J. theor. Biol. (1994) **169**, 391–402

Rethinking "Shape Space": Evidence from Simulated Docking Suggests that Steric Shape Complementarity is not Limiting for Antibody–Antigen Recognition and Idiotypic Interactions

JORGE CARNEIRO† AND JOHN STEWART

Unité d'Immunobiologie, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris, France

(*Received on* 30 *November* 1993, *Accepted in revised form on* 25 *March* 1994)

The concept of "shape space" is based on the assumption that the relevant properties of individual molecules can be adequately specified by a finite list of N parameters; and that c_{ij}, the affinity between molecules i and j, can be specified by an equation of the form: $c_{ij} = f(x_i, x_j)$, where x_i and x_j are N-dimensional vectors representing the absolute positions of molecules i and j in an objective, referential "shape space", and f is an appropriate function. We have performed simulated docking of the combining sites of immunoglobulin molecules, based on their crystallographic structures. The results suggest that shape complementarity cannot account for the specificity of idiotypic interactions, since in the simulations each pair of docked proteins had a buried surface area as great as that occurring in known complexes. It therefore seems likely that the atomic interactions accounting for the specificity of immunoglobulin recognition are highly relational. This casts doubt on the basic assumptions underlying the shape-space concept, at least in the simple form hitherto used in theoretical modelling of the immune system. In order to be realistic, the dimensionality N would have to be high (more than 20), and the function f would be irregular and discontinuous. Alternatively, if the equation $c_{ij} = f(x_i, x_j)$ is interpreted as a purely formal construction in an abstract "inversion space", its validity is entirely relative to the empirical affinity matrix on which the construction is based. We conclude that at present there is no sure way of adequately characterizing the internal structure of idiotypic affinity matrices; and that models of the immune system should therefore aim at being generic and robust with respect to the structure of the idiotypic affinity matrices of unselected immunoglobulins.

Introduction

In the field of theoretical immunology, there is a rather widespread conviction that antibody–antigen specificity, and protein–protein recognition in general, involves some kind of steric complementarity between particular regions on each of the interacting molecules (Perelson, 1989; Stewart & Varela, 1989; Kazatchkine & Coutinho, 1993). This notion is not just a trivial reminiscence of Ehrlich's early "side-chain" theory (Ehrlich, 1901), but seems to have some foundation in modern biophysics. The non-covalent interactions that are believed to lead to an enthalpic stabilization of protein complexes are "weak" short-range interactions, notably, van der Waals forces, hydrogen bonds and interactions between ionic charged groups (Novotny *et al.*, 1989; Pellegrini & Doniak, 1993). Such interactions can only become effective if the chemical groups involved actually approach each other closely, and *a priori* this would be expected to require global shape complementarity. Similarly, the so-called hydrophobic effect, i.e. the entropic stabilization of the complexed state by removal of water molecules from the interface area (Sharp *et al.*, 1991), clearly implies a substantial degree of steric complementarity in order to avoid empty spaces that could be filled by water (Cherfils *et al.*, 1991). The early reports on "low" resolution X-ray structures of antibody–antigen and protease–inhibitor complexes reinforced the latter

† Author to whom correspondence should be addressed.

0022-5193/94/160391 + 12 $08.00/0

view, since the bound molecules showed very small changes in main chain conformation as compared to their free forms, and a very "tight" complementarity between the partners was suggested by the apparent absence of water in the more or less extensive interface region [with around 1500 Å^2 of solvent-accessible area being buried upon binding (Davies & Padlan, 1990)]†. The notion that shape complementarity is necessary for binding is so well established that induced fit is commonly seen as a major conformational change undergone by molecules in order to assume complementary three-dimensional profiles.

THE "SHAPE SPACE" CONCEPT

However, if the hypothetical notion that molecular affinity is due to shape complementarity has been transformed into a widespread and largely unexamined assumption, this is not due to direct stereochemical evidence alone. A major role has been played by the theoretical concept of "shape space": initially advanced by Perelson & Oster (1979), this seminal concept has constituted ever since one of the most powerful tools to organize modelling in theoretical immunology (Segel & Perelson, 1988; Perelson, 1989; DeBoer *et al.*, 1992). The basic idea is that the features that determine the combining properties of any possible epitope‡ with any other can be specified by N parameters, where N is a fixed and limited number. Any given epitope could thus be represented by a point whose co-ordinates are the values of each of those parameters in an N-dimensional space which was called "shape space". The great attraction of the shape-space concept stems from the fact that if the number N of parameters necessary and sufficient to completely describe each and every possible epitope is finite and relatively small, then properties such as similarity and mutual affinity between molecules are readily derived from the relative positions of the points that represent them in a low-dimensional shape space.

As it is used in modelling, the shape-space concept can be expressed mathematically as follows:

$$c_{ij} = f(\boldsymbol{x}_i, \boldsymbol{x}_j), \tag{1}$$

where c_{ij} is the affinity between molecules i and j, $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ are N-dimensional vectors representing the positions of molecules i and j in shape space, and f is an appropriate function. This equation is open to two quite distinct interpretations, and we consider that it is most important not to confuse them.

One interpretation, which we present first because it is easier to understand and thus has heuristic value, is what we shall call a "realistic" shape space. On this view, the vectors $\boldsymbol{x}$ describe quite specific, identifiable features of the epitope that are amenable to direct experimental measurement. In their original paper, Perelson & Oster (1979) explicitly propose as an illustrative example the geometrical parameters that describe the surface of the epitope (hence the appellation "shape" space). In a subsequent article referring to the shape-space concept, Perelson built on a suggestion already included in the initial paper, and proposed a "generalized shape" including distributions and intensity of charges, potential hydrogen bonds, complementary chemical groups, etc. (Perelson, 1989). The characteristics and three-dimensional distribution of all the chemical groups which are potentially able to form local interactions with other groups or solvent molecules can, at least conceptually, be parametrized into this "generalized shape", and ascribed to one or more dimensions of the shape space. Nevertheless, it should be noted that the notion according to which protein binding specificity relies on some sort of steric complementarity (not just geometrical complementarity but also a physicochemical one), and that the binding properties can be inferred from this "generalized shape" which is unique and well defined for any given epitope, remains basic to all such "realistic" interpretations of the shape-space concept that have been envisaged to date.

