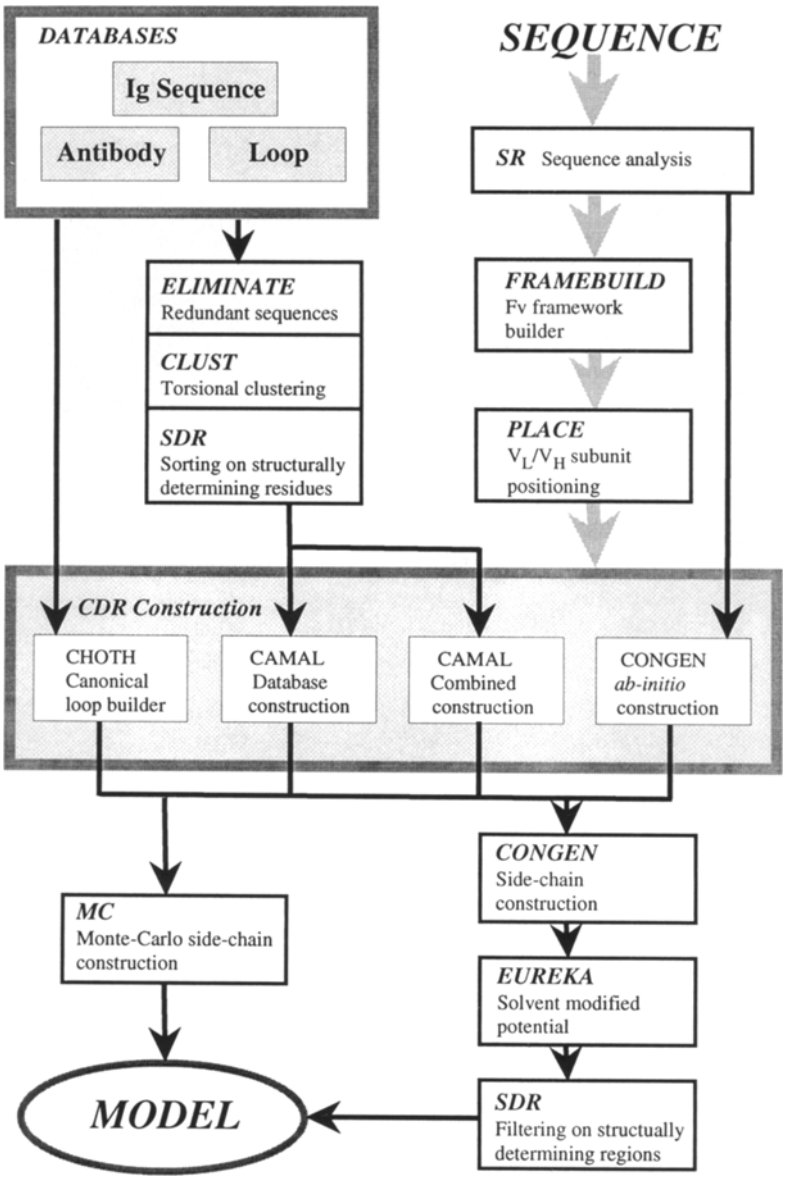

Fig. 2. Flowchart of the AbM modeling procedure. Individual steps in the procedure are in bold italics.

Modeling the CDRs

CDRs comprise three loops from the light chain and three loops from the heavy chain. The CDRs display the most variation in both sequence and structure with each loop interconnected by a framework strand. The topology of the loops, their takeoff trajectory from the framework, and their packing against each other determine the shape of the combining site and, hence, antigen recognition.

The method, used in our laboratory to model CDR loops, is embodied in an algorithm called CAMAL (*see* refs *12,13* for early versions of this algorithm and ref *6*), which employs either a date-base search or a combined data-base and/or conformational search to generate CDR backbone conformations or conformational searching alone by *ab initio* methods. The advantage of the combined approach is that not only is the existing knowledge base of antibody and other loop conformations utilized, but if necessary, the entire phase space for some or all loop conformations may be searched.

