PROTEINS: Structure, Function, and Bioinformatics 66:000–000 (2007)

Interaction of Malaria Parasite-Inhibitory Antibodies With the Merozoite Surface Protein $MSP1_{19}$ by Computational Docking

Flavia Autore,[1] Sara Melchiorre,[1] Jens Kleinjung,[3] William D. Morgan,[2] and Franca Fraternali[1,4]*

[1]*Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, Complesso Universitario Monte Sant'Angelo, via Cinthia, 80126, Naples, Italy*

[2]*Division of Parasitology, National Institute for Medical Research, Mill Hill, London, NW7 1AA, United Kingdom*

[3]*Division of Mathematical Biology, National Institute for Medical Research, Mill Hill, London, NW7 1AA, United Kingdom*

[4]*Bioinformatics Unit, Randall Division of Cell and Molecular Biophysics, New Hunt's House, London, SE1 1UL, United Kingdom*

ABSTRACT **Merozoite surface protein 1 (MSP1) of the malaria parasite *Plasmodium falciparum* is an important vaccine candidate antigen. Antibodies specific for the C-terminal maturation product, $MSP1_{19}$, have been shown to inhibit erythrocyte invasion and parasite growth. Specific monoclonal antibodies react with conformational epitopes contained within the two EGF-like domains that constitute the antigen $MSP1_{19}$. To gain greater insight into the inhibitory process, the authors selected two strongly inhibitory antibodies (designated 12.8 and 12.10) and modeled their structures by homology. Computational docking was used to generate antigen–antibody complexes and a selection filter based on NMR data was applied to obtain plausible models. Molecular Dynamics simulations of the selected complexes were performed to evaluate the role of specific side chains in the binding. Favorable complexes were obtained that complement the NMR data in defining specific binding sites. These models can provide valuable guidelines for future experimental work that is devoted to the understanding of the action mechanism of invasion-inhibitory antibodies. Proteins 2007;66:000–000.** © 2006 Wiley-Liss, Inc.

Key words: antibody–antigen complex; docking; molecular dynamics; NMR; malaria; *Plasmodium falciparum*; MSP1 (merozoite surface protein 1)

INTRODUCTION

Malaria is one of the most widely spread diseases, with more than 40% of the world population at risk of being infected. The development of effective malaria vaccines is therefore a mandatory public health challenge.[1] Acute clinical malaria from infection with *Plasmodium falciparum* is associated with replication of the asexual blood-stage parasite (merozoite) in circulating erythrocytes. After invasion, the merozoite multiplies within the host cell causing its rupture and to liberate several newly formed merozoites. These, in turn, invade other erythrocytes thus dramatically increasing the level of parasitemia that causes the typical symptoms of malaria. The merozoite expresses a number of surface proteins, one or more of which are thought to mediate the initial interaction between the parasite and host erythrocyte.[2,3] Merozoite surface protein 1 (MSP1) is one of the most extensively studied molecules of *P. falciparum*[4] and is present on the surface of the released merozoite in the form of a multicomponent protein complex derived via proteolytic processing.[5,6] In between the merozoite release and the completion of erythrocyte invasion, the membrane-bound component ($MSP1_{42}$) of this surface complex is further cleaved at a single site to form two fragments, $MSP1_{33}$ and $MSP1_{19}$. As a consequence, the majority of the complex is shed from the parasite surface, leaving only $MSP1_{19}$ exposed. Therefore, this protein is considered a very promising malaria vaccine candidate.[7] Several monoclonal antibodies (mAbs) directed against $MSP1_{19}$ have been identified, and some inhibit invasion of erythrocytes in vitro,[8] while some inhibit both secondary processing of MSP1 and erythrocyte invasion.[9] Those antibodies that are specific for $MSP1_{19}$ but are not inhibitory can be classified as blocking antibodies (that interfere with the binding of inhibitory antibodies) and neutral antibodies (which bind to $MSP1_{19}$ but neither inhibit invasion nor block the binding to inhibitory antibodies).[10,11]

Abbreviations: CDR, complementarity determining region; El, electrostatic potential; LJ, Lennard–Jones potential; mAb, monoclonal antibody; MD, molecular dynamics; VH, variable heavy chain; VL, variable light chain; SR, short range.

The Supplementary Material referred to in this article can be found at http://www.interscience.wiley.com/jpages/0887-3585/suppmat/

Grant sponsors: Medical Research Council, European Commission through the EUROMALVAC consortium and CEINGE Advanced Biotechnology consortium

*Correspondence to: Franca Fraternali, Dipartimento di Chimica Organica e Biochimica, Università di Napoli Federico II, Complesso Universitario Monte Sant'Angelo, via Cinthia, 80126, Naples, Italy. E-mail: franca.fraternali@kcl.ac.uk

Received 10 April 2006; Revised 4 August 2006; Accepted 9 August 2006

Published online 00 Month 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21212

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 2

The structure of *P. falciparum* $MSP1_{19}$ has been solved by NMR,[12] and its homologues from *Plasmodium cynomolgi*[13] and *Plasmodium knowlesi*[14] have been solved by X-ray crystallography. The molecule is composed of two EGF-like domains forming a disk-shaped structure. The binding mode to specific antibodies was subject of different structural studies. Recently, the crystal structure of $MSP1_{19}$ in complex with the Fab-fragment of mAb G17.12 has been solved.[15] This antibody is certainly not inhibitory, but it has not yet been determined whether it is a neutral or a blocking antibody. A recent study has been devoted to the discrimination between blocking, inhibitory, and neutral antibodies for $MSP1_{19}$ mapping *via* mutagenesis the epitopes for inhibitory mAbs 12.8 and 12.10[16] and for blocking mAbs such as 1E1 and 7.5.[17]

NMR epitope mapping has been performed with several $MSP1_{19}$-specific antibodies using chemical shift perturbation as well as the more precise cross-saturation method.[18,19] From this experiment, six residues have been individuated for antibody 12.10 (Gln 6, Cys 7, Asn 15-Cys 18) and five residues for antibody 12.8 (Gln 14-Cys 18). This technique is convenient for mapping protein–protein interactions, since there is one observable probe for each residue in the antigen. In practice, a fraction of the backbone amide signals may not be observed because of exchange line broadening or signal overlap. One limitation of this method is that it provides no information on contacts to the side chains of antigen residues. To investigate this further, we decided to model the antibody structures, and to investigate protein–protein interactions with the antigen (including potential side chain contacts) by means of docking calculations. Prior knowledge of the epitope on the antigen, as obtained from NMR cross-saturation, provides reliable restraints to guide these calculations.[20]

The rigid body docking algorithm ZDOCK proved to be very effective in predicting protein–protein complexes in previous studies.[21] The effectiveness of the method relies on a scoring function that properly takes into account the major determinants of protein–protein association, such as shape, electrostatic complementarities and desolvation energies.[22,23] The method profits from a very efficient search engine, based on a Fast Fourier transformation algorithm,[21] and was found to be one of the best protein–protein interaction predictors for antibody–antigen prediction.[24] Precise mapping of the epitopes that are recognized by the mAbs 12.8 and 12.10 has important implications for rational vaccine design and development of antibody-based therapeutics. Recently, the importance of local protein flexibility on protein–protein association has been shown,[25] and therefore we decided to refine and optimize the resulting complexes from the docking procedure by performing Molecular Dynamic simulations in explicit solvent. The use of MD simulations for the unbound state of protein–protein complexes has shown that core interface residues are less mobile than the rest of the surface and therefore rigid docking procedures do not account for different mobilities.[26] This implies that while rigid docking is acceptable for core residues, it is not ideal for the residues on the surface of the protein. We observed that the docked complexes are not optimally positioned. The two proteins tend to be in general more distant in the model than in the reference complex, probably because of ZDOCK's weak electrostatic term (see also Discussion). Therefore, we opted for an MD refinement that should correct for the relative positioning through a more accurate force field and added residue mobility at the interface.

The methodology used here may prove to be a valuable general approach to the problem of antibody–antigen interactions, as in cases where experimental structure resolution is difficult.

METHODS

Sequence Alignments and Modeling of the Antibodies

Sequence alignments were generated by using the T-Coffee method[27] that uses the Lalign algorithm[28] from the FASTA package.[29] The best alignments were selected according to not only the alignment score but also the length of the hypervariable loop gaps in the complementarity determining region (CDR). When multiple choices were possible, we opted for the template structure with the highest resolution. Three-dimensional models were generated using the SWISS-MODEL web server by means of the project mode option that allows for submitting a manually optimized modeling request. This way it is possible to control a wide rage of parameters, for example, selection of template and gap placement in the alignment.[30] Since only the variable regions of the antibody were sequenced, the antibodies were modeled in the form of single chain variable fragments (scFv), with the light and heavy chains connected by a flexible linker segment.

NMR Filter

The cross-saturation technique involves excitation of protons in one macromolecule (in this case the antibody), followed by transfer of magnetization to protons within the second macromolecular component of the complex (the antigen in this case). Cross-saturation is detected by monitoring the backbone amide cross peaks in a HSQC or TROSY correlation spectrum of the antigen, which is perdeuterated and uniformly labeled with ^{15}N. The maximum range of saturation transfer has been estimated at 7 Å, indicating that for each antigen residue where a strong cross-saturation effect is observed, at least one antibody proton is located within this distance from the backbone amide proton of that residue. The nomenclature used for the models is the following: models 12.10_2FBJ and 12.10_1JHL were built using templates 2FBJ[31] and 1JHL,[32] respectively; model 12.8_1IND was built using template 1IND[33] and model 12.8_1GGI using template 1GGI.[34]

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 3

Complex Selection and Clustering

The mAbs were submitted to a docking procedure with antigen $MSP1_{19}$ using the program ZDOCK 2.3.[21,22,35] Only the framework of the antibody was excluded from the docking searches.