Unfortunately, these "realistic" interpretations suffer from two major drawbacks. First, in the present state of stereochemical knowledge, the implicit goal of actually identifying the N parameters in the vectors $\boldsymbol{x}$, and then specifying the function f by a "bottom up" approach based on a solid understanding of the actual mechanisms involved, is totally remote. Second, the very difficulty of this task, coupled with the stereochemical knowledge that we do have, overwhelmingly suggests that in any realistic model, the dimensionality of N would be high (at the very least of the order of 10); worse, many of the parameters involved would not be simply geometrical;

† At present, this view may be changing as higher resolution refinements of X-ray structures seem to reveal the presence of very precisely organised hydrogen-bonded water molecules in the buried interface area (R. Mariuza, conference at the Pasteur Institute, 1993), but we will come back to this later in the discussion.

‡ In this particular context, and for the following discussion, we provisionally define an "epitope" as simply being a discrete region of one molecule that is involved in binding to other molecules. This definition regroups, without distinction, such categories as "antigenic determinants", "combining regions", "idiotopes", "paratopes", "catalytic site", etc, since these are purely "functional" categories which are related not to the structural basis of affinity but to the molecules they are part of in particular situations (respectively antigens, antibodies, antibodies when recognizing antigen, antibodies when being recognized by other antibodies, enzymes, etc).

and worse still, the function f would be highly complex, irregular and discontinuous.

In view of these difficulties, it is worth examining an alternative interpretation of eqn (1), which is also suggested in Perelson & Oster (1979). In order to distinguish this alternative approach clearly from the "realistic" interpretation, it will be convenient to rewrite the equation as:

$$b_{ij} = g(y_i, y_j). \tag{2}$$

For any given set of immunoglobulins ($i = 1, \ldots n$), under physiological conditions of ionic composition and temperature of the aqueous medium, there will exist *empirically* a definite affinity matrix b_{ij}. The construction of a set of vectorial representations (the y_i) and a function g that satisfy eqn (1) can now be considered as a *purely formal* operation, belonging to the class of problems known as "reverse engineering". Generically, such problems have the awkward property that there is no way of logically deducing a unique optimal solution: if approximate solutions are acceptable (as they are here), then *as a matter of principle* there are always an indefinite number of qualitatively different solutions. This being so, the usual procedure is to *assume*, more or less arbitrarily, a functional form for g that has desirable properties, and then to optimize the construction of the vectors y and the parameters in g. We propose to label the M-dimensional space of the vectors y in such a construction an "inversion space".

What is the relation between an "inversion space" of this sort and a "realistic shape space"? The major conceptual difference is that a realistic shape space is grounded in knowledge of the actual underlying biophysical mechanisms, whereas an inversion space is merely a tool for providing an economical description of the empirical matrix b_{ij} which was taken as the starting-point for its construction. The N dimensions of the vectors x in a realistic shape space constitute an absolute domain of description of objective, intrinsic properties of individual molecules. A suitable quasi-random distribution of points in a realistic shape space would therefore provide a reasonable basis for generating simulated affinity matrices for the immunoglobulins of unselected B-lymphocyte clones as these are generated by the bone-marrow. The M dimensions of the vectors y in an inversion space, by contrast, are *entirely abstract* and *a priori* bear no necessary relation to identifiable biophysical properties of individual molecules. An inversion space is also entirely relative to the particular set of immunoglobulins and the empirical matrix, b_{ij}, on which its construction is based. Thus, not only is the distribution of points in an inversion space unlikely to be random, but the very structure of the space will reflect any processes such as repertoire selection (or co-selection) which may have affected the original set of immunoglobulins. An inversion space based on limited empirical data cannot therefore be reliably used to generate simulated affinity matrices of unselected immunoglobulins.

It should now be clear that a "realistic shape space" and an "inversion space" are indeed conceptually distinct entities that should not be confused. The great popularity of the shape-space concept, as it has been used both by ourselves and others in theoretical modelling of the immune system, is based on the major assumption that these two spaces do in fact coincide, so that it is legitimate to pass freely from one to the other as convenient. Thus, it has been widely assumed that the mutual affinity between "shapes" A and B can be calculated by a simple Gaussian metric based on the distance $\|A - B'\|$ where B' is the symmetrical image of B through some point in shape space (De Boer *et al.*, 1992; Detours, personal communication). The simplicity of this assumption could be justified as an *a priori* choice for g in the context of the construction of an inversion space, except that it has not been shown that this functional form can provide an acceptable approximation to any empirical affinity matrix b_{ij}; conversely, this absence of reference to substantial data on empirical affinity matrices (which are indeed lacking), and the assumption of a random distribution of unselected epitopes in the space in question, could be justified in the context of a realistic shape space. The assumption is highly attractive because relations of similarity and mutual affinity between epitopes then correspond to simple topological relations in the "shape space" in question, so that it is possible to *visualize* the extent and the internal structure of the immune repertoire as it emerges during the course of (simulated) ontogeny. Unfortunately, however, these appealing features do not in themselves provide any guarantee that the underlying assumptions are valid. The danger is obvious: the heuristically stimulating visions of the organization of the immune system based on the shape-space concept may be illusions based on little else than wishful thinking.

Precisely because of its importance, we think that the time may have come to proceed to a lucidly critical examination of the assumptions on which the shape-space concept is based. The major assumption—that a realistic shape space and an inversion space do in fact coincide—is surely highly optimistic; it is not, however outrageously absurd. In particular, it could conceivably be justified if specific affinity really does derive from geometrical shape

complementarity between the two molecules concerned. The aim of this paper is to examine this possibility in the light of a multitude of recent reports on the simulation of protein binding specificity using "docking algorithms" based on shape complementarity (for review see Cherfils & Janin, 1993).

SIMULATED PROTEIN DOCKING

As Connolly (1986) has pointed out, protein docking takes the three-dimensional structures of two proteins as given, and addresses two related but separate questions. The first question is to predict *whether* the two proteins associate; in the event that they do, the second question is to predict the actual structure of the complex.† Most authors working in the field of protein docking are stereochemists who have concentrated on the second question, i.e. the reconstitution of known complexes in the configuration found in X-ray crystals, by docking the two involved proteins using their bound or free structures (Cherfils & Janin, 1991; Jiang & Kim, 1991; Shoichet & Kuntz, 1991; Connolly, 1992). Such calibration of the algorithms (Cherfils & Janin, 1993) is a reasonable first step, and makes it possible to estimate the relative importance of the criteria used for evaluating the docked structures (Connolly, 1992). However, a more general understanding of specificity also requires an answer to the first question: given a set of protein structures, it should be possible to predict which pairs can associate and form stable complexes and to discriminate the latter from those complexes that are thermodynamically unstable and simply will not be formed. The detailed work to be presented in this paper is designed to complement the available bibliography on protein-docking by examining this general question.