Canonical Method

Although CDRs are the most variable parts of the Fv, it was proposed by us (*7,8*) that CDRs of the same length may adopt similar conformations. Chothia and Lesk (*9,10*) proposed the concept of canonical loops where certain CDR conformations are determined by a few residues in a loop of a particular length, occasionally including residues from the framework. These key residues may form particular hydrogen bonding patterns or certain packing arrangements, or form unusual backbone torsional angles. Within the current definition of canonical classes, it has been noted (*6,11*) that variation may occur either because of an inadequate canonical rule or possibly because of inaccuracies in X-ray structures. At present, no canonical structures exist for CDR H3, although such classes may be identified as the structural base increases. This CDR occupies a central position in the antibody-combining site, and hence, accurate modeling of its conformation is critical.

Data-base Method

A knowledge base of known antibody structures (*12*) is used to construct a data base of C-α distance constraints from within the loops. A distance data base containing all entries from the Brookhaven Protein Databank (PDB) (*14*) is searched for loops of the same length that satisfy the antibody distance constraints within 3.5 SD. Redundancies are removed at this stage by torsional clustering and may be further filtered on structurally determining regions (SDR) (*15*).

Ab initio Reconstruction

Phase space is saturated using the conformational search program CONGEN (*15*), which has been implemented in AbM (*12,13*). The program searches conformational space by rotation on a torsional grid about the ψ and ϕ torsion angles, and about the χ torsion angle during side-chain reconstruction. Phase space is restricted by Ramachandran energies using glycine, alanine, and proline maps to cover all residue types. Construction commences at the base of the loop and proceeds toward the center with the central three residues being constructed by the Go and Scheraga (*16*) chain closure algorithm.

Selection of Backbone Construction Method

Where a canonical loop exists, it is used as the method of choice in the loop construction. Otherwise, construction may be based on data-base searching alone if a sufficient number of loops are found and loop length is in the range ($5 < n > 8$), data base plus *ab initio*, where the central s section of the data-base loop is reconstructed by CONGEN and ($n > 7$), or *ab initio* using conformational search alone (CONGEN) and the loop range is ($n < 6$).

Side-Chain Construction

Side-chain positions in canonical loops are, in general, well positioned. A template is used to overlay as many atoms as possible in the constructed side chain with the parent atom. Remaining atoms are reconstructed by either the iterative side-chain algorithms in CONGEN (*15*) or by a Monte Carlo simulated annealing (MC) algorithm (*6*). The MC method has recently become our method of choice in side-chain construction. Side chains from the data base, combined or *ab initio* backbone methods are completely reconstructed using either CONGEN or MC. The MC method uses a simple nonbond and torsional term as its objective function. This function is suitable for buried side chains in packed cores (*17,18*), but can result in many low-energy conformations on the surface. There is a need to take into account packing factors, especially with bulky hydrophobic residues, which, when found on the surface, are shielded by hydrophillic side chains. This effect may be incorporated into the objective function by a hydrophobic shielding term or as in AbM by accessing conformations after the simulation with a packing/accessibility screen.

Combining-Site Construction

The combining site is modeled by initially placing canonical loops onto the framework (*see 6*). Since these loops tend to be well conserved in their packing arrangements, this allows us to construct this part of the combining site with greater confidence. Side chains are then added to the canonical loops by the template-CONGEN/MC method described above. Loops derived from data-base construction are placed on the framework in isolation or with other CDR loops, and the central region reconstructed if required. These loops are ranked by an energy-screening procedure using a solvent modified Eureka (*19*) force field. Cluster analysis is performed on all loops, and the five lowest energy *different* conformations are selected for further filtering. A unique conformation is selected from these five conformations using structure determining regions (*6,15*). If a unique conformation is not found after filtering by SDR, then all low-energy conformations should be similar, and the lowest energy conformation can be chosen. Loops derived by CONGEN alone are screened using the solvent modified force field and the lowest energy conformation chosen.

Although AbM allows loops to be modeled in the presence or absence of other loops, we have found that modeling in the presence of other loops, with which we have confidence, restricts the phase space allowed for them to explore. This has the effect of preventing the possibility of low-energy conformations occupying space that would normally be denied them, resulting in better agreements with X-ray structures.

The aim of modeling is to produce antibody-combining sites of an accuracy that will enable us to make rational modifications to the combining site to improve properties, such as specificity and affinity, and to design new functionality, such as catalytic antibodies and metallo-enzymes. To do this, it is necessary not only to understand the structure of the combining site, but also the interaction of antibody with its antigen.