Complexes were selected on the basis of experimental NOE couplings (see NMR filter) and classified by cluster analysis as explained in the following. From NMR experiments lists of cross-saturation antigen HN atoms in possible contact with ANY of the antibody's H atoms are given.

For antibody 12.10 the HN of residues 6, 7, 15, 16, 17, 18 (List 1), while for antibody 12.8 the HN of residues 14, 15, 16, 17, 18 (List 2) have been identified by these NMR experiments. These few restraints are clearly not sufficient for an unambiguous complex selection; therefore, a number of analyses (distances from the hypervariable regions, solvent accessibilities, manual inspections) are necessary to be able to propose models.

An HN-H (antigen–antibody) distance below the cut off of 7 Å (see NMR filter) was counted as a restraint satisfaction. The total number of restraint satisfactions for each of the 2000 model structures was taken as the rank of the structure and a list sorted by structure rank was compiled. The top 17 models (selected structures) for 12.8 and the top 9 models for 12.10 of the sorted list were analyzed in detail by clustering. Clustering was performed in two dimensions: structural similarity as first dimension and constraint satisfaction similarity as second dimension. Structural similarity values were obtained from structural alignments of all pairs of selected structures, using a modified contact-based alignment program.[36] The root mean square deviation (RMSD) values of all $n(n-1)/2$ pairs form a triangular matrix (excluding diagonal elements). Constraint satisfaction similarity was calculated the same way by calculating the difference of restraint satisfactions between all pairs, yielding another triangular matrix. The two triangular matrices were combined in a nonsymmetrical square matrix, and two-dimensional clustering was performed using the "heatmap" function of the R environment.[37] This function performs hierarchical clustering in two dimensions based on the distance matrix values and generates a false color image of the cluster matrix.

The qualitatively same result was obtained when the triangular difference matrices were both scaled to the value range [0:1], but the grouping of favorable solutions was more scattered.

The docking of mAb G17.12 with antigen $MSP1_{19}$ was performed as a test case to check the reliability of the filtering through external restraints and the clustering applied. To test the validity of using a model for the antibody in the docking procedure, we have additionally constructed a model for the mAb G17.12 sequence from the template 1IAI with 83% sequence identity. We have also performed the docking on the modeled structure that we will call MG17.12 and have compared the results with the ones from the original complex (see Results).

For the analysis of the data relative to G17.12, an artificial dataset of restraints had to be created (in this case there are no NOE experiments but one can assume to use the restraints from the crystal structure in an analogous way). Specifically, the distance between 20 HN atoms in the antigen (residues 8–15, 21–28, and 37–41) and 58 hydrogen atoms of the antibody's heavy chain loops (11 on H1, 11 on H2, and 36 on H3) was calculated. In this case the top 50 models were considered for further analysis.

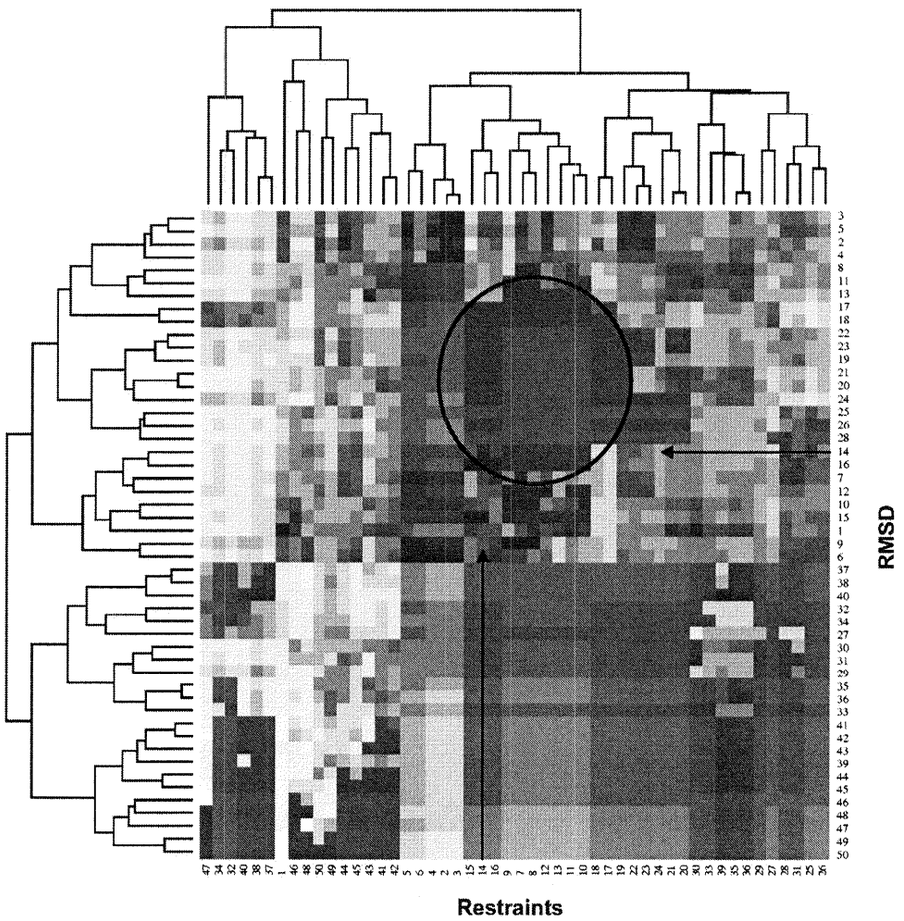
Clustering allowed us to divide the solution space into subgroups and to investigate whether the grouping is similar in terms of structural similarity and restraint satisfaction. We expect favorable subsolutions to be structurally similar and to have similar rank. Moreover, we expect favorable solutions to have a high number of restraint satisfactions, meaning a low rank. Therefore, in Figure 1 and Figure S4(A,B), Supplementary Materials, we were selecting for clusters of intense red cells (low distance) with low-ranking labels. F1

Molecular Dynamics Simulations

All simulations were performed with the GROMACS package[23] using the GROMOS96 force field.[38] The structures were solvated in a box of SCP water,[39] in a solution of 50 m*M* NaCl. All simulated boxes contained about 17,000 water molecules. Simulations were carried out at a constant temperature of 300 K. The Berendsen algorithm was applied for temperature and pressure coupling.[40] After a first steepest descent energy minimization with positional restraints on the solute, the LINCS algorithm was used to constrain the bonds[41] and to carry out an initial 200 ps simulation with the positions of the solute atoms restrained by a force constant of 3000 kJ/(mol nm^2) to let the water diffuse around the molecule and to equilibrate. The particle mesh Ewald method (PME)[42] was used for the calculation of electrostatic contribution to non bonded interactions (grid spacing of 0.12 nm). A reaction field method[43] was used for the calculation of electrostatic contributions in the simulations of the X-ray structure, and ions were added to neutralize the system.

During a number of simulation trials we observed the flexible loop of the antigen encompassing residues 65–75 to be rapidly attracted toward the antibody and to directly interact with the side chains of loops L1 and L2 of antibody 12.10 and with loops H1 and H2 of antibody 12.8. This attraction force would govern the antigen positioning at the binding surface. In some cases, as for antibody 12.10, it would even obstruct a proper docking to the CDRs. Considering that the NMR data rule out significant interactions with this loop, we decided to keep the residues 65–75 restrained to the NMR conformation during the simulation with a force constant of 3000 kJ/(mol nm^2).

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 4



COLOR

Fig. 1. Selection of docked structure by two-dimensional clustering. Number labels indicate the rank of the complex (see text). The root means square deviation dimension was obtained from a distance matrix of structural alignments of all pairs of the 50 top ranking structures, while the restraints dimension was obtained by calculating the difference matrix of restraint satisfactions between all pairs. Clustering was performed by using the "heatmap" function of the R environment,[37] which generates a false color image of the clustering matrix. The circle indicates the region of the subclusters with low pairwise distances (dark red).

RESULTS

Test of the Docking and Complex Selection Procedure

To test the adopted docking procedure, in particular the filtering of structures based on external restraints, we chose as a test case a known complex of the antigen $MSP1_{19}$ with the Fab-fragment of mAb G17.12, recently solved by X-ray crystallography.[15]

We separated antigen and antibody and generated 2000 docked structures, starting from randomly chosen relative orientations. Favorable complexes were preselected by a combination of restraint filtering and two-dimensional-clustering as follows: structures were ranked by the number of HN-H (antigen–antibody) distances below 7 Å for a given list of 20 HN atoms and 58 H atoms (see Methods for details). These restraints are chosen in analogy with the NMR data that are available for the unknown complexes of antigen $MSP1_{19}$ with antibodies 12.10 and 12.8. The 50 top-ranked complexes were clustered according to their pairwise structural differences (first cluster dimension) and their pairwise restraint–satisfaction differences (second cluster dimension). The cluster map in Figure 1 highlights a favorable subgroup within those top ranking complexes that show low pairwise differences in both dimensions.

One favorable representative of this selection was submitted to 5 ns of MD simulation. As a reference, we also performed 2 ns of MD on the native structure, to confirm that the structure would be stable under the adopted conditions (data not shown).

