Acta Cryst. (1991). B47, 549–555

Use of Molecular Replacement in the Solution of an Immunoglobulin Fab Fragment Structure

BY T. BIZEBARD

Laboratoire de Biologie Physicochimique, CNRS URA 1131, *Université Paris-Sud, Bat* 433, 91405 *Orsay CEDEX, France, and Laboratoire de Physique, Centre Pharmaceutique, CNRS UPR* 180, 92296 *Chatenay Malabry CEDEX, France*

Y. MAUGUEN

Laboratoire de Physique, Centre Pharmaceutique, CNRS UPR 180, 92296 *Chatenay Malabry CEDEX, France*

J. J. SKEHEL

National Institute of Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, England

AND M. KNOSSOW

Laboratoire de Biologie Physicochimique, CNRS URA 1131, *Université Paris-Sud, Bat* 433, 91405 *Orsay CEDEX, France*

(*Received* 22 *October* 1990; *accepted* 4 *February* 1991)

Abstract

Molecular-replacement efficiency depends highly on structural and sequence homologies between available models and the molecule in the crystal being studied. The structure of the Fab fragment of an antibody specific for an influenza virus hemagglutinin was determined by molecular replacement and the Fv and the CH1:CL parts were localized separately. When rotation functions were calculated using known Fv structures as probes, a solution could not be found; this turns out to be due to an insufficient structural homology between the structure and the probes. When the structural homology between the Fv part and its model was enhanced by combining known structures of Fv domains based on sequence information, the right orientation was determined and confirmed by translation-function results. In the cases described here, a high contrast of the translation function was the most reliable criterion to detect a molecular-replacement solution.

Introduction

An often followed procedure to determine the three-dimensional structure of a protein involves multiple isomorphous replacement; given the uncertainty of finding adequate heavy-atom derivatives, this can be a lengthy process. The growing number of proteins whose three-dimensional structure has been determined increases the probability that a new protein being investigated will have a structure close enough to a known one. Molecular replacement (MR) can then be used, thus decreasing significantly the time required to solve the structure.

The objective of MR methods is to locate a molecule in an unknown structure by using a model which is supposed to deviate as little as possible from that molecule [for a recent review on molecular replacement, see Rossmann (1990)]. This location involves the determination of six parameters, three rotations and three translations. When those have been determined, one can position the model in the unit cell of the unknown structure, and calculated phases can then be used to complete the determination of the structure.

The likelihood of finding a model suitable for molecular replacement is highest in 'superfamilies' of proteins; this is true in the case of Fab fragments, which belong to the immunoglobulin superfamily; immunoglobulins are characterized by a common fold, several of their structures are known and coordinates have been deposited in the Protein Data Bank (Bernstein, Koetzle, Williams, Meyer, Brice, Rodgers, Kennard, Shimanouchi & Tasumi, 1977). Although it is probably the fastest method in such cases, problems arise in using MR in the solution of Fab fragments structures because of the variability of their sequences. That variability can be correlated to the classification of immunoglobulin polypeptide chains. Each immunoglobulin is characterized by the class of its two polypeptide chains (heavy and light),

0108-7681/91/040549-07$03.00

which can easily be determined by immunological methods (Friguet, Djavadi-Ohaniance & Goldberg, 1989). The light chain comprises two domains and the heavy chain is made up of at least four domains; Fab fragments are constituted of an immunoglobulin light chain associated with the two NH_2-terminal domains of a heavy chain. Whereas the sequence and, therefore, the structural similarity between homologous domains in a given class can be very high, it decreases significantly across classes.

In order to gain more insight into the molecular mechanism which allows influenza viruses to escape neutralization by the humoral immune response (Wiley & Skehel, 1987), we have undertaken to determine the three-dimensional structure of the Fab fragment of a mouse antibody (HC19) directed against the main viral antigen, influenza hemagglutinin (Hong Kong/68/H3N2). This should enable us to correlate the conformation of the antibody combining site to the nature of the mutation(s) that allow the virus to escape recognition by HC19 immunoglobulin. One such mutation has already been identified as the change Ser→Leu of amino acid 157 of the hemagglutinin (Daniels, Douglas, Skehel & Wiley, 1983).