ANTIBODY RECOGNITION

Antibody recognition depends on the formation of a complex with antigen that is both topologically and energetically favorable. For macromolecular interactions, typically 1500 A^2 of surface area are buried on complex formation (*20*) and it is thought that the driving force of antibody–antigen interactions is largely the result of hydrophobic interactions with electrostatic interactions providing specificity (*21–23*). Macromolecular association is a dynamic process with changes in both entropic and enthalpic terms. Molecules move through solvent and loose rotational and translational entropy on complex formation, yet water displaced from the interface gains in entropy. Hydrogen bonds are both lost and formed. Salt bridges may be formed and side chains immobilized in the interface. It has been shown that the binding of antibodies to macromolecules is characterized by second-order rate constants in the range $2–4 \times 10^6$ $M^{-1} \cdot s^{-1}$ (*24*). Whether diffusion alone, lengthy collisions, and/or hydrophobic or electrostatic steering can account for the rapid association kinetics is still a matter of debate (*25*).

The Problem

The geometry of macromolecular association of static molecules is governed by six degrees of freedom, five rotational and one translational. Completely exploring the phase space available to these two static molecules is at present computationally intractable. Algorithms to predict docking of macromolecules or macromolecule/ligands have in the past tried to find ways to circumvent an entire exploration of phase space (*26–28*). Conformational flexibility has been reported in some antibody–antigen complexes (*29,30*) and would increase the already complex problem from 6 degrees of freedom to many thousands. Attempts have been made to introduce flexibility into algorithms, but only for very small systems (*31*).

For these reasons, most algorithms only attempt to solve the molecular recognition problem for static molecules. There are two major problems in developing a docking algorithm that is of general use. First, a small number of putative complexes must be generated within a reasonable time and within this ensemble must lie a candidate that represents a close approximation to the native complex. Second, these putative complexes must be evaluated in a manner such that correct orientations may be distinguished from incorrect orientations. We have developed a shape-independent topological analysis of macromolecules utilizing graph theory that fulfills these goals. The method is extensible to allow electrostatic, hydrophobic, hydrogen bonding, or other parameters to be evaluated in isolation or in concert. A flow chart of the algorithm is shown in Fig. 3.

The Algorithm

Proteins are in general asymmetric molecules, and any topological analysis that depends on locality of their surfaces must be independent of the overall shape of the molecule. It is possible to obtain shape independence by generating an equi-momental ellipsoid based on the moments of inertia of each molecule (*32*). Because it is desirable to remove atomic-scale bumps and to produce a surface that is capable of being differentiated, molecular surfaces are smoothed by rational B-splnes (*33*). Topological features, such as surface pits or surface peaks, can easily be identified by reference to a "sea-level" surface ellipsoid and higher-order features may also be identified (*34*). Other types of features may include energetic terms, such as hydrophobic or electrostatic potentials, mapped to the molecular surfaces. Such features may be represented as a graph where vertices represent topological features and edges the distance between them. Initial graphs from the antibody may be used to search for subgraph isomorphisms in the antigen graph framework, which covers the entire antigen surface (*35,36*). An example of such matching graphs can be seen in Fig. 4A, where a feature graph has been generated on the contoured molecular surface of the antibody-combining site from the HyHEL5 complex (*23*), and a corresponding subgraph isomorphism shown in Fig. 4B. Similarly, a simple graph comprising the strongest electrostatic potentials mapped to the molecular surface is shown in Fig. 4C and D. The extensibility of the method in its ability to combine graphs of different features is shown in Fig. 5, where electrostatic potentials and topological features are merged in the same graph. Putative isomorphisms are evaluated by energy screening and accessible surface burial. As an example, the above complex was subjected to the analysis, and the lowest energy structure obtained was predicted with an all-atom RMS deviation of 1.5 A from the crystal structure. A similar analysis has been reported on the D1.3 antibody and its antigen lysozyme with an RMS deviation of 1.4 A all atom from the crystal structure in a preliminary report (*39*). The algorithm is extremely fast with run times of minutes on a work station.