Using the well-calibrated Wodak–Janin docking algorithm, we therefore obtained a square pseudo-stabilizing energy matrix from a set of antibodies (D1.3, E225, McPC603, HyHEL10, J539) and another protein [hen egg lysozyme (HEL)], for which the X-ray structures are known. Since it is well established that only some of these proteins actually cross-react to form stable complexes (whose structure is also known), this seemed to be a good way to evaluate to what extent shape complementarity in the strict sense can account for the specificity of protein–protein recognition. The results presented here are not intended to be an exhaustive study on this matter, but together with the results of others they raise serious questions as to the validity of the assumption that specific affinity is due to shape complementarity. In conclusion, we shall discuss the implications of these findings for the shape-space concept and theoretical modelling in immunology.

† The heuristic value of the shape-space concept is precisely that it allows theoretical immunologists to avoid involvement with questions of this sort by presupposing a computationally convenient answer.

Methods

The atomic co-ordinates for the molecules and complexes used in this work are from X-ray structures, and were obtained from protein data bank (PDB) files or from the authors [D1.3 and E225 complex (Bentley *et al.*, 1990)]. "Simplified protein" models are obtained from the X-ray structures by replacement of each amino acid residue by a sphere centered on the center of gravity of its side chain, including the Cα. The radii of the sphere are from Levitt (1976). Docking is performed on these simplified protein models.

The DOCK package was used, which was a kind gift from J. Janin and J. Cherfils. The procedure has been described in detail previously in Cherfils *et al.* (1991) and we will present it here in outline only. Like any other docking algorithm, DOCK proceeds by bringing two molecules into contact and attributing a score to the contact. Each molecule is kept rigid; under these conditions, the pair has only six relational degrees of freedom.

The six rigid-body parameters that define the position and orientation of one molecule relative to the other are five angles and a distance (Fig. 1): θ_1 and φ_1 locate the centre O_2 of molecule 2 relative to molecule 1; θ_2 and φ_2 do the same for O_1; χ is a spin angle about the centre line; and r is the centre-to-centre distance. All angles are set to zero in the

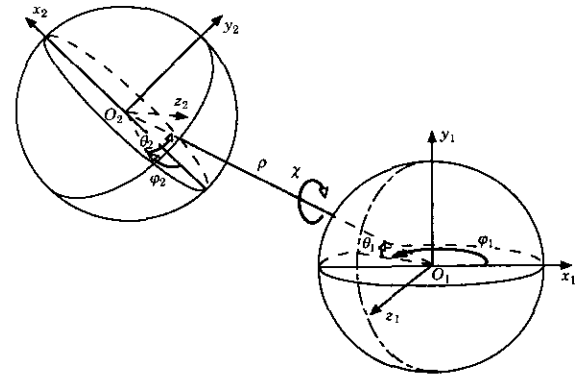

FIG. 1. The degrees of freedom in rigid-body docking. ρ is the distance between the centres of mass of the two proteins O_1 and O_2. φ_1 (φ_2) and θ_1 (θ_2) are the longitude and latitude of O_2 (O_1) in the reference frame of molecule 1 (2). χ is the spin angle that establishes the relative orientation of the molecules along the axis defined by the centres O_1 and O_2.

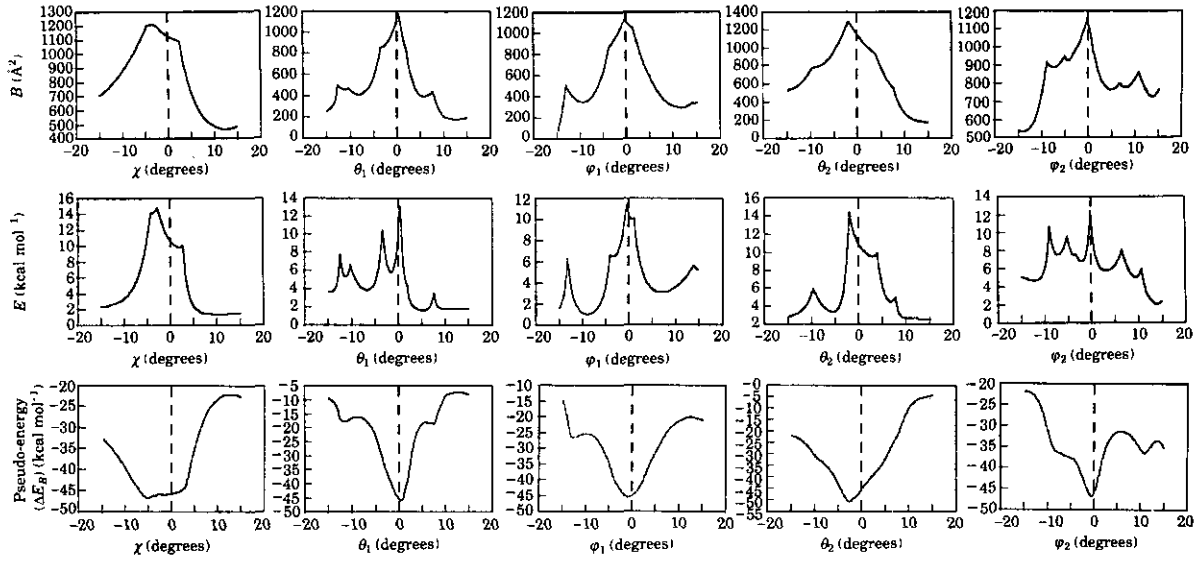

FIG. 2. Single-dimensional exploration of the parameter space for the complex E225–D1.3. Each angle θ_1, φ_1, θ_2, φ_2, or χ was explored individually in a range of -15 to $+15$ around the orientation found in the X-ray crystal structure (0,0,0,0,0).

orientation found in X-ray structure of the D1.3–E225 complex. The zero position for "artificial" complexes was obtained by superposing the Cα from the framework region (Chlothia & Lesk, 1987) of each Fv on those of D1.3 as positioned in the complex with E225, using a least-square fit by Kabsch (1976).