In Figure 2(A), a ribbon representation of the X-ray structure of complex mAbG17.12-$MSP1_{19}$ is shown. The binding region as calculated with POPS[45,46] is shown in dark grey. In Figure 2(B), a superimposition of the selected docked structure (black) and the X-ray structure (grey) is shown, while in Figure 2(C) the superim-

F2

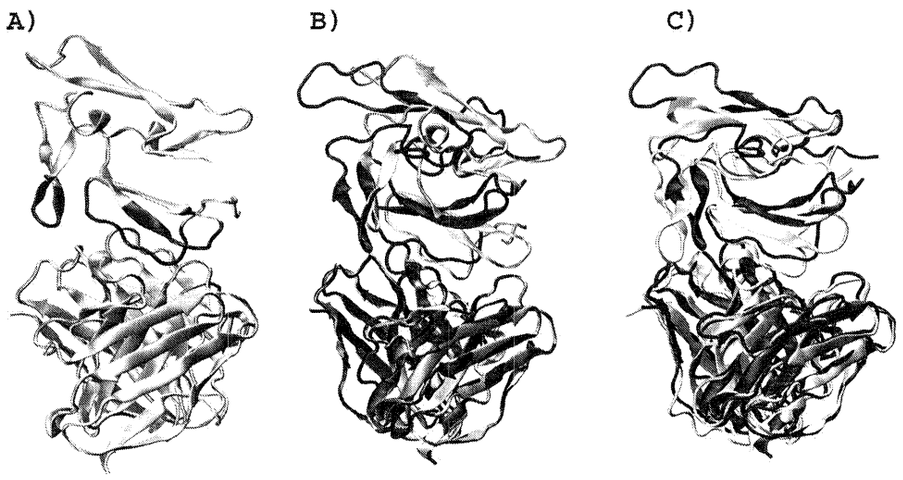


Fig. 2. (**A**) Ribbon representation of the X-ray structure. The residues of the binding-site buried upon complex formation as calculated by POPScomp[44] are displayed in black. (**B**) Superimposition of the X-ray structure (grey) and the docked complex (black). (**C**) Superimposition of the docked complex before (black) and after (grey) 5 ns of MD.

position of the docked structure (black) and the final conformer after 5 ns of MD (grey) is displayed.

The docked complex, although very similar to the X-ray structure, is initially not in its optimal configuration, as can be seen from the comparison of the distances between the centers of mass of antibody and antigen
F3 [Fig. 3(A)]. These are about 30 Å because of the size of the antibody.

The X-ray structure shows a value of 31 Å for this distance, while the docked complex shows an initial value of 33 Å. During the MD simulation, antigen and antibody move closer to each other and the distance between their centers of mass shortens by 2.0 Å, after about 3800 ps becomes very close to the X-ray value, and toward the end stabilizes around 30 Å.

T1 In Table I, the solvent accessible surface areas buried upon formation of the two complexes (X-ray and final conformer from MD) are reported. By comparison with the X-ray structure we found that 50% of the original contacts formed by the antibody and 60% of those in the antigen were preserved. As shown already in Figure 2(A), the antibody contains contacts in all surface regions that are buried in the X-ray complex, and it manages to draw also the loop of the second EGF domain [residues 65–73, in black in Fig. 2(A)] closer to the antibody with respect to the structure from ZDOCK. The time evolution of the number of hydrogen bonds that are formed between the antigen and the CDRs of the light (grey line) and heavy (black line) chains of the antibody is shown in Figure 3(B). During MD, we observe an increase of the number of hydrogen bonds at the interface, reaching a total number of eight at the end of the simulation. For the X-ray structure the total number of antibody–antigen hydrogen bonds was reported to be 11.[15]

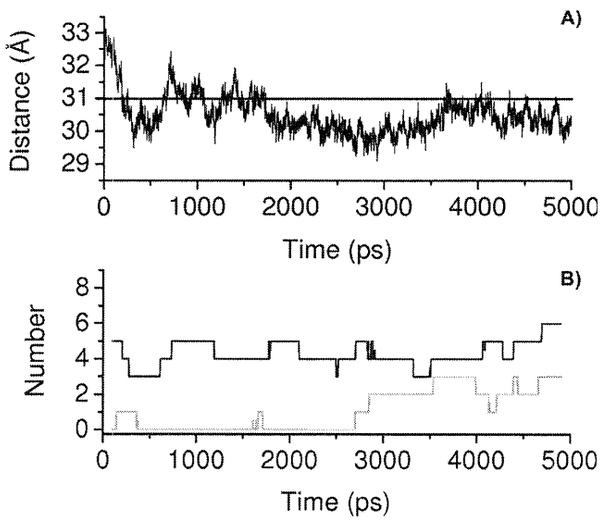


Fig. 3. (**A**) Time evolution of the distance between the centers of mass of antibody and antigen of the docked complex. The distance of centers of mass in the X-ray structure is represented by a continuous line at 31 Å. (**B**) Number of hydrogen bonds formed between the antigen and the CDRs of the antibody in the docked complex. The hydrogen bonds formed with the heavy chain and the light chain are displayed in black and in grey, respectively.

For the modeled structure of G17.12 onto the template FAB 730.1.4 (MG17.12), the docking performed with the modeled structure (MG17.12) gave rise to a best complex very similar to the one obtained with the original crystal structure. The superimposition between the complex from the modeled structure and the X-ray structure is reported in Figure S1, Supplementary Materials. The complexes from G17.12 and MG17.12 have a similar

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 6

TABLE I. Solvent Accessible Surface Areas Buried Upon Formation of the Complex Between Antigen $MSP1_{19}$ and Antibody G17.12

Complex ZDock	ΔSASA Å^2	X-ray	ΔSASA Å^2	CDR
(a) Comparison of the selected conformer from ZDOCK with the X-ray structure				
Antibody				
THR H3O	25.6			H1
THR H31	**43.9**	**THR H31**	**14.3**	**H1**
TRP H50	**13.4**	**TRP H50**	**38.3**	**—**
		ASN H52	19.7	H2
HISH 54	**56.1**	**HIS H54**	**38.4**	**H2**
		GLN H59	31.8	—
ASN H99	13.3			H3
TYR H100	**34.9**	**TYR H100**	**21.7**	**H3**
TYR H100a	**66.4**	**TYR H100a**	**70.5**	**H3**
ARG H100b	**51.1**	**ARG H100b**	**42.1**	**H3**
		SER L30	34.2	L1
		SER L31	22.3	L1
TYR L33	**40.9**	**TYR L33**	**52.5**	**—**
		HIS L93	23.9	L3
		HIS L94	19.1	L3
		TYR L97	17.3	L3
Antigen				
		VAL 8	47.1	
LYS 9	**108.5**	**LYS 9**	**69.0**	
LYS 10	**16.4**	**LYS 10**	**56.6**	
GLN 11	**144.0**	**GLN 11**	**168.0**	
CYS 12	12.0			
PRO 13	**52.7**	**PRO 13**	**21.8**	
GLN 14	**132.4**	**GLN 14**	**37.7**	
		ASP 23	32.9	
		GLU 24	45.5	
		ARG 25	27.3	
GLN 36	35.2			
ASP 39	**100.7**	**ASP 39**	**110.5**	
LYS 40	36.9			
(b) Comparison of the last conformer after 5 ns of MD simulations of the X-ray structure				
Antibody				
THR H28	23.8			H1
THR H30	49.7			**H1**
THR H31	**23.9**	**THR H31**	**14.3**	H1
		TRP H50	38.3	H2
		ASN H52	19.7	H2
HIS H54	**54.9**	**HIS H54**	**38.4**	**H2**
		GLN H59	31.8	—
TYR H100	**20.8**	**TYR H100**	**21.7**	**H3**
TYR H100a	**73.5**	**TYR H100a**	**70.5**	**H3**
ARG H100b	**23.1**	**ARG H100b**	**42.1**	**H3**
ASP H101	18.0			H3
		SER L30	34.2	L1
		SER L31	22.3	L1
SER L32	16.3			L1
TYR L33	**79.6**	**TYR L33**	**52.5**	**—**
PHE L92	13.4			L3
HIS L93	**11.5**	**HIS L93**	**23.9**	**L3**
		HIS L94	19.1	**L3**
TYR L97	**20.0**	**TYR L97**	**17.3**	**L3**

TABLE I. (Continued)

Complex ZDock	ΔSASA Å^2	X-ray	ΔSASA Å^2	CDR
Antigen				
		VAL 8	47.1	
LYS 9	**77.3**	**LYS 9**	**69.0**	
LYS 10	**76.8**	**LYS 10**	**56.6**	
GLN 11	**12.3**	**GLN 11**	**168.0**	
CYS 12	47.8			
PRO 13	**67.1**	**PRO 13**	**21.8**	
GLN 14	**101.5**	**GLN 14**	**37.7**	
		ASP 23	32.9	
GLU 24	**101.5**	**GLU 24**	**45.5**	
		ARG 25	27.3	
CYS 28	12.7			
ASP 39	**106.8**	**ASP 39**	**110.5**	
LYS 40	29.6			

In bold: matching proton restraints of the ZDOCK complex and the X-ray structure.

distance in rms from the X-ray complex: 2.8 and 3.2 Å, respectively. We decided therefore to use the original complex for further analysis.

On the basis of the above results, we considered the adopted complex modeling procedure to be reliable in detecting plausible complexes, and therefore we applied it to the modeling of the unknown complex structures of MSP_{19} and mAbs 12.8 and 12.10.

Modeling of the Antibodies

Inorder to model the structures of single chain variable fragments (scFv) of the two mAbs 12.8 and 12.10 (see Methods section), we selected the templates according to the guidelines of Morea et al.[47] The key role of loop H3 in the antigen recognition specificity restricts the choice of templates to the ones with the same length of this loop. In Table II, the PDB codes of these potential template structures are reported together with the H3 loop sequence and sequence identities. The structures that were finally selected as modeling templates are shown in bold. For mAb 12.10 we selected templates 1JHL and 2FBJ, because they share the most similar sequence in the central part of the H3 loop, which should be critical to the interaction with the antigen. The nomenclature used for the models is specified in the Methods section. Complete sequence alignments are reported in Figure 4. In Figure 5 the electrostatic surfaces of the modeled antibodies and of the antigen are displayed. The chosen view is from the top of the binding site of the CDR loops. The surfaces are quite different in the binding region: 12.8_1IND presents a more pronounced positive surface (blue) in the central region, while 12.8_1GGI presents a more negative surface (red) on the right side of this surface. For mAb 12.10, both modeled surfaces are on average more hydrophobic than 12.8, with 12.10_1JHL presenting a cavity with negative electrostatic potential in the central binding site region.