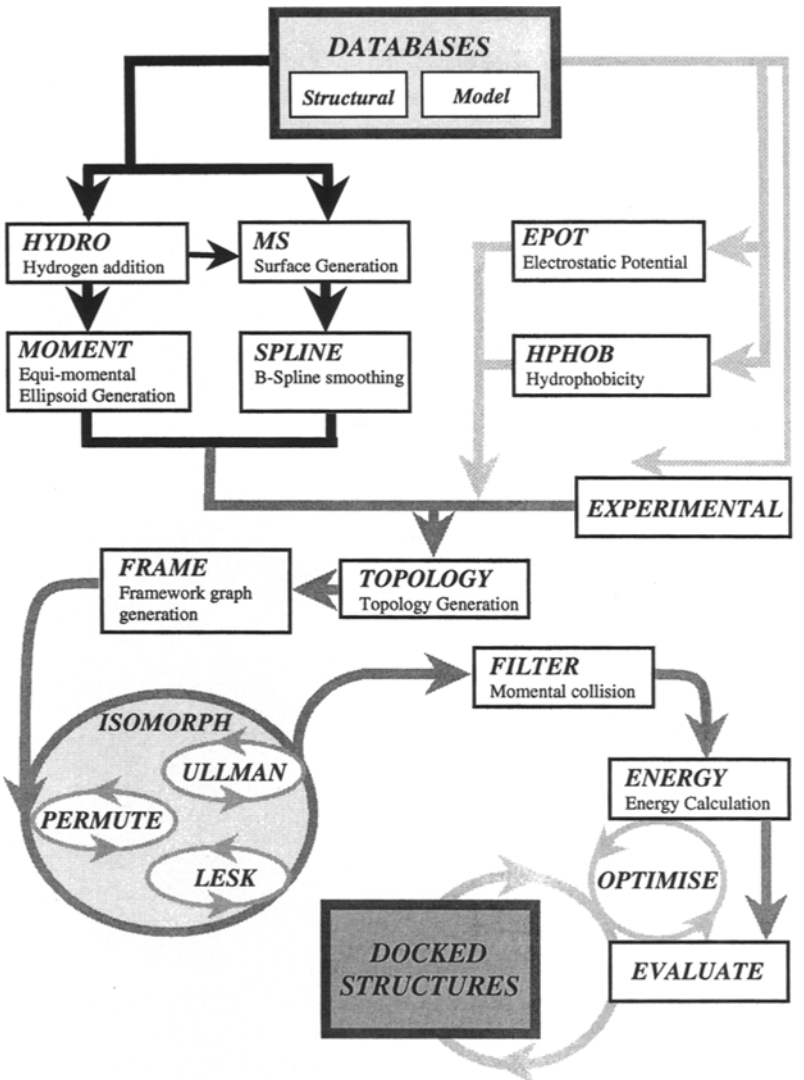

Fig. 3. Outline of the molecular recognition simulation procedure. Individual steps in the procedure are in bold italics.

The Future

We have recently completed the generation of a topological data base of the entire Brookhaven PDB. Such a data base may be used not only for docking experiments, but also as a vehicle for the study of crossreactivity, template design in the reverse engineering of unknown antigens, or drug design. We are in the process of developing a strategy for the replacement