Given a set of angles θ_1, φ_1, θ_2, φ_2 and χ, the two molecules are docked by translation along the line of centres. The translation brings the two molecules as close as possible with no overlap between residues. The condition for overlap is defined by:

$$d/(r_i + r_j) < s,$$

where r_i and r_j are Levitt radii and d is the inter-residue distance. The parameter s renders the spheres "soft": this is an approximation that is assumed to account for the flexibility of the side chains. Janin and colleagues (Cherfils *et al.*, 1991) have calibrated this parameter and shown that a value of 0·75 is appropriate.

The score function to evaluate the docked complexes is a pseudo-energy ΔE_B, calculated as a function of two quantities: an approximate interface area, B, and an approximate repulsive energy, E. B is an underestimate of the surface area buried within a complex, and equals the sum of the solvent accessible surface areas of the two free components minus that of the complex, calculated with an analytical approximation (Wodak & Janin, 1980). E is an estimate of molecular overlaps calculated from the repulsive components of Levitt's soft 6-8 potentials (Levitt, 1976). E and B are brought to the same scale by:

$$\Delta E_B = E - \gamma B,$$

where the constant γ converts surface area to energy (like a surface tension). Experimental values for γ (Chlothia, 1974; Sharp *et al.*, 1991) range from 25 to 47 cal mol^{-1} Å^{-2}. We used $\gamma = 50$ cal mol^{-1} Å^{-2} as suggested by previous work, and we confirmed that smaller values give insufficient weight to B.

The five-parameter space is explored by Monte Carlo simulated annealing using the Metropolis method (Metropolis *et al.*, 1953), in a range of $-10{\cdot}0$ to $+10{\cdot}0$ degrees for each angle θ_1, φ_1, θ_2 or φ_2, and between -180 to $+180$ degrees for χ. Simulated annealing follows a random iterative path in parameter space, going downwards in energy ΔE_B. To avoid freezing in local minima upward jumps are also allowed but with a probability $p(T)$ that depends on the pseudo energy step δE_B:

$$p(T) = \exp(-\delta E_B/kT)$$

The initial temperature is set so that $p(T)$ is around 0·5 and is lowered progressively in steps of 10%, with either 100 successful jumps or 5000 orientations tested at each temperature. Each simulation includes 18 temperature steps. When the success rate falls to less than 1%, a restart is made by choosing a new random conformation.

Results

In order to confirm the effectiveness of the docking procedure considered here we studied its performance on the E225–D1.3 idiotype–anti-idiotype complex. As

can be seen in Fig. 2, by changing each parameter individually a minimum of the pseudo-energy ΔE_B (a maximum of buried area B) can be found very close to the real complex configuration (zero). Moreover, the buried surface area is highly sensitive to changes in each and every parameter, clearly demonstrating the capacity of the algorithm to discriminate effectively between regions showing different degrees of shape complementarity.

The orientation space of the pair E225 and D1.3 was then explored by simulated annealing in a limited volume corresponding to the combining regions of the two immunoglobulin molecules. In such relatively restricted volume, ranging from $\pm 10°$ around the zero position for angles θ_1, φ_1, θ_2 and φ_2, and $\pm 180°$ for χ, we found 23 different configurations characterized by a buried surface area greater than 1000 Å^2. In Table 1, we list only the ten configurations with a pseudo-energy ΔE_B lower than 55 kcal mol^{-1}. We would like to note that amongst the configurations that show extensive shape complementarity, a configuration can be found that closely resembles the actual X-ray structure. However, the corresponding value of ΔE_B or B is just a median one, and alternative orientations lead to improved interface complementarity. Similar observations have been repeatedly reported by other authors on docking of other antibody–antigen or protease–inhibitor complexes: a "pseudo-stable" configuration more or less close to the real X-ray structure can be found, but multiple "false positive" configurations are also consistently found.

We then searched for complementary regions in molecules that are known not to associate. In a first step, the combining region of E225, or the one from D1.3, were docked onto themselves (it is known that neither of these immunoglobulins are "autobodies"). The search was limited to the same volume of the parameter space used previously to study the real pair E225–D1.3. The "auto-docking" of both E225 or D1.3 in this restricted parameter range revealed the existence of several regions showing excellent values of shape complementarity as compared to real complexes (Table 1). In the case of the artificial pair E225–E225 the situation is dramatic with the number of "false positive" configurations exceeding the one found in the real pair E225–D1.3. We would like to emphasize that these "false positive" configurations do not simply result from malfunction of the docking algorithm. As illustrated in Fig. 3, the algorithm accurately selected the relatively few "false positive" configurations amongst the very large number of random configurations explored.

This striking result was consistently reproduced when we extended the study to three additional immunoglobulin molecules and the protein HEL. In Table 2 we present a pseudo-energy (ΔE_B) matrix between D1.3, E225, HyHEL10, J539, McPC603 and HEL. The docking was restricted to the immunoglobulin combining site, or the region recognised by D1.3 in HEL, using always the same parameter range. It is notable that extensive shape complementarity was systematically found between each and every pair of proteins, even though the study was restricted to a relatively small region of their surfaces. A curiosity is

TABLE 1

Artificial complexes from docking by simulated annealing†

	Rank	E (kcal mol^{-1})	B (Å^2)	ΔE_B (kcal mol^{-1})
fvE225–fvD1.3				
	X-ray crystal‡	−10·76	1127·3	−45·6
	1	10·76	1446·6	−59·3
	2	10·98	1379·5	−58·0
	3	7·84	1306·8	−57·5
	4	7·76	1283·2	−56·4
	5	9·26	1307·3	−56·1
	6	9·62	1308·4	−55·8§
	7	11·64	1344·7	−55·6
	8	10·70	1326·1	−55·6
	9	11·29	1337·7	−55·6
	10	13·45	1371·0	−55·1
fvE225–fvE225				
	1	11·35	1557·0	−66·5
	2	12·94	1584·7	−66·3
	3	11·90	1545·9	−65·4
	4	12·27	1413·3	−58·4
	5	12·83	1422·6	−58·3
	6	12·41	1398·2	−57·5
	7	10·18	1347·6	−57·2
	8	8·54	1308·8	−56·9
	9	9·78	1329·5	−56·7
	10	10·55	1319·5	−55·4
	11	6·6	1238·0	−55·3
	12	8·84	1278·8	−55·1
fvD1.3–fvD1.3				
	1	5·93	1574·6	−72·8
	2	11·37	1335·4	−55·4

† Artificial complexes were generated by docking the combining site of the fv molecule on the right to the one on the left. The five-parameter space was explored by simulated annealing in a range restricted to $\pm 10°$ for the four orientational parameters θ_1, φ_1, θ_2 and φ_2. The pseudo-stabilizing energy, ΔE_B, is the score function of the docking algorithm, calculated as:

$$\Delta E_B = \text{score function} = E - 0{\cdot}50\,B,$$

where E is a repulsive component proportional to steric overlap, and B is the surface area buried within the complex. All configurations studied were clustered according to ΔE_B values and relative distance in parameter space. The configurations in the table are the best score representatives of the top rank clusters, and are more than 8·7° apart in the parameter space.