T2

F4, F5

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 7

TABLE II. Parameters of Potential Template Structures for the Modeling of Antibodies 12.8 and 12.10

PDB code	Resolution Å	Identity %	Length (H3)	Sequence (H3)
12.8	**—**	**—**	**10**	CVQSVIDNWG
1IND[33]	**2.2**	**41% (30%)**	**10**	CASHRFVHWG
1GGI[34]	**2.8**	**48% (50%)**	**10**	CVQEGYIYWG
12.10	**—**	**—**	**14**	CARNWAYWFEDVWG
1BBD[51]	2.8	66% (36%)	14	CDGYYSYYDMDYWG
1JHL[32]	**2.4**	**66% (43%)**	**14**	CTRDDNYGAMDYWG
2FBJ[31]	**1.9**	**52% (43%)**	**14**	CARLHYYGYNAYWG
7FAB[52]	2.0	45% (57%)	14	CARNLIAGGIDVWG
1MFB[53]	2.1	51% (36%)	14	CTRGGHGYYGDYWG
1JEL[54]	2.8	65% (57%)	14	CARVMGEQYFDVWG

The length and the sequences of H3 loops are reported for each potential template structure. Percentages of identity with 12.8 and 12.10 given relative to the entire sequence; identities in parentheses are given relative to the H3 loop only. In bold: structures that were selected as templates.

The surface of 12.10_2FBJ is on the other hand more hydrophobic. The negatively charged surface of the $MSP1_{19}$ region, as identified by epitope mapping, should prove to be more complementary for interactions with either more hydrophobic surfaces or surfaces with fewer negative spots in the electrostatic potential.

We have also modeled the sequence of G17.12 mAb with the template FAB 730.1.4, as a control test (see methods section) and the sequence alignments are reported in Figure S2, Supplementary Materials. The rms deviation on the C^{α} atoms between the model and the X-ray mAbs is 1.39 Å, and the CDRs loops are very similar in sequence; therefore, we expect a similar behavior in terms of binding/selectivity.

Docked Structures

Docking of 12.10 and $MSP1_{19}$

We performed ZDOCK calculations with the antigen $MSP1_{19}$ and the two modeled antibodies 12.10_2FBJ and 12.10_1JHL. For the latter, none of the 2000 docked complexes that were passing the NMR filter (see Methods section) presented the antigen in a suitable position for binding, that is, located more or less at the centre of the binding CDR loops. Therefore, no further analysis was performed on this complex. For 12.10_2FBJ a total of nine complexes were satisfying the NMR data (Fig. S3 and Table SI, Supplementary Materials).

The clustering procedure described in the Methods section of these nine complexes is shown Figure S4(A) and we selected complex 1 (low distance, intense red cells with low ranking labels).

For this structure, 5 ns of MD simulations were performed, and the final conformer was used for further analysis. In Figure 6(A,B) the selected ZDOCK complex and the final MD conformer are shown superimposed. As for the case of the X-ray complex, the MD simulation has the effect of pulling the antigen closer to the antibody. In Table IIIa the areas that were buried by formation of the complex as calculated by POPScomp[44] are reported. All of the residues indicated by the NMR cross-saturation data (Gln 6, Cys 7, Asn 15-Cys 18) are in contact with the heavy chain of the antibody, and as can be seen from Table IIIb, four of the six residues of the NMR data set are close to the H3 loop. Specifically, for the antigen residues with strong cross-saturation, the model shows interaction of the side chain of Arg H57 (flanking loop H2) with residues Gln 6 and Cys 7, of side chains Tyr H100b and Trp H100c (loop H3) with residues Asn 15 to Cys 18, and residues Asn 15 and Ser 16 also near hydrophilic side chains of loops L1 and L3. These results are summarized in Figure 6(C): the NMR-predicted interaction surface is colored in blue, while the residues identified by POPScomp are plotted in orange. In addition, the residues that did not show sufficient burial despite respecting the NMR filter are colored in yellow. In Table IVa, Lennard–Jones (LJ) and electrostatic contributions (El) are reported for the intramolecular energies of the antibody and the antigen (mAb-mAb and $MSP1_{19}$-$MSP1_{19}$ contributions respectively), together with the intermolecular antigen–antibody contributions. For the latter term, which reflects the relative strength of the interaction, the El and LJ contributions become increasingly more favorable during the MD simulation, suggesting that the system relaxes toward a more favorable structure. The energetic contributions of the loops contain most of the entire binding interaction of the antibody: therefore, while these become more favorable, the antigen manages to position itself better at the CDRs. The final binding mode shows one side of the discoidal $MSP1_{19}$ facing the CDR loops, with the antigen loop comprising residues 65–75 pointing toward the left of the central binding site, facing antibody loops L2 and L1 [Fig. 6(A)].

Docking of 12.8 and MSP119

As for antibody 12.10 we performed ZDOCK calculations with the antigen $MSP1_{19}$ and the two modeled antibodies 12.8_1IND and 12.8_1GGI. Both gave rise to a series of complexes respecting the NMR filter: 16 complexes for 12.8_1IND and 15 complexes for 12.8_1GGI (Fig. S5 and Table SII, Supplementary Materials).

Additional NMR information was obtained by performing 3D-TROSY-NOESY experiments (W. Morgan, unpublished results) for the complexes of the 12.8 and 12.10 mAb

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 8

A)

```
                    10          20          30          40          50          60
12.8_VH     QIHLVQSGPELRKPGETVKISCKAS[GYTFTNY]GMNWVKQDTGKGLKWMGWI[NTFSG]EPTY
1IND_VH     EVTLVESGGDSVKPGGSLKLSCAAS[GFTLSGE]TMSWVRQTPEKRLEWVATT[LSGGG]FTFY
            :: **:** :  *** ::*:** ** *:*::.   *.**:* . * *:*:.    : .*  . *
                    70          80          90          100         110
12.8_VH     ADDFKGRFAFSLETSATTAYLQINNLKNEDMATYF[CVQSVIDNWG]QGTTLTVSS
1IND_VH     SASVKGRFTISRDNAQNNLYLQLNSLRSEDTALYF[CASHRFVHWG]HGTLVTVS-
            : ..****::* :.: .. ***:*.*:.** * ** *..  : :** :** :***
                    10          20          30          40          50
12.8_VL     QIVLTQSPAIMSASPGEKVTMTCSA[SSSI----DYM]HWYQQKPGTSPKRWIY[DTSKL]ASG
1IND_VL     -AVVTQESA-LTTSPGETVTLTCRS[STGAVTTSNYA]NWVQEKPDHLFTGLIG[GTNNR]APG
              *:**..* :::****.**:** : *:.     :*  :* *:**.   .  *  .*.:  *.*
               60         70          80          90          100
12.8_VL     VPARFSGSGSGTSYSLTISSMEAEDAATYYCHQ[RSSYPFTF]GSGTKLEIKR-
1IND_VL     VPARFSGSLIGDKAALTITGAQTEDEARYFCAL[WYSNLWVF]GGGTKLTVLGQ
            ********  * . :***:. ::** * *:*     *  :.* *.**** :
```

B)

```
                    10          20          30          40          50
12.8_VH     QIHLVQSGPELRKPGETVKISCKAS[GYTFTNY]GM--NWVKQDTGKGLKWMGWI[NTFSG]EP
1GGI_VH     QVQLKESGPGILQPSQTLSLTCSFS[GFSLSTY]GMGVSWIRQPSGKGLEWLAHI[-FWDG]DK
            *::* :*** : :*.:*:.::*. * *::::.* **  .*::* :****:*:. *   :.* :
             60         70          80          90
12.8_VH     TYADDFKGRFAFSLETSATTAYLQINNLKNEDMATYF[CVQSVIDNWG]QGTTLTVSS
1GGI_VH     RYNPSLKSRLKISKDTSNNQVFLKITSVDTADTATYY[CVQEGYIYWG]QGTSVTVSS
             *  .:*.*: :* :** . .:*:*..:.. * ***: ***.    ** ***::****
                    10          20          30          40          50
12.8_VL     QIVLTQSPAIMSASPGEKVTMTCSA[SSSID-----YM]HWYQQKPGTSPKRWIY[DTSKL]AS
1GGI_VL     DIVLTQSPGSLAVSLGQRATISCRA[SESVDDDGNSFL]HWYQQKPGQPPKLLIY[RSSNL]IS
            :*******. ::.* *::.*::* * *.*:*     :: ******** .**  **  :*:*  *
                60          70          80          90          100
12.8_VL     GVPARFSGSGSGTSYSLTISSMEAEDAATYYCHQ[RSSYPFTF]GSGTKLEIKR-
1GGI_VL     GIPDRFSGSGSRTDFTLTINPVEADDVATYYCQQ[SNEDPLTF]GAGTKLEIKRA
            *:* ******* *.::***..:**:*.*****:*  .. *:** *:********
```

C)