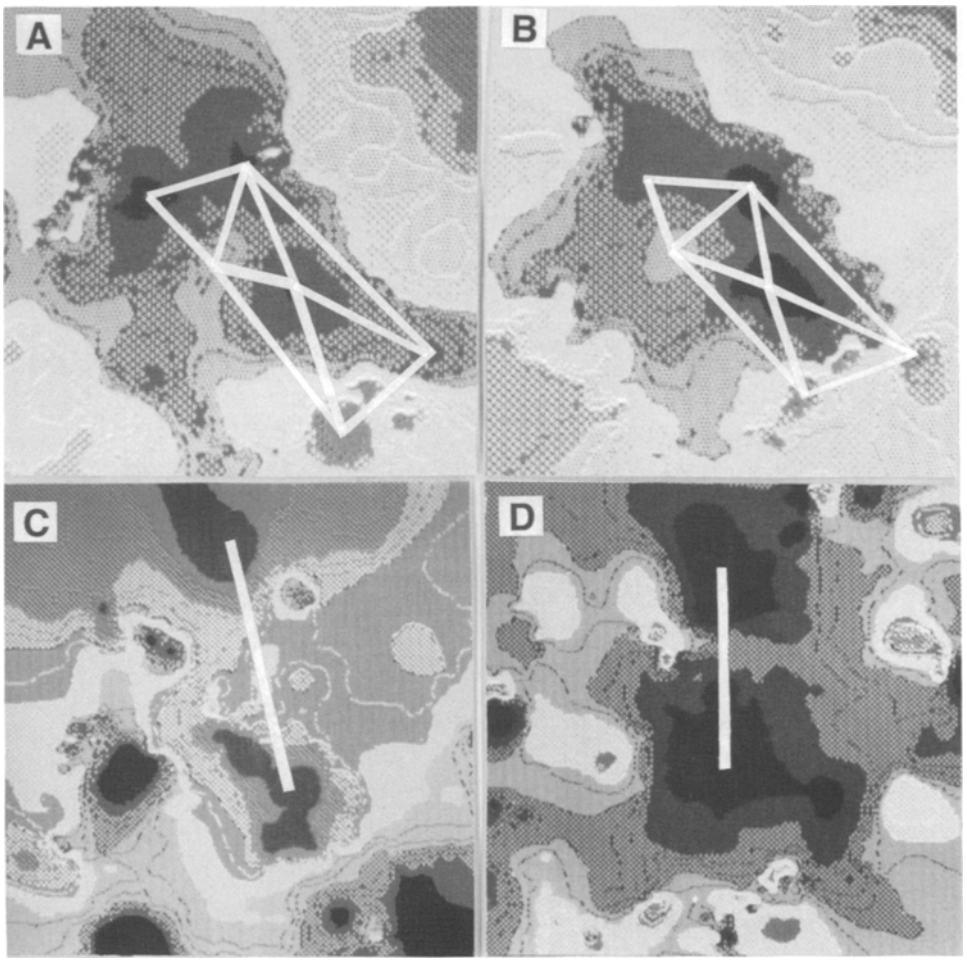

Fig. 4. Contour maps of the interface region of the antilysozyme complex HyHEL5. Topological maps (graphs) of significant points on the molecular surfaces (pits and peaks) are shown in (**A**) Hy5 and (**B**) HEL. The vertices of the graphs shown are the features that match each other, for example, a pit in Hy5 is filled by matching peak in HEL. The edge in the graphs represents the linear distance between features. Electrostatic potentials (*37,38*) contoured to the molecular surface are shown in (**C**) Hy5 and (**D**) HEL. We depict a simple graph of two vertices and one edge joining as shown in (**C**) the two largest positive electrostatic potentials and in (**D**) the two largest negative electrostatic potentials.

of antibodies by nonpeptide mimetics, using a combination of combining-site design and molecular docking algorithms using feature templates. If such mimetics exhibited high specificity and affinity for their targets, they would provide a credible alternative to antibodies, which, if exploited in pharmaceutical applications, could result in a new generation of therapeutic drugs.