‡ E225 and d1.3 were docked in the angular orientation found in the X-ray crystal.

§ The local minimum closest (less than 5°) to the X-ray crystal angular orientation.

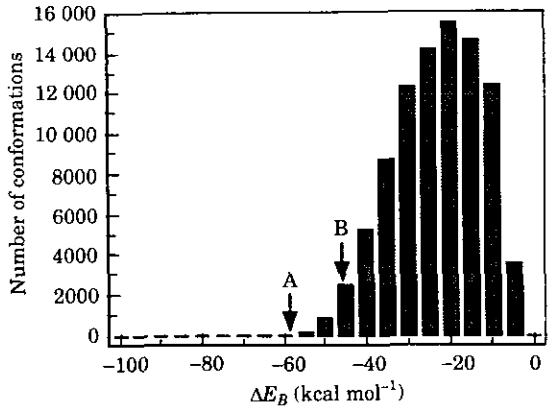

FIG. 3. The histogram of all the individual pseudo-energies of the D1.3–E225 complex encountered in the course of the simulated docking procedure as presented in Table 1. The arrows indicate (A) the absolute minimum and (B) the local minimum close to X-ray crystal configuration.

the tendency to find larger interface areas when the molecules are tested against themselves. In a few cases the pairing was symmetric, but overall this does not seem to be a sufficient explanation for this definite tendency (data not shown).

Discussion

The overall conclusion of a few years of protein docking seems to be that although shape complementarity might be a necessary requirement for productive protein–protein binding, it is not a sufficient condition. More precisely, shape complementarity is far from being a limiting criterion in the discrimination between real and artificial complexes. Attempts to reproduce known complexes by docking have led to repeated findings of false positive complex configurations showing extensive shape complementarity comparable to (and often greater than) the one found in the actual X-ray crystal configuration. This observation is very consistent and independent of the algorithm used to dock the proteins and to evaluate the interface areas and complementarity of surface patches: Wodak–Janin (Wodak & Janin, 1980; Cherfils *et al.*, 1989), Shoichet & Kuntz (1991), Jiang & Kim (1991), Connolly (1992) and Walls & Sternberg (1992). The results reported in the present paper indicate that this is not a specific feature found only when the docking counterparts are molecules known to actually bind. Indeed, extensive shape complementarity, as evaluated by the score function used, was found between different surface patches of any pair of molecules we tested, and although the docking was restricted to a very small region of the molecules, still several of these "false positive" configurations were found. In other words, our results show that the frequency distribution of complementary patches is basically the same irrespective of whether the docking counterparts are known to bind or not; this suggests that the probability of finding any given three-dimensional profile in the surface of any protein might be surprisingly high.

There are basically two contrasting interpretations of the general failure to accurately predict protein–protein complexes by docking algorithms. The first interpretation invokes a lack of accuracy of current methods of prediction. The high redundancy in simulated antibody fits could actually be due to flaws in the Wodak–Janin docking algorithm itself. Indeed we cannot rule out rigorously the possibility that the protein pairs that scored *in machina* as positive would not have significant surface complementarity in the real world. We are convinced that similar results would have been found with any other docking algorithm, since the one we used (which should be noted underestimates surface areas) has been shown previously to perform just as well as any other. However, more generally, it is true that since the simulation of protein–protein recognition in total physico-chemical detail is not computationally

TABLE 2

The matrix of pseudo-energies (ΔE_B, kcal mol^{-1}) for all pairwise combinations of a set of five immunoglobulins (D1.3, E225, J539, HyHEL 10, McPC603) and one protein (HEL)

	D1.3	E225	J539	HyHEL 10	McPC603	*HEL*
D1.3	−72·8	**−59·3**	−62·9	−56·9	−46·7	**−56·8**
E225	**−59·3**	−66·5	−54·6	−69·8	−58·1	−61·3
J539	−62·9	−54·6	−59·2	−58·1	−41	−53·5
HyHEL 10	−56·9	−69·8	−58·1	−70·6	−57·3	**−66·8**
McPC603	−46·7	−58·1	−41	−57·3	−59·4	−53·5
HEL	**−56·8**	−61·3	−53·5	**−66·8**	−53·5	−62·3

The figures in bold type correspond to molecules which actually form high-affinity complexes. The values shown are the highest single values obtained by the simulated docking procedure presented in Table 1.

feasible at the present time, very strong simplifications must be made. Such simplifications may result in a fatal loss of precision, or even worse, be far away from the reality. According to this view, our incapacity to understand protein recognition should be overcome as computational power is increased, or as more realistic simplifications are made that correctly grasp the essential features of protein–protein interactions. Both these attitudes can be found in the literature, in response to the accumulating reports that the use of criteria based on shape complementarity leads to redundancy. Some authors have developed different algorithms for analysing shape-complementarity [an extreme case being the systemic search performed on the Fab–Lysozyme system by Walls & Sternberg (1992) with the aid of a connexionist machine]; others have started to adopt new criteria in simulated docking [a good example being the recently reported work of Pellegrini & Doniak (1993)].

There is one version of this first interpretation that is particularly relevant for our examination of the shape-space concept. It could be argued that although shape complementarity *per se* cannot fully account for the specificity of protein recognition, it might nevertheless play a role in the process of complex formation, representing an important prerequisite that must be complemented by additional features, typically ion–ion interactions, hydrogen bonds or van der Waals forces. Using the shape-space metaphor this is equivalent to saying that geometrical definitions might represent only a reduced number of dimensions, and that additional parameters of the "generalized shape" lead to the "real" higher-dimensional space. This line of argument has been followed by the workers in the field of protein docking when trying to reproduce the X-ray structure of known protein complexes. Docking based on shape complementarity is interpreted as a very efficient first step in the "purification" of the actual configuration amongst the millions to billions of potential ones. Thus, in all studies reported so far, the shape complementarity criteria allowed the preselection of a much reduced set of "false positives", that always include the actual configuration, or more precisely configurations very close to it. Although the number of complementary "false positives" is great, representing typically 0·01–0·1% out of hundreds of millions trials, such indeterminacy would not represent an insuperable problem if the inclusion of additional criteria based on more detailed physical-chemical interactions and realistic energetic potentials were to allow the accurate selection of actual complexes amongst the false positive complexes preselected by docking.