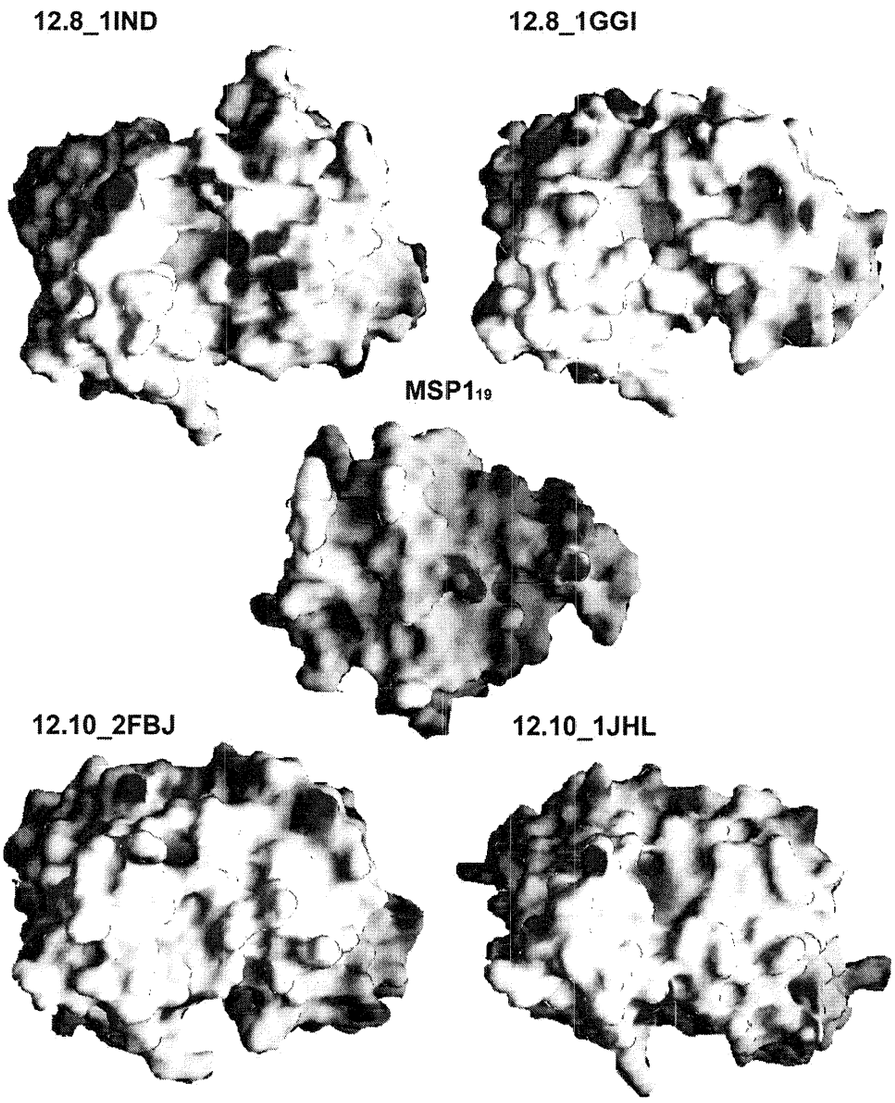
```
                    10          20          30          40          50          60
12.10_VH    QVQLQQPGAELVKPGASVKLSCKAS[GYTFTSY]WIHWVKQRPGQGLEWIGEI[IPNNG]RSNY
2FBJ_VH     EVKLLESGGGLVQPGGSLKLSCAAS[GFDFSKY]WMSWVRQAPGKGLEWIGEI[HPDSG]TINY
            :*:* :.*. **:**.*:**** ** *: *:.* *: **:* **:********  *:.*   **
                    70          80          90          100         110
12.10_VH    NEKFKSKATLTVDKSSSTAYMQLSSLTSEDSAVYF[CARNWAYWFFDVWG]AGTTVTVSS
2FBJ_VH     TPSLKDKFIISRDNAKNSLYLQMSKVRSEDTALYY[CARLHYYGYNAYWG]QGTLVTVS-
            . .:*.*  :: *::..: *:*:*.: ***:*:*: ***   * :   **  ** ****
                    10          20          30          40          50          60
12.10_VL    DIVMTQSHKFMSTSIGDRVSITCKA[SQDVSSA]VAWYQQRPGQSPKLLI[YSASY]RYTGVPD
2FBJ_VL     EIVLTQSPAITAASLGQKVTITCSA[SSSVSS-]LHWYQQKSGTSPKPWI[YEISK]LASGVPA
            :**:***  : ::*:*::*:***.* *..***  : ****:.* ***  * *. *    :***
                    70          80          90          100
12.10_VL    RFAGSGSGTDFTFTISSVQAEDLAVYYCQ[QHSSTPL]TFGAGTKLELK
2FBJ_VL     RFSGSGSGTSYSLTINTMEAEDAAIYYCQ[QWTYPLI]TFGAGTKLELK
            **:******.:::**.:::*** *:**** * : . : **********
```

D)

```
                    10          20          30          40          50          60
12.10_VH    QVQLQQPGAELVKPGASVKLSCKAS[GYTFTSY]WIHWVKQRPGQGLEWIGEI[IPNNG]RSNY
1JHL_VH     QVQLQQSGAELVRPGASVKLSCKAS[GYTFISY]WINWVKQRPGQGLEWIGNI[YPSDS]YTNY
            ******.*****:************ **** ** **:***************:*  *.:.  :**
                    70          80          90          100         110
12.10_VH    NEKFKSKATLTVDKSSSTAYMQLSSLTSEDSAVYF[CARNWAYWFFDVWG]AGTTVTVSS
1JHL_VH     NQKFKDKATLTVDKSSSTAYMQLSSPTSEDSAVYY[CTRDDNYGAMDYWG]QGTTVTV--
            *:***.****************** ********: *:*:  *  :* **  ******
                    10          20          30          40          50          60
12.10_VL    DIVMTQSHKFMSTSIGDRVSITCKA[SQDVSSA]VAWYQQRPGQSPKLLI[YSASY]RYTGVPD
1JHL_VL     DIELTQSPSYLVASPGETITINCRA[SKSISKS]LAWYQEKPGKTNNLLI[YSGST]LQSGIPS
            ** :*** .:: :* *: ::*.*:* *:.:*.: :****::**:: :*** **.*    :*:*.
                    70          80          90          100
12.10_VL    RFAGSGSGTDFTFTISSVQAEDLAVYYCQ[QHSSTPL]TFGAGTKLELK-
1JHL_VL     RFSGSGSGTDFTLTISSLEPEDFAMYICQ[QHNEYPW]TFGGGTKLEIKR
            **:*********:****::.**:*:* ** **.. *  ***.*****:*
```

Fig. 4. Sequence alignments of loop H3 of the antibodies 12.8 and 12.10 with loop H3 of the selected templates. The alignments were obtained using T-Coffee.[27] Residues are numbered according to Chotia et al.[52] The CDR regions are highlighted in parentheses and in italic format.

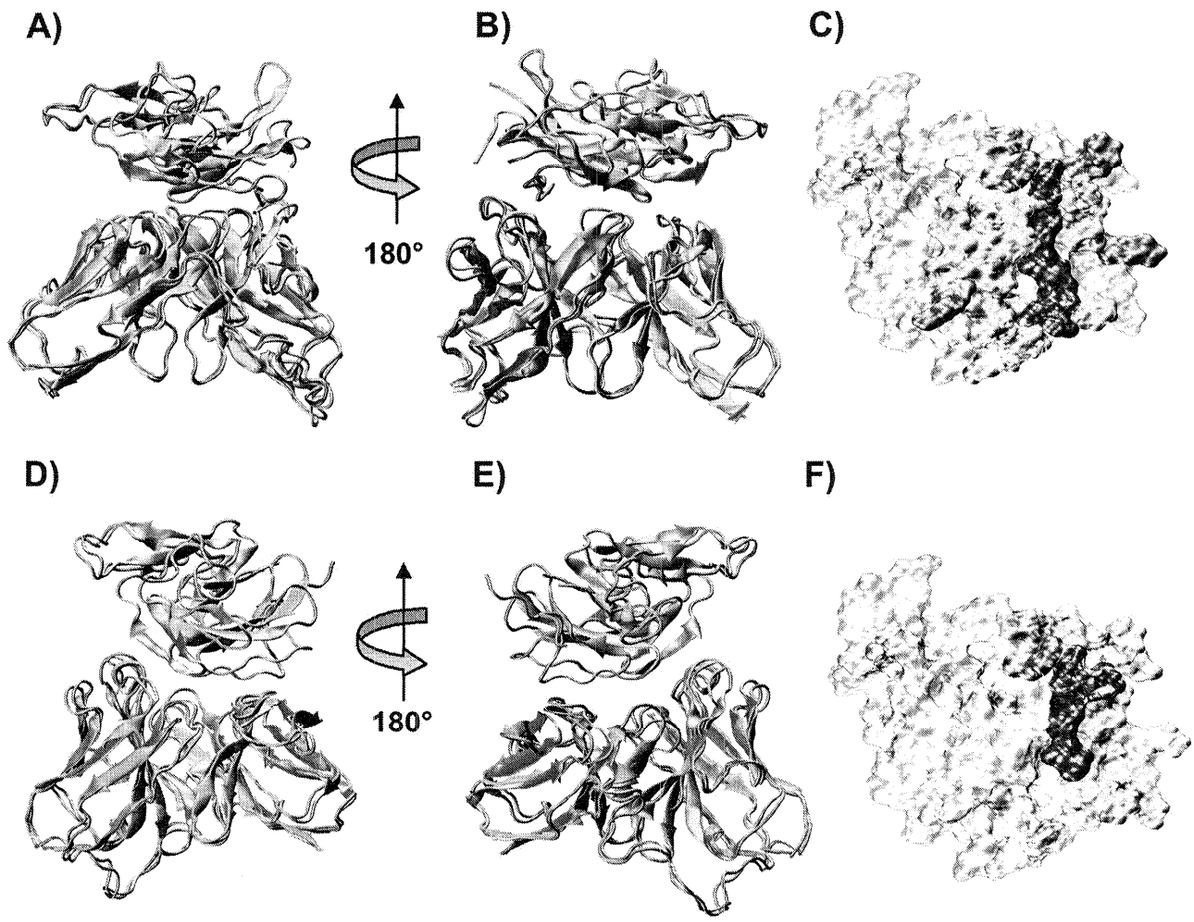
J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 9