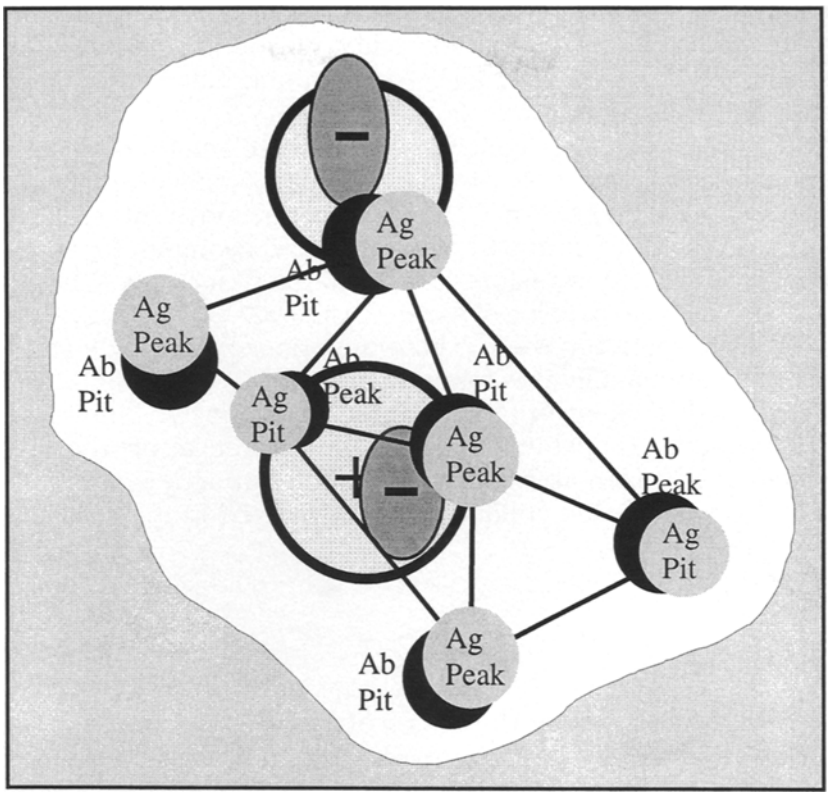

Fig. 5. Schematic representation of the graphical relationshipo between the topological features of HyHEL5 at the interface and the electrostatic potential features as depicted from Fig. 4. To simplify the diagram, only the antibody topographic features (vertices) are shown in the form of a graph and their adjoining edges. The schematic also depicts the strongest electrostatic potentials taken from Fig. C and D, and these too may be linked in a single topographic/electrostatic potential graph.

MODELING OF METAL BINDING SITES

Our ability to extend the natural limits of antibodies relies not only on their ability to model combining sites, but also on our ability to design novel functions into those combining sites in such a manner as to fulfill the required design specifications. An example of such design has been the creation of a metallo-antibody that exhibited allosterism between the antibody-combining site and the metal binding (*40*). The incorporation of a metal binding site into an antibody may result in a perturbation of function of one or both. Our model system was the HyHEL5 antibody, which forms a complex with hen egg lysozyme (HEL) (23). Zinc was chosen as

the metal, since it exhibits hard/soft behavior and provides greater flexibility in the choice of liganding amino acids. Ligand geometry and optimal ligand length were obtained from crystallographic data of small molecules and metallo-protein structures.

Although all six CDRs contribute to the free energy of stabilization, the amino terminal portion of CDR L1 contributes unfavorable interactions and does not have any critical CDR-CDR interactions. This section of the L1 loop was chosen to form part of a tetrahedral geometry for metal binding along with an N-terminal residue from the framework. The residues from L1 were mutated as follows: Ala25→Cys, Ser27→Glu, Val 29→Cys, and Ile2→His from the framework. These mutations were made and the Fv purified. The mutated Fv was found to bind Zn(II), Co(II), and Cd(II) and to retain its high-antigen affinity and specificity. The two binding sites exhihited negative cooperativity, and the presence of these metals modulated the antigen association constant by up to a factor of 10. This allosteric behavior of the antibody may be utilized in the design of new and novel biosensors.

CONCLUSION

By using algorithms for the design of antibody-combining sites, with the docking and template algorithms and our growing knowledge of combining-site formation and recognition processes, we have attempted to show that progress has been made in our understanding of ways to extend the natural limits of antibody-combining sites. We are already addressing the problems of flexibility, the design of metal binding sites, of catalytic antibodies, of combining sites with completely new functions, the reverse engineering of unknown antigens, the replacement of antibodies by nonpeptide mimetics, and the humanization of mouse monoclonal antibodies. These and other applications will become amenable as we combine multidisciplinary algorithms to enable us to exploit more fully the abundant repertoire of possible conformations that antibodies may adopt.

REFERENCES

1. Wu, T. and Kabat, E. (1970), *J. Exp. Med.* **132,** 211.
2. Lesk, A. M. and Chothia, C. (1982), *J. Mol. Biol.* **160,** 325.
3. Chothia, C. and Janin, J. (1981), *Proc. Natl. Acad. Sci. USA* **78,** 4146.
4. Novotny, J., Bruccoleri, R. E., Newell, J., Murphy, D., Haber, E., and Karplus, M. (1983), *J. Biol. Chem.* **258,** 4433.
5. AbM: Antibody modeling. An automated program based on CAMAL and marketed by Oxford Molecular Ltd., Oxford, UK, and Oxford Molecular Inc., Palo Alto, CA.