It is therefore important to emphasize that further "purification" of the complementary configurations preselected by docking, by refining the energetic potentials and using programs like X-PLOR that take into account detailed interatomic interactions and explore specific side-chain conformations, has revealed that many of these "false complexes" would actually be expected to be quite "stable" (Cherfils *et al.*, 1991). In our own study, we performed a simple analysis of the relative three-dimensional distributions of charged ionic groups in the "false complexes" we obtained. The results gave no hint that this refinement would solve the problem; in some cases it actually made it worse, by improving the free energy of "false complexes" compared to that of the known complex (results not shown). Another possible refinement would be to take into account the fact that molecules are not rigid structures, but that the chains are flexible so that "induced fit" is a real possibility. However, it seems clear that this refinement could only increase the degree of redundancy, so that far from solving the problem of accounting for specificity on the basis of complementary shapes, it would actually make it worse.

It is of course methodologically impossible to rule out categorically the possibility that future refinements of docking algorithms might eventually lead to accurate predictions of protein–protein specificity. In this respect, we would like to make a constructive suggestion. In future work in this area, it would be useful to compare systematically actual X-ray crystal structures with the structures of false positives from simulated docking. Hopefully, some regular differences will become apparent; but as long as this is not the case, we can only conclude that we do not adequately understand what actually accounts for the specificity of binding.

However, this enterprise is manifestly of great difficulty, and the failures to date to exclude false positives are in themselves sufficiently striking that it is important to envisage a radically alternative interpretation. According to this second interpretation, there is something intrinsically flawed in trying to match "absolute" properties of the involved proteins. After all, the existing docking algorithms do seem to be working, and quite robustly, because they correctly include the actual configuration in a subset of only 0·01–0·1% of all configurations. Thus, it could be that these algorithms are already correctly grasping an essential feature of the reality of protein–protein interactions: some degree of shape complementarity may indeed be usually necessary for binding, but this is not even a partial explanation of specificity because the required degree of complementarity will always

occur between any two proteins whatsoever. If this is the case, then "more of the same"—i.e. the type of further refinement of docking algorithms mentioned above—will fail to solve the problem.

The alternative view, then, is that the specificity of protein recognition is inherently relational. Proteins in aqueous solution are correctly seen as very dynamical structures, and complex formation occurs when molecules co-bind, as thermodynamical stability emerges from the network of interactions between each and every residue (and indeed, in the search for a thoroughly reductionist explanation, between each and every molecular orbital) with its neighbours in both molecules and in the solvent (the extent of the relevant neighbourhood is certainly not well defined). From this point of view, the formation of a particular protein complex is stringently dependent on the unique protein–solvent environment that the involved proteins co-produce. A classical illustration of this concept is the case of "induced fit", where major conformational differences are seen between the structures of free molecules and those of (co)bound molecules that form very precise interatomic interactions.† Of course, as already noted, conformational plasticity will increase the frequency of productive interactions; we are therefore suggesting that the background frequency of precise intermolecular interactions is *a priori* extremely low, and that conformational rearrangement (most frequently "trivial" side-chain rearrangements and in extreme cases "induced fit") is always necessary to raise this frequency to that observed in specific protein–protein recognition. The notion that conformational changes play a role in antibody recognition is certainly not new and was reviewed in Tainer *et al.* (1985) and Colman (1988). The most significant evidence is provided by the correlation of segmental mobility with the location of antigenic determinants in several proteins (Westhof *et al.*, 1984; Getzoff *et al.*, 1987; Geysen *et al.*, 1987) and the actual reports on conformational changes of antibody and antigen molecules upon binding (Colman *et al.*, 1987; Sheriff *et al.*, 1987; Bentley *et al.*, 1990). The fundamental role of the solvent is certainly not limited to hydrophobic effects, since higher precision refinements of X-ray structures of different protein complexes are revealing hydrogen-bonded water molecules in the interface between the bound proteins, making very precise bridges between groups, and that are certainly involved in an enthalpic stabilization of the complex. Such water molecules, which make a sort of "sandwich" complex, illustrate very well the relational nature of binding, as they connect very specific groups (hydrogen donors or acceptors) in both molecules, which rearrange into very precise orientations. This is the extreme picture of protein–protein recognition specificity that we would like to draw here, where very dynamical molecular structures interact with each other in very unique ways that depend strictly on the environment that each ensemble creates. The uniqueness in (co)binding would emerge from multiple interactions at the quantic level where phenomena are "all or none" rather than continuous, and require precise "overlappings" of molecular orbitals (MO). At this level, the relational nature of intermolecular interactions is very clear: electron density redistributes within the new context, and MO coalesce, losing individual existence, as novel MO emerge from the former.

Our relational view of protein binding renders *a priori* difficult the prediction of the formation of protein complexes. The minimal realistic model which could potentially illustrate the relational nature of the phenomenon would be the inclusion of side-chains and water molecules in a dynamical way. The huge number of degrees of freedom in such a system would render its simulation computationally unfeasible. If our view, while not being new, is certainly discouraging for chemical-physicists, the relational (co)binding image can be nevertheless a useful heuristic tool for theoreticians and immunologists in general, who deal indirectly with protein–protein interactions. Most of all, it may contribute to irradicate the "original sin" represented by the "lock and key" image. It should be noted that whenever immunologists meet to discuss specificity and related subjects, they inevitably end up illustrating their points by drawing more or less complex "locks" and "keys" (this is usually after paying lip service to the idea that what is meant is "not only shape" but all sorts of binding properties of the molecules). However, if the relational picture we are proposing has any validity, then the lock and key drawings, which are supposed to be heuristic, will be not only wrong but positively misleading.