COLOR

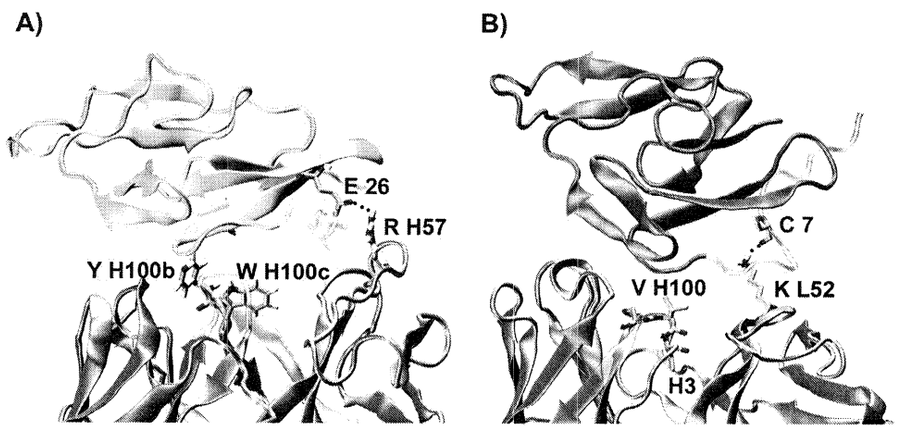
Fig. 5. Comparison of the electrostatic potential of the studied antibodies and $MSP1_{19}$. Each molecule is viewed from the top of the binding-site. The potential is ranged from −10 kT (red) to the maximal positive value +9 kT (blue).

fragments with perdeuterated, $^{13}C/^{15}N$-labeled $MSP1_{19}$, under conditions similar to those used previously.[18] Intermolecular NOEs from signals showing methyl group-region chemical shifts to the antigen could be identified in the spectrum of the mAb 12.8 complex. These must originate in the mAb fragment, since the aliphatic side chains in the antigen are perdeuterated, although as the mAb fragment cannot be isotope-labeled, there is no assignment information for this component (no NOEs with methyl region chemical shifts were observed with the Fab 12.10 complex). Two antigen residues that showed strong cross-saturation in the 12.8 complex, Asn 15 and Ser 16, showed backbone amide NOE cross peaks with a pair of methyl signals. One of these methyl signals showed an additional NOE to the side chain amide protons of a residue in the antigen, tentatively identified as Gln 14. From these data, we conclude that there is a close contact between antigen residues Asn 15 and Ser 16 and an aliphatic residue of the antibody, probably Val, Ile, or Leu. The 12.8 docking results were manually inspected according to this hypothesis. Among the selected complexes, none was found to have valine, isoleucine, or leucine in short reach from the amides of residues 15 and 16 and we therefore generated 1000 additional complexes for each antibody. On the surface of the CDRs, we can identify Ile 29 in loop L1, Leu 53 in loop L2, and Val 100 in loop H3 as possible candidates for the interaction in question. It should be noted that Ile L29 and Leu L53 are situated at



COLOR

Fig. 6. Panels (**A**), (**B**), and (**C**) refer to complex mAb 12.10-MSP1$_{19}$. (A) and (B) show the superimposition of the complex before (light blue) and after 5 ns of MD simulation (purple). Panel (C) displays the surface of the antigen buried upon complex formation, as calculated by POPS-comp.[44] The residues predicted by NMR experiments are displayed in blue, in yellow are shown the residues that, after MD simulation, still pass the NMR filter. In orange are displayed the remaining residues that bury a ΔSASA $\geq$ 10 Å^2 due to the formation of the complex (values shown in Table III). Panels (**D**), (**E**), and (**F**) refer to complex mAb 12.8-MSP1$_{19}$. (D) and (E) show the superimposition of the complexes before (orange) and after 5 ns of MD simulation (green). Panel (F) is colored like panel (C) (values shown in Table V).



COLOR

Fig. 7. Panel (**A**): Ribbon representation of complex mAb 12.10-MSP1$_{19}$. In licorice rendering Arg H57 (H2 loop) forming a salt bridge with Glu 26 of MSP1$_{19}$, and Trp H100c that shields the cavity in the central binding site region with negative electrostatic potential. Residues Tyr H100b and Trp H100c are found to be in contact with the antigen. Panel (**B**): Ribbon representation of complex mAb 12.8-MSP1$_{19}$. In licorice rendering, Lys L52 forming a hydrogen bond with the carbonyl oxygen of Cys 7 of MSP1$_{19}$. The central residues in loop H3 (Ser H99-Val H100-Ile H100a) are highlighted.

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 11

TABLE III. Residues of Antigen $MSP1_{19}$ in Contact With mAb 12.10_2FBJ After 5 ns of MD Simulation

(a) Accessible surface areas buried upon formation of the antigen-antibody complex

Antigen		Antibody		
Complex 12.10-$MSP1_{19}$	ΔSASA Å^2	Complex 12.10-$MSP1_{19}$	ΔSASA Å^2	CDR
ASN 1	29.1	SER H31	24.8	H1
ILE 2	52.6	TYR H33	25.8	H1
GLN 4	83.2	ILE H52	25.6	H2
IHS 5	10.2	ASN H54	56.9	H2
GLN 6	96.0	ASN H55	60.2	H2
GLN 11	20.8	ARG H57	172.9	—
CYS 12	22.1	ASN H59	63.7	—
GLN 14	50.8	TYR H100b	34.4	H3
ASN 15	39.4	ASP L28	27.8	L1
TYR 34	19.8	SER L30	23.2	L1
LYS 40	14.0	SER L67	22.0	—
GLU 43	34.4	SER L92	32.9	L3
TYR 84	16.0	SER L93	45.8	L3
PHE 87	18.9	THR L94	54.7	L3

(b) Residues complying to the NMR filter

Antigen	Antibody	
Residue	Residue	CDR
ASN 1	SER H31	H1
ILE 2	SER H31	H1
	TYR H33	—
SER 3	TYR H33	—
GLN 4	ARG H57	—
HIS 5	ARG H57	—
GLN 6	ARG H57	—
	ASN H59	—
CYS 7	ARG H57	—
VAL 8	ARG H57	—
	LYS H65	—
CYS 12	SER L93	L3
	THR L94	L3
PRO 13	THR L94	L3
GLN 14	SER L93	L3
	THR L94	L3
ASN 15	TYR H100b	H3
	SER L30	L1
	SER L92	L3
	SER L93	L3
	THR L94	L3
SER 16	TYR H100b	H3
	SER L30	L1
	SER L92	L3
	SER L93	L3
GLY 17	TYR H100b	H3
	TRP H100c	H3
	SER L93	L3
CYS 18	TRP H100c	H3
	SER L93	L3
LEU 31	TYR H100b	H3
GLU 43	SER L67	—
PRO 85	TYR H33	H1
LEU 86	TYR H33	H1
PHE 87	TYR H100b	H3
ASP 88	TYR H100b	H3

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 12

TABLE IV. Energy Decomposition Analysis of the Docked Complexes

(a) Energy decomposition for mAb12.10-$MSP1_{19}$

Complex 12.10-$MSP1_{19}$	Loops ZDOCK	Antibody ZDOCK	Loops 1 ns MD	Antibody 1 ns MD	Loops 5 ns MD	Antibody 5 ns MD
E1-SR: mAb-mAb	−4152.5	−9797.2	−4840.3	−10267.9	−5053.8	−10860.5
LJ: mAb-mAb	−3533.8	−7542.7	−3519.9	−7856.1	−3491.7	−7945.2
E1-14: mAb-mAb	12464.4	26577.5	13028.8	27635.1	12936.9	27557.7
LJ-14: mAb-mAb	−142.1	−317.6	206.3	401.8	146.4	356.4
E1-SR: mAb-$MSP1_{19}$	**−8.3**	**−8.6**	**−83.2**	**−84.4**	**−104.6**	**−104.0**
LJ: mAb-$MSP1_{19}$	**−85.0**	**−87.0**	**−248.7**	**−256.5**	**−329.7**	**−332.1**
E1-SR: $MSP1_{19}$-$MSP1_{19}$		−3528.9		−4342.9		−4180.7
LJ: $MSP1_{19}$-$MSP1_{19}$		−1637.8		−2811.6		−2929.7
E1-14: $MSP1_{19}$-$MSP1_{19}$		10646.7		10085.0		10067.6
LJ-14: $MSP1_{19}$-$MSP1_{19}$		1575.7		109.5		110.6

(b) Energy decomposition for mAb12.8-$MSP1_{19}$; same as for (a)

E1-SR: mAb-mAb	−1915.8	−10626.7	−2041.4	−10626.8	−1775.0	−10335.7
LJ: mAb-mAb	−1218.7	−6968.4	−1308.8	−7701.1	−1375.4	−7895.5
E1-14: mAb-mAb	5473.8	26280.2	5806.1	27300.3	5651.5	27255.6
LJ-14: mAb-mAb	−69.7	−356.4	157.5	498.3	93.4	344.1
E1-SR: mAb-$MSP1_{19}$	**1.7**	**28.6**	**−33.4**	**−35.0**	**−96.6**	**−116.6**
LJ: mAb-$MSP1_{19}$	**40.2**	**14250.8**	**−125.6**	**−131.2**	**−212.8**	**−220.8**
E1-SR: $MSP1_{19}$-$MSP1_{19}$		−3512.0		−4346.9		−4448.5
LJ: $MSP1_{19}$-$MSP1_{19}$		−1641.2		−2792.7		−2923.5
E1-14: $MSP1_{19}$-$MSP1_{19}$		10646.4		10173.9		10183.1
LJ-14: $MSP1_{19}$-$MSP1_{19}$		1576.8		97.7		68.8

The CDR loops have been decomposed from the entire antibody, in order to estimate their contribution to the binding. Values have been reported for the starting configuration (ZDOCK), after 1 ns of MD and at the end of the 5 ns of MD. The intermolecular energy contributions (antigen-antibody) are given in bold. Energies are given in units of kJ/mol.

the edge of the binding region. Therefore, to observe a contact between these residues and residues 15 and 16 of the antigen, the antigen itself would not be anymore centrally located with respect to the binding site, and in particular it would not be in contact with the H3 loop. With these concepts in mind we were able to select only one complex (from 12.8_1IND) with Asn 15 and Ser 16 of $MSP1_{19}$ at a short distance from residue Val H100 of the antibody. From the clustering procedure (as described in the Methods section) of the total of 17 acceptable complexes, our selected complex is number 10 (Fig. S4, Supplementary Materials). All complexes between 1 and 11 are equally good with respect to the clustering, but complex 10 is the only one among them that has the antigen centered relative to the antibody's binding loops.