6. Pedersen, J., Searle, S., Henry, A., and Rees, A. R. (1993), *Immunomethods* **1,** 126.
7. Darsley, M. J., de la Paz, P., Phillips, D. C., Rees, A. R., and Surtton, B. J. (1985), in *Investigation and Exploitation of Antibody Combining Sites, Methodological Surveys in Biochemistry and Analysis,* vol. 15, Reid, E., Cooke, G. M. W., and Morre, D. J., (eds.), Plenum, New York, p. 63.
8. de la Paz, P., Sutton, B. J., Darsley, M. J., and Rees, A. R. (1986), *EMBO J.* **5,** 415.
9. Chothia, C. and Lesk, A. M. (1987), *J. Mol. Biol.* **196,** 901.
10. Chothia, C., Lesk, A. M., Tramontano, A. M., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M., and Poljak, R. J. (1989), *Nature* **342,** 877.
11. Steipe, B., Pluckthun, A., and Huber, R. (1992), *J. Mol. Biol.* **225,** 739.
12. Martin, A. C. R., Cheetham, J. C., and Rees, A. R. (1989), *Proc. Natl. Acad. Sci. USA* **86,** 9268.
13. Martin, A. C. R., Cheetham, J. C., and Rees, A. R. (1991), *Methods Enzymol.* **203,** 121.
14. Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977), *J. Mol. Biol.* **112,** 535.
15. Bruccoleri, R. E. and Karplus, M. (1987), *Biopolymers* **26,** 137.
16. Go, N. and Scheraga, H. A. (1970), *Macromolecules* **3,** 178.
17. Lee, C. and Levitt, M. (1991), *Nature* **352,** 448.
18. Lee, C. and Subbiah, S. (1991), *J. Mol. Biol.* **217,** 373.
19. Viner, R. C. (1989), Ph.D. thesis, University of Bath, UK.
20. Davies, D. R., Padlan, E. A., and Sheriff, S. (1990), *Annu. Rev. Biochem.* **59,** 439.
21. Novotny, J. (1991), *Mol. Immunol.* **28,** 201.
22. Janin, T. and Chothia, C. (1990), *J. Biol. Chem.* **265,** 16027.
23. Sheriff, S., Siverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., and Davies, D. R. (1987), *Proc. Natl. Acad. Sci. USA* **84,** 8075.
24. Ward, E. S., Gussow, D., Griffiths, A. D., Jones, P. T., and Winter, G. (1989), *Nature (Lond.)* **341,** 544.
25. Northrup, S. H. and Erikson, H. P. (1992), *Proc. Natl. Acad. Sci. USA* **89,** 3338.
26. Kuntz, I. D., Blaney, J. M., Oatley, S. J., Langridge, R., and Ferrin, T. E. (1982), *J. Mol. Biol.* **161,** 269.
27. Walls, P. H. and Sternberg, M. J. E. (1992), *J. Mol. Biol.* **228,** 277.
28. Pellegrini, M. and Doniach, S. (1993), *PROTEINS: Structure, Function and Genetics* **15,** 436.
29. Bhat, T. N., Bentley, G. A., Fischmann, T. O., Boulot, G., and Poljak, R. J. (1990), *Nature (Lond.)* **347,** 483.
30. Rini, J. M., Schulze-Gahmen, U., and Wilson, I. A. (1992), *Science* **255,** 959.
31. Leach, A. R. and Kuntz, I. D. (1992), *I. Comp. Chem.* **13,** 730.
32. Taylor, W. R., Thornton, J. M., and Turnell, W. G. (1983), *J. Mol. Graphics* **1,** 30.
33. Colloc'h, N. and Mornon, J. P. (1990), *J. Mol. Graphics* **8,** 133.
34. Haralick, R. M., Watson, L. T., and Laffey, T. J. (1983), *Int. J. Robotics Res.* **2,** 50.

35. Lesk, A. M. (1979), *Commun. ACM* **22,** 219.
36. Ullman, J. R. (1976), *J. ACM.* **23,** 31.
37. Warwicker, J. and Watson, H. C. (1982), *J. Mol. Biol.* **157,** 671.
38. Rogers, N. K. and Sternberg, M. J. E. (1984), *J. Mol. Biol.* **174,** 527.
39. Webster, D. M. and Rees, A. R. R. (1993), *Miami Short Reports* **3,** 94.
40. Staunton, D., Jones, A. E., and Rees, A. R. (1993), *Protein Eng.* **6,** 93.