IMPLICATIONS FOR THE SHAPE-SPACE CONCEPT

The shape-space concept has given sterling service in the field of theoretical immunology, and we do not wish to suggest that it should be simply jettisoned. Even less would we wish to dismiss the broad qualitative idea, to which the shape-space concept has greatly contributed, that the immune repertoire possesses a definite internal structure based on a

† It should be noted in this context that attempts to reproduce known complexes by docking the structures of the free molecules are frequently doomed to failure, even when considerable plasticity is given to the structure (soft docking).

network of idiotypic interactions which develop through ontogeny. We do however wish to sound a cautionary note: if we are not to be led astray by mere wishful thinking, some hard rethinking of the shape-space concept is in order.

A major conclusion that we draw from the work presented in this paper is that a minimally realistic shape space, and an inversion space with a moderately low dimensionality of the vectors y and a reasonably simple, well-behaved function g, have virtually no chance of coinciding. Thus, in future work, it is not really acceptable to pass blithely back and forth between a realistic shape space and an inversion space, taking advantage of the best of both worlds while conveniently neglecting the particular difficulties and limitations inherent in each. We feel that when using eqn (1) [or (2)] in any particular instance, a definite commitment should be made to one or other interpretation, which implies taking on board the particular difficulties and limitations involved.

If the "realistic shape space" interpretation is chosen, then a major consequence is that the N dimensions of the vectors x should describe objective, intrinsic properties that characterize individual molecules independently of whether the molecule is free or in this or that complex. We have argued that the tertiary structure, and even more importantly the interatomic forces that are actually the immediate causes of binding specificity, are *not* intrinsic properties of this sort. At the present time, the only intrinsic characteristics of the required sort in which one can have any confidence are the amino acid sequences of the proteins involved.

However, in order to characterize individual immunoglobulins, it would be necessary to specify at least the amino acid sequences of the CDR regions that vary significantly from molecule to molecule, so that the dimensionality of the corresponding vector space would be high (of the order of 50). Moreover, although we do not doubt that under precisely given physiological conditions amino acid sequences do ultimately determine binding, the path is long and complex indeed, so that the possibility of actually predicting binding from sequences is quite remote.† It is, however, well known that single amino acid substitutions can lead to radical differences in binding specificity (Stark & Katon, 1991). Thus, in the terms of eqn (1), we can be virtually certain that the function f in any realistic shape space would be highly discontinuous.

In the context of theoretical immunology, these considerations render the use of a realistic shape space a singularly unattractive proposition. We shall therefore turn to the alternative interpretation, that of an "inversion space". In our view, the bulk of hitherto published work invoking the "shape space" concept is indeed best interpreted as a preliminary move in this direction.‡ In this case, however, a cautionary note is in order. Given that it is not grounded in "bottom up" realism, an inversion space (and the corresponding function g) runs a serious risk of being quite arbitrary if it is not properly derived by controlled reverse engineering from actual data on one or more empirical affinity matrices. The empirical data on idiotypic affinity matrices is frustratingly scanty, but it is not non-existent. In particular, it is known that many immunoglobulins are multi-specific, and capable of binding to epitopes that are chemically quite unrelated to each other (Inman, 1978). Some preliminary attempts have been made to take into account multi-specificity of this sort by postulating that immunoglobulin molecules possess multiple epitopes (Perelson & Oster, 1979; Stewart & Varela, 1991) or that "shape-space"§ has multiple centres of symmetry (Detours, personal communication). However, the general problem of the systematic construction of an inversion space has not yet received much serious attention. For example, it is not clear whether an individual molecule can be adequately represented in an inversion space by a discrete number of points, or whether it should be represented by one or several curves or volumes (Perelson and Oster referred to such questions very briefly). More generally, it is to be noted that the very possibility of successfully‖ constructing an inversion space is anything but a foregone conclusion. If the empirical affinity matrix is practically random, it will be impossible; and if, as we have argued, the specificity of antibody–antigen binding and idiotypic interactions is a relational property which is stringently dependent on each particular pair of

† The difficulties involved in predicting the binding between a pair of proteins are very likely of the same nature as those involved in predicting the precise three-dimensional atomic structure of a single protein in aqueous solution: in both cases, the problem stems from the highly *relational* nature of the interactions involved.

‡ A hint of this is given in the original paper by Oster & Perelson (1979), and also by Stewart & Varela (1991), but without further development.

§ In the present context, we interpret this as a potential inversion space.

‖ By "successful", we mean that an acceptable approximation to an empirical affinity matrix of dimension n can be obtained from an inversion space where the dimensionality M of the vectors y is fairly small compared to n, and the function g is reasonably simple and well behaved.

proteins involved, it is likely to prove arduous. It may turn out that the successful construction of an inversion space is only possible in cases where the empirical affinity matrix possesses a marked internal structure by virtue of a process of coselection during ontogeny of the lymphocyte clones involved. This would be a most interesting finding in its own right, but it should be recalled that in this case the resulting inversion space could not be used to generate affinity matrices of relatively unselected immunoglobulins.

What then can be concluded from this? It seems that at the present time we have no way, either through the analysis of empirical data or by theoretical reasoning, to be sure of adequately characterizing the internal structure of idiotypic affinity matrices. The generation of simulated matrices is thus a hazardous process. The only secure conclusion that can be drawn for theoretical immunology is that models of the immune system should aim at being generic and robust with respect to the structure of these matrices. In practice, this means that models should control their dependence on this unknown factor by employing a number of matrices generated by widely different mechanisms, including the quasi-random attribution of affinities between pairs of immunoglobulins.

We are most grateful to J. Cherfils and J. Janin for their generosity in making the package DOCK available to us, and for stimulating discussions. We thank the group of R. Poljak at the Pasteur Institute for their support, especially P. Alzari who strongly encouraged us to start and pursue this work and G. Bentley for the use of the co-ordinates of D1.3–E225 complex. We also wish to thank A. Coutinho's group for providing us with the environment where this work could be carried out, and C. Bhimarao for discussions which were most useful in developing the concept of an "inversion space". J. Carneiro thanks the financial support of Junta Nacional de Investigação Científica e Tecnológica-Programa Ciência, Lisbon (grant BD/2319/92-ID).

REFERENCES

BENTLEY, G. A., BOULOT, G., RIOTTOT, M. M. & POLJAK, R. J. (1990). Three-dimensional structure of an idiotope-anti-idiotope complex. *Nature, Lond.* **348,** 254–257.

CHERFILS, J., DUQUERROY, S. & JANIN, J. (1991). Protein-protein recognition analysed by docking simulation. *Proteins* **11,** 271–280.