For this complex, 5 ns of Molecular Dynamics simulations were performed and further analyses were done on the final complex. In Figure 6(D,E) the selected ZDOCK complex and the final MD conformer are shown superimposed. As in the previous cases, the MD simulation has the effect of pulling the antigen closer to the antibody. In Table Va the areas that are buried upon formation of the complex as calculated by POPScomp[44] are reported. Also in this case most of the antigen residues indicated by the NMR data are in contact with the antibody's heavy chain CDRs. Because loop H3 of antibody 12.8 is particularly short and lacks long side-chains, most of the interactions are found to form with loops H1 and H2. Table Vb shows that residue Gln 14 is in contact with Val H100, and this should represent one of the key interactions for this complex. The Val H100 side chain is still too far from the backbone amides of antigen residues Asn 15 and Ser 16 to generate a strong NOE, while residues Gly 17 and Cys 18, where cross-saturation was observed, are not sufficiently close to the antibody protons. These residues were close to the antibody in the ZDOCK complex, but moved away during the MD refinement. This effect could be ascribed to some repulsion between residues Lys 9 and Lys 10 of the antigen and residue Lys L52 of the antibody. The electrostatic interaction as described in the MD force field could amplify such an effect. The overall results are summarized in Figure 6(F). In Table IVb LJ and electrostatic contributions are reported analogously to Table IVa. As for the 12.10 complex, the electrostatic and intermolecular LJ contributions become increasingly more favorable during the simulated time (Table IVb), with larger contributions to the binding from CDR loops. We observe for this antibody a slightly different binding mode when compared with 12.10, with the antigen binding primarily to the first EGF-like domain to loops L2, H1 and H3. The antigen loop comprising residues 65–75 points toward loop H2 and is tilted with a larger distance from H2 than from L2.

CONCLUSIONS

We present here the docking refinement and analysis of two strongly inhibitory antibodies (designated 12.8

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 13

TABLE V. Residues of Antigen $MSP1_{19}$ in Contact With mAb 12.8 After 5 ns of MD Simulation

(a) Accessible surface areas buried upon formation of the antigen-antibody complex

Antigen		Antibody		
Complex 12.8-$MSP1_{19}$	ΔSASA Å^2	Complex 12.8-$MSP1_{19}$	ΔSASA Å^2	CDR
VAL 8	24.8	GLY H26	17.7	H1
LYS 9	34.0	TYR H27	10.4	H1
GLN 11	65.5	THR H28	57.3	H1
PRO 13	14.2	ASN H31	29.8	H1
GLN 14	113.9	TYR H32	47.5	H1
ASN 15	19.2	THR H53	10.7	H2
TYR 34	18.7	PHE H54	50.4	H2
ASP 39	20.3	VAL H100	35.1	H3
LYS 40	46.5	SER L51	18.9	L2
CYS 41	10.2	LYS L52	96.2	L2

(b) Residues complying to the NMR filter

Antigen	Antibody	
Residue	Residue	CDR
CYS 7	LYS L52	L2
VAL 8	LYS L52	L2
LYS 9	SER L51	L2
	LYS L52	L2
	LEU L53	L2
LYS 10	LEU L53	L2
GLN 11	TYR L48	—
	ASP L49	—
	LYS L52	L2
	LEU L53	L2
CYS 12	TYR L48	—
	LYS L52	L2
PRO 13	TYR H32	H1
	TYR L48	—
GLN 14	THR H28	H1
	TYR H32	H2
	VAL H100	H3
ASN 15	THR H28	H1
	TYR H32	H1
SER 16	THR H28	H1
	TYR H32	H1
GLY 38	GLN H1	—
ASP 39	GLN H1	—
	ILE H2	—
LYS 40	GLN H1	—
	ILE H2	—
	GLY II26	H1
	TYR H27	H1
	THR H28	H1
	TYR H32	H1
CYS 41	TYR H27	H1
	THR H28	H1
	TYR H32	L1
	SER H93	H3
VAL 42	TYR H27	H1
	THR H28	H1
GLU 43	THR H28	H1

and 12.10) to the antigen structure $MSP1_{19}$ as solved by NMR.[12] The antibodies were modeled by comparative homology modeling and the docking results were screened by selecting the structures with the help of data from NMR cross-saturation and chemical shift perturbation experiments,[18,19] complemented by additional

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 14

NOE experiments. Molecular Dynamics simulations in explicit solvent were performed on the selected complexes to refine their structures to evaluate the role of specific side chains in the binding. The trial exercise using the Fab G17.12-$MSP1_{19}$ complex X-ray structure indicated the feasibility of obtaining successfully docked structures using this NMR-based docking/MD approach. With this exercise we mainly wanted to test our filter of structures based on external restraints. Moreover, the additional docking performed with the modeled Fab G17.12 gave rise to a complex very similar to the one obtained with the original X-ray structure. This finding supports the use of models in the case of absence of X-ray structures for the antibody, but this procedure is of course limited to a sufficiently high sequence identity between the sequences of model and template, in particular in the binding region.

This issue is very important, especially in relation to new structural genomics/system biology initiatives that attempt the structural reconstruction of protein–protein interaction networks.[53]

Our computed complexes show two slightly different binding modes for the two antibodies 12.10 and 12.8: complex 12.10-$MSP1_{19}$ shows antigen binding with one face of its discoidal shape, lying flat on the surface of the CDRs; complex 12.8-$MSP1_{19}$ presents the antigen in a tilted position, anchored to the antibody mainly through its first EGF-like domain. The second EGF domain is more distant from the H2 loop. It is noteworthy that previous experimental evidences had pointed out that antibody 12.10 can only bind in the presence of both $MSP1_{19}$ EGF domains, while for 12.8 only the presence of domain 1 is sufficient to observe binding.[54] For this antibody we stress that loop H3 is particularly short and without long side-chain residues that could be specific for the binding. This could be the cause of the observed binding to mainly one EGF-like domain of $MSP1_{19}$.

From a detailed analysis of our docked complexes we underline the importance of a positively charged residue on the surface of the antibody. In particular antibody 12.10 has residue Arg H57 on the H2 loop that forms a salt bridge with Glu 26 of the antigen [Fig. 7(A)]. Antibody 12.8 contains residue Lys L52 that forms a hydrogen bond with the carbonyl oxygen of Cys 7 of $MSP1_{19}$ [Fig. 7(B)]. Our analysis pointed out the necessity of a positively charged electrostatic spot on the CDRs surface that seems to be a general requirement for the recognition and binding of this antigen. Accordingly also mAb G17.12 has been observed to form a salt bridge with $MSP1_{19}$ through its residue Arg H98.[15]

Each docking experiment produced a limited number of results consistent with the NMR data and the general expectation of typical antibody binding modes (recognition through the CDR loops). This also in reason of the few restraints obtained from cross-saturation experiments. In principle a few restraints would be enough to define the binding mode precisely if these were distributed about the vertices of a triangle covering one face of the antigen. Once a binding mode is adopted having one of the discoidal faces lying flat on the surface of the CDRs, we can observe that the NMR data are concentrated on one half of this surface, therefore the docked solutions are only mildly restrained [Fig. 6(C,F)]. For the antibody 12.10 we had two more restraints (on a distant spot from the main ones (HN 6, 7 *vs* HN 15, 16, 17, 18)) and therefore the docking procedure proved to be more straightforward.

Although it would be desirable to crystallize one or both of the 12.8 and 12.10 mAb complexes to verify the modeled structures, useful information can be derived from the current models. They provide a more detailed picture of the binding interface, and predict intermolecular interactions that can be used, for example, as a guide for mutagenesis work. The models will be valuable for understanding the mechanism of action of invasion-inhibitory antibodies and for studying therapeutic effects of antibody reagents derived from these antibody sequences.

ACKNOWLEDGMENTS

This work was supported by Medical Research Council and by European Commission through the EUROMALVAC consortium. F. A. was supported by CEINGE Advanced Biotechnology consortium. The authors thank Annalisa Pastore and Andres Ramos for the use of computer facilities.

REFERENCES

1. Engers HD, Godal T. Malaria vaccine development: current status. Parasitol Today 1998;14:56–64.
2. Holder AA. Proteins on the surface of the malaria parasite and cell invasion. Parasitology 1994;108 Suppl:S5–S18.
3. Barnwell JW, Galinski MR. The adhesion of malaria merozoite proteins to erythrocytes: a reflection of function? Res Immunol 1991;142:666–672.
4. Holder AA, Guevara Patino JA, Uthaipibull C, Syed SE, Ling IT, Scott-Finnigan T, Blackman MJ. Merozoite surface protein 1, immune evasion, and vaccines against asexual blood stage malaria. Parassitologia 1999;41:409–414.
5. McBride JS, Heidrich HG. Fragments of the polymorphic Mr 185,000 glycoprotein from the surface of isolated *Plasmodium falciparum* merozoites form an antigenic complex. Mol Biochem Parasitol 1987;23:71–84.
6. Holder AA, Sandhu JS, Hillman Y, Davey LS, Nicholls SC, Cooper H, Lockyer MJ. Processing of the precursor to the major merozoite surface antigens of *Plasmodium falciparum*. Parasitology 1987:94:199–208.
7. Holder AA, Riley EM. Human immune response to MSP-1. Parasitol Today 1996;12:173,174.
8. Blackman MJ, Heidrich HG, Donachie S, McBride JS, Holder AA. A single fragment of a malaria merozoite surface protein remains on the parasite during red cell invasion and is the target of invasion-inhibiting antibodies. J Exp Med 1990;172:379–382.
9. Blackman MJ, Scott-Finnigan TJ, Shai S, Holder AA. Antibodies inhibit the protease-mediated processing of a malaria merozoite surface protein. J Exp Med 1994;180:389–393.
10. Guevara Patino JA, Holder AA, McBride JS, Blackman MJ. Antibodies that inhibit malaria merozoite surface protein-1 processing and erythrocyte invasion are blocked by naturally acquired human antibodies. J Exp Med 1997;186:1689–1699.
11. Nwuba RI, Sodeinde O, Anumudu CI, Omosun YO, Odaibo AB, Holder AA, Nwagwu M. The human immune response to *Plasmodium falciparum* includes both antibodies that inhibit merozoite surface protein-1 secondary processing and blocking antibodies. Infect Immun 2002;70:5328–5331.
12. Morgan WD, Birdsall B, Frenkiel TA, Gradwell MG, Burghaus PA, Syed SE, Uthaipibull C, Holder AA, Feeney J. Solution