DISCUSSION

D. Webster

Hansen: How will the fact that either the antibody or the antigen often changes conformation on binding affect this overall procedure, given, I assume, that your ultimate goal is to try to do everything *de novo* and dock things together without a crystal structure to guide you.

Webster: With the docking, obviously if we have flexibility in the molecules, that is going to complicate the problem. The algorithm can cope with a certain amount of flexibility because we can alter the edge parameters for the allowed tolerances. The problem really is in evaluating those putative structures, because if we have some flexibility, we get overlapping of side chains and clashes in the energies go very high. That is obviously a problem, but possibly one that is not as great as a lot of people think it might be. Analysis of the side-chain positions of all the eight hen lysozymes plus the three from complexes has revealed very little flexibility in the residues that are actually involved in the interface of the two molecules. Having said that, there is some recent evidence that there is some flexibility. In some antibody–antigen complexes, for example, study of D1.3 in Poljak's lab has indicated that the root mean square deviation between the docked and the undocked complex may be about 1.5–2 A. The algorithm could easily cope with that amount of flexibility. However, recently I spoke to Peter Colman, and he said that in the case of NC41, the antineuraminidase, there is a very large difference in the angle between the V_L and the V_H between the docked and the undocked structure of around 6 or 7° and that the topology on the surface is quite different. Obviously, we would not be able to account for that. Docking experiments rely on static structures. We are working on ways to get around this. For example, you can use soft potentials in your evaluation functions. Also, if you do an analysis of rotamer space, you can identify potential flexible side chains and eliminate those from the evaluation. These points are just ignored in constructing the putative complex because you have other significant features that are invariant. If you obtain a large change in the backbone, that is a completely different problem for which nobody has an answer at the moment.

Hansen: It seems it is going to be case-by-case evaluation of the validity of the structures.

Webster: Yes.

Paul: The algorithms you use are dependent on crystal structures. Do *ab initio* approaches yield similar results? Is it reasonable to say that crystal structures reflect solution conformation of antibodies found at much lower physiological concentrations?

Webster: *Ab initio* approaches are not the answer. We do less well with *ab initio* approaches than we do using a canonical data-base or combined approach. We have done experiments where we have constructed all the side chains using canonical loops, L1–H2, then a normal construction for H3, and then a complete CONGEN construction of L1–H2. We found that the root mean square deviations for those differ by about 1 A. So, for the canonical loops, we would typically get under 1 A deviation, backbone, all atom, for those loops. Add on another 1 A for *ab initio* approaches. In the cases of noncanonical structures, we do combine the data-base approach with an *ab initio* approach, in that we use what we can from the data base to construct the bottom of the loops, and then we construct the central portion. It is only for very small loops that we can just use a straight *ab initio* approach. Certainly, we obtain useful information from the data base, and that is why we use the combined approach.

Paul: Is there independent evidence that the crystal structures, which determine the outcome of molecular modeling, are indeed correct?

Webster: That is really something that should be directed toward a crystallographer. Obviously we know there are crystal packing effects in crystals. The solution structures may be different. I would assume that there are more constraints within crystals, and therefore, you will see some differences in solution.

Paul: The suggestion is that since *ab initio* approaches are not dependent on crystal structures, they may yield structures that more closely reflect the solution structures.

Webster: I see what you mean. Yes, but we have to evaluate the structure somehow. We say we have a correct structure in relation to a crystal structure. Whether that is a true structure is a different question, and that is something that we cannot answer at the moment because we do not have that information.

Tramontano: You showed a slide where the electrostatic interactions were clearly matched in the docking experiment. My question is whether the sequence information shows that these electrostatic sites are located on a particular CDR loop or whether they occur anywhere in the combining site.

Webster: This is not something I have really looked at closely. I do know that you do get certain specificities, for example with some anti-DNA

antibodies in which the charge may tend to localize in the loops in H3, because of a frame shift in the germline.

Wu: If you look at the charges in the CDRs, these seem to be all over the place.

Tramontano: There is no localization?

Webster: The only time that I have seen localization is with the anti-DNA antibodies. You always find a high proportion of charge in the CDR loops.

Tramontano: in CDRH3?

Webster: Yes.

Wu: There is no special pathology in the case of these antibodies.