CHERFILS, J. & JANIN, J. (1993). Protein docking algorithms: simulating molecular recognition. *Curr. Op. struct. Biol.* **3,** 265–269.

CHLOTHIA, C. (1974). Hydrophobic binding and accessible surface area in proteins. *Nature, Lond.* **248,** 338–339.

CHLOTHIA, C. & LESK, A. M. (1987). Canonical structures for the hypervariable regions of immunoglobulins. *J. molec. Biol.* **196,** 901–917.

COLMAN, P. M. (1988). Structure of antibody-antigen complexes: implications for immune recognition. *Adv. Immun.* **43,** 99–132.

COLMAN, P. M., LAVER, W. G., VARGHESE, J. N., BAKER, A. T., TULLOCH, P. A., AIR, G. M. & WEBSTER, R. G. (1987). Three-dimensional structure of a complex of antibody with influenza virus. *Nature, Lond.* **326,** 358–363.

CONNOLLY, M. L. (1986). Shape complementarity at the hemoglobin $\alpha 1\beta 1$ subunit interface. *Biopolymers* **25,** 1229–1247.

CONNOLLY, M. L. (1992). Shape distributions of protein topography. *Biopolymers* **32,** 1215–1236.

DAVIES, D. R. & PADLAN, E. A. (1990). Antibody-antigen complexes. *A. Rev. Biochem.* **59,** 439–473.

DE BOER, R., HOGEWEG, P. & PERELSON, A. S. (1992). Growth and recruitment in the immune network. In: *Theoretical and Experimental Insights into Immulogy* (Perelson, A. S. & Weisbuch, G., eds) pp. 223–247. Berlin: Springer-Verlag.

EHRLICH, P. (1901). *The Collected Papers of Paul Ehrlich*, Vol. 2. New York: Pergamon Press.

GEYSEN, H. M., TAINER, J. A., RODDA, S. J., MASON, T. J., ALEXANDER, H., GETZOFF, E. D. & LERNER, R. A. (1987). Chemistry of antibody binding to a protein. *Science* **235,** 1184–1190.

GETZOFF, E. D., GEYSEN, H. M., RODDA, S. J., ALEXANDER, H., TAINER, J. A. & LERNER, R. A. (1987). Mechanisms of antibody binding to a protein. *Science* **235,** 1191–1190.

INMAN, J. K. (1978). The antibody combining region: speculations on the hypothesis of general multispecificity. In: *Theroetical Immunology*, Vol. 8 (Bell, G. I., Perelson, A. S. & Pembley, G. H., eds) pp. 243–277. New York: Addison-Wesley.

JIANG, F. & KIM, S.-H. (1991). "Soft-docking": matching of molecular surface cubes. *J. molec. Biol.* **219,** 79–102.

KABSCH, W. (1976). A solution for the best rotation to relate two sets of vectors. *Acta Crystallog. sect. A* **32,** 922–923.

KAZATCHKINE, M. & COUTINHO, A. (1993). Clonal taxonomy versus network physiology. Are lymphocytes concerned with our definition of idiotypes? *Immun. Today* **14,** 513–515.

LEVITT, M. (1976). A simplified representation of protein conformations for rapid simulation of protein folding. *J. molec. Biol.* **104,** 59–107.

METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H. & TELLER, A. (1953). Equation of state calculation by fast computing machines. *J. chem. Phys.* **21,** 1087–1092.

NOVOTNY, J., BRUCCOLERI, R. E. & SAUL, F. A. (1989). On the attribution of binding energy in antigen-antibody complexes McPC 603, D1.3, and HyHEL-5. *Biochemistry* **28,** 4735–4749.

PELLEGRINI, M. & DONIACH, S. (1993). Computer simulation of antibody binding specificity. *Proteins* **15,** 436–444.

PERELSON, A. S. (1989). Immune network theory. *Immun. Rev.* **110,** 5–35.

PERELSON, A. S. & OSTER, G. F. (1979). Theoretical studies of clonal selection: minimal antibody repertoire size and reliability of self-non-self discrimination. *J. theor. Biol.* **81,** 645–670.

SEGEL, L. A. & PERELSON, A. S. (1988). Computations in shape-space: a new approach to immune network theory. In: *Theoretical Immunology*, Part 2 (Perelson, A. S., ed.) pp. 321–343. RedWood City, CA: Addison-Wesley.

SHARP, K. A., NICHOLLS, A., FINE, R. F. & HONIG, B. (1991). Reconciling the magnitude of the microscopic and macroscopic hydrophobic effects. *Science* **252,** 106–109.

SHERIFF, S., SILVERTON, E. W., PADLAN, E. A., COHEN, G. H., SMITH-GILL, S. J., FINZELL, B. C. & DAVIES, D. R. (1987). Three-dimensional structure of an antibody-antigen complex. *Proc. natn. Acad. Sci. U.S.A.* **84,** 8075–8079.

SHOICHET, B. K. & KUNTZ, I. D. (1991). Protein docking and complementarity. *J. molec. Biol.* **221,** 327–346.

STARK, S. E. & KATON, A. J. (1991). Antibodies that are specific for a single amino acid interchange in a protein epitope use structurally distinct variable region. *J. expl Med.* **174,** 613–624.

STEWART, J. & VARELA, F. (1989). Exploring the meaning of connectivity in the immune network. *Immun. Rev.* **110,** 37–71.

STEWART, J. & VARELA, F. (1991). Morphogenesys in shape space. *J. theor. Biol.* **153,** 477–498.

TAINER, J. A., GETZOFF, E. D., PATERSON, Y., OLSON, A. J. & LERNER, R. A. (1985). The atomic mobility component of protein antigenicity. *A. Rev. Immun.* **3,** 501–535.

WALLS, P. H. & STERNBERG, M. J. E. (1992). A new algorithm to

model protein-protein recognition based on surface complementarity: application to antibody-antigen docking. *J. molec. Biol.* **228,** 277–297.

WESTHOF, E., ALTSCHUH, D., MORAS, D., BLOOMER, A. C., MONDRAGON, A., KLUG, A. & VAN REGENMORTEL, V. (1984). Correlation between segmental mobilty and the location of antigenic determinants in proteins. *Nature, Lond.* **311,** 123–126.

WODAK, S. J. & JANIN, J. (1980). Analytical approximation to the accessible surface area of proteins. *Proc. natn. Acad. Sci. U.S.A.* **77,** 1736–1740.