J_ID: Z7E Customer A_ID: 21212 Cadmus Art: PROT2970 Date: 20-NOVEMBER-06 Stage: I Page: 15

structure of an EGF module pair from the *Plasmodium falciparum* merozoite surface protein 1. J Mol Biol 1999;289:113–122.
13. Chitarra V, Holm I, Bentley GA, Pétres S, Longacre S. The crystal structure of C-terminal merozoite surface protein-1 at 1.8 A resolution, a highly protective malaria vaccine candidate. Mol Cell 1999;3:457–464.
14. Garman SC, Simcoke WN, Stowers AW, Garboczi DN. Structure of the C-terminal domains of merozoite surface protein-1 from *Plasmodium knowlesi* reveals a novel histidine binding site. J Biol Chem 2003;278:7264–7269.
15. Pizarro JC, Chitarra V, Verger D, Holm I, Pétres S, Dartevelle S, Nato F, Longacre S, Bentley GA. Crystal structure of a Fab complex formed with PfMSP1–19, the C-terminal fragment of merozoite surface protein 1 from *Plasmodium falciparum*: a malaria vaccine candidate. J Mol Biol 2003;328:1091–1103.
16. Jennings RM.Murine and human antibody responses to *Plasmodium falciparum* merozoite surface protein-1. PhD Thesis, London: University College London; 2004.
17. Uthaipibull C, Aufiero B, Syed SE, Hansen B, Guevara Patino JA, Angov E, Ling IT, Fegeding K, Morgan WD, Ockenhouse C, Birdsall B, Feeney J, Lyon JA, Holder AA. Inhibitory and blocking monoclonal antibody epitopes on merozoite surface protein 1 of the malaria parasite *Plasmodium falciparum*. J Mol Biol 2001;307:1381–1394.
18. Morgan WD, Frenkiel TA, Lock MJ, Grainger M, Holder AA. Precise epitope mapping of malaria parasite inhibitory antibodies by TROSY NMR cross-saturation. Biochemistry 2005;44:518–523.
19. Morgan WD, Lock MJ, Frenkiel TA, Grainger M, Holder AA. Malaria parasite-inhibitory antibody epitopes on *Plasmodium falciparum* merozoite surface protein-1(19) mapped by TROSY NMR. Mol Biochem Parasitol 2004;138:29–36.
20. Wodak SJ, Méndez R. Prediction of protein-protein interactions: the CAPRI experiment, its evaluation and implications. Curr Opin Struct Biol 2004;14:242–249.
21. Chen R, Li L, Weng Z. ZDOCK: an initial-stage protein-docking algorithm. Proteins 2003;52:80–87.
22. Chen R, Weng Z. A novel shape complementarity scoring function for protein-protein docking. Proteins 2003;51:397–408.
23. Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: A message-passing parallel molecular dynamics implementation. Comput Phys Commun 1995;91:43–56.
24. Chen R, Tong W, Mintseris J, Li L, Weng Z. ZDOCK predictions for the CAPRI challenge. Proteins 2003;52:68–73.
25. Ehrlich LP, Nilges M, Wade RC. The impact of protein flexibility on protein-protein docking. Proteins 2005;58:126–133.
26. Smith GS, Sternberg MJE, Bates PA. The relationship between the flexibility of proteins and their conformational states on forming protein-protein complexes with an application to protein-protein docking. J Mol Biol 2005;347:1077–1101.
27. Notredame C, Higgins DG, Heringa J. T-Coffee: a novel method for fast and accurate multiple sequence alignment. J Mol Biol 2000;302:205–217.
28. Huang X, Miller W. A time-efficient, linear-space local similarity algorithm. Adv Appl Math 1991;12:337–357.
29. Pearson WR, Lipman DJ. Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 1988;85:2444–2448.
30. Schwede T, Kopp J, Guex N, Peitsch MC. SWISS-MODEL: an automated protein homology-modeling server. Nucleic Acids Res 2003;31:3381–3385.
31. Suh SW, Bhat TN, Navia MA, Cohen GH, Rao DN, Rudikoff S, Davies DR. The galactan-binding immunoglobulin Fab J539: an X-ray diffraction study at 2.6-A resolution. Proteins 1986;1:74–80.
32. Chitarra V, Alzari PM, Bentley GA, Bhat TN, Eiselé JL, Houdusse A, Lescar J, Souchon H, Poljak RJ. Three-dimensional structure of a heteroclitic antigen-antibody cross-reaction complex. Proc Natl Acad Sci USA 1993;90:7711–7715.
33. Love RA, Villafranca JE, Aust RM, Nakamura KK, Jue RA, Major JG, Radhakrishnan R, Butler WF. How the anti-(metal chelate) antibody CHA255 is specific for the metal ion of its antigen: X-ray structures for two Fab'/hapten complexes with different metals in the chelate. Biochemistry 1993;32:10950–10959.
34. Alzari PM, Spinelli S, Mariuzza RA, Boulot G, Poljak RJ, Jarvis JM, Milstein C. Three-dimensional structure determination of an anti-2-phenyloxazolone antibody: the role of somatic mutation and heavy/light chain pairing in the maturation of an immune response. EMBO J 1990;9:3807–3814.
35. Chen R, Weng Z. Docking unbound proteins using shape complementarity, desolvation, and electrostatics. Proteins 2002;47: 281–294.
36. Kleinjung J, Romein J, Lin K, Heringa J. Contact-based sequence alignment. Nucleic Acids Res 2004;32:2464–2473.
37. Team RDC. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2005.
38. Daura X, Mark AE, van Gunsteren WF. Parametrization of aliphatic CHn united atoms of GROMOS96 force field. J Comput Chem 1998;19:535–547.
39. Berendsen HJC, Postma JPM, van Gunsteren WF, Hermans J, Pullman B.Interaction models for water in relation to protein hydration. In:Pullman B, editor. Intermolecular forces. Dordrecht: Reidel; 1981. pp 331–342.
40. Berendsen HJC, Postma JPM, van Gunsteren WF, Dinola A, Haak JR. Molecular dynamics with coupling to an external bath. J Chem Phys 1984;81:3684–3690.
41. Hess B, Bekker H, Berendsen HJC, Fraaije J. LINCS: a linear constraint solver for molecular simulations. J Comput Chem 1997;18:1463–1472.
42. Darden T, Perera L, Li L, Pedersen L. New tricks for modelers from the crystallography toolkit: the particle mesh Ewald algorithm and its use in nucleic acid simulations. Struct Fold Des 1999;7(3).
43. Tironi IG, Sperb R, Smith PE, Vangunsteren WF. A generalized reaction field method for molecular-dynamics simulations. J Chem Phys 1995;102:5451–5459.
44. Kleinjung J, Fraternali F. POPSCOMP: an automated interaction analysis of biomolecular complexes. Nucleic Acids Res 2005; 33:W342–W346. (Web server issue).
45. Fraternali F, Cavallo L. Parameter optimized surfaces (POPS): analysis of key interactions and conformational changes in the ribosome. Nucleic Acids Res 2002;30:2950–2960.
46. Cavallo L, Kleinjung J, Fraternali F. POPS: a fast algorithm for solvent accessible surface areas at atomic and residue level. Nucleic Acids Res 2003;31:3364–3366.
47. Morea V, Tramontano A, Rustici M, Chothia C, Lesk AM. Conformations of the third hypervariable region in the VH domain of immunoglobulins. J Mol Biol 1998;275:269–294.
48. Tormo J, Stadler E, Skern T, Auer H, Kanzler O, Betzel C, Blaas D, Fita I. Three-dimensional structure of the Fab fragment of a neutralizing antibody to human rhinovirus serotype 2. Protein Sci 1992;1:1154–1161.
49. Saul FA, Poljak RJ. Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 A resolution. Proteins 1992;14: 363–371.
50. Zdanov A, Li Y, Bundle DR, Deng SJ, MacKenzie CR, Narang SA, Young NM, Cygler M. Structure of a single-chain antibody variable domain (Fv) fragment complexed with a carbohydrate antigen at 1.7-A resolution. Proc Natl Acad Sci USA 1994;91: 6423–6427.
51. Prasad L, Sharma S, Vandonselaar M, Quail JW, Lee JS, Waygood EB, Wilson KS, Dauter Z, Delbaere LT. Evaluation of mutagenesis for epitope mapping. Structure of an antibody-protein antigen complex. J Biol Chem 1993;268:10705–10708.
52. Chothia C, Lesk AM, Tramontano A, Levitt M, Smith-Gill SJ, Air G, Sherriff S, Padlan EA, Davies D, Tulip WR, Colman PM, Spinelli S, Alzari PM, Poljak J. Conformations of immunoglobulin hypervariable regions. Nature 1989;342:877–883.
53. Aloy P, Russel RB. Structural systems biology: modelling protein interactions. Nat Rev Mol Cell Biol 2006;7:188–197.
54. Chappel JA, Holder AA. Monoclonal antibodies that inhibit *Plasmodium falciparum* invasion in vitro recognise the first growth factor-like domain of merozoite surface protein-1. Mol Biochem Parasitol 1993;60:303–311.