The HC19 immunoglobulin belongs to class IgG1,λ for which no structure has been published so far. We show here that, using the knowledge of the antibody's class, we were able to build a model that led to a solution of that structure by MR; we also discuss the criteria that allowed us to detect that solution.

Models

Fab fragments are obtained by limited proteolysis of immunoglobulins and comprise the NH_2-terminal parts of the two polypeptide chains [noted light (L) and heavy (H)] that constitute these molecules. Both polypeptide chains consist of compact domains, approximately 110 amino acids long, related by elongated segments. There are two such domains in each chain of a Fab fragment. The NH_2-terminal domains of both chains are more variable and are named variable (VL and VH for light and heavy chain respectively), whereas the other domains are noted constant (CL and CH1 for those that contribute to the Fab fragment). All domains have the characteristic immunoglobulin fold.

Structural models of the domains of immunoglobulins are provided by published structures of Fab fragments; seven of these have had their coordinates deposited in the Protein Data Bank. These are: KOL (Marquart, Deisenhofer, Huber & Palm, 1980), NEW (Saul, Amzel & Poljak, 1978), McPC603 (Segal, Padlan, Cohen, Rudikoff, Potter & Davies, 1974), J539 (Suh, Bhat, Naira, Cohen, Rudikoff, Rao & Davies, 1986), HyHEL5 (Sheriff, Silverton, Padlan, Cohen, Smith-Gill, Finzel & Davies, 1987), HyHEL10 (Padlan, Silverton, Sheriff, Cohen, Smith-Gill & Davies, 1989) and R19.9 (Lascombe, Alzari, Boulot, Saludjian, Tougar, Berek, Haba, Rosen, Nisonoff & Poljak, 1989). Coordinates of two other Fab fragments have been provided to us prior to publication: D1.3 (T. Fischmann & R. J. Poljak, 1990, personal communication) and NC41 (W. R. Tulip & P. M. Colman, 1990, personal communication). These nine data sets provide the basis for our MR studies.

The heavy and light chains of a Fab fragment associate through non-covalent interactions. Such interactions are most extensive between CL and CH1 domains and between VL and VH domains; the (CL,CH1) and (VL,VH) dimers constitute distinct structural subunits in the Fab fragment, and they each have approximate twofold symmetry which relates the light and heavy chains. The angle between the approximate twofold axes is called the elbow angle and it varies from one Fab fragment to another; the range covered so far in published structures is between 115 and 180° (Huber & Bennett, 1987). Because of this quaternary structural variability, the models which were found most useful in MR studies of Fab fragments are the constant and variable domain dimers (Cygler & Anderson, 1988*a*). They have been used in this work.

The likelihood of being able to match a model to an unknown structure is expected to be highly correlated to the sequence homology of the proteins involved. As we have already mentioned, sequence homology is highest among immunoglobulin domains belonging to the same class or subclass. HC19 clone belongs to the IgG1,λ class. No Fab fragment of known structure shares this subtype (Table 1). After initial failure in locating the variable-domain dimer by MR using variable-domain dimers of published structures, we have built mixed dimers using variable domains from different structures, to maximize the sequence homology between our model and the HC19 variable-domain dimer. No structure of a Fab fragment from a mouse immunoglobulin comprising a λ light chain has been published so far; we therefore used a λ human VL model for our VL domain because the homology of human λ class VL domain is higher than that of mouse κ class to mouse λ class VL domain (40% compared to 30%, as determined after manual alignment of representative sequences). Two human λ VL models are available in the Protein Data Bank, but we preferred KOL VL domain chain since NEW light-chain variable domain has a seven-residue atypical deletion (Poljak, Amzel, Avey, Chen, Phizackerley & Saul, 1973). Using the mouse Fab fragment structures that have already been determined, we built three different

Table 1. *Fab fragment subtypes for structures providing the models for molecular replacement in this study*

Fab fragment	Species	Class	
		Heavy chain	Light chain
KOL	Human	G	λ
NEW	Human	G	λ
MCPC603	Mouse	A2	κ
J539	Mouse	A2	κ
HYHEL5	Mouse	G1	κ
HYHEL10	Mouse	G1	κ
R19.9	Mouse	G2b	κ
D1.3	Mouse	G1	κ
NC41	Mouse	G2a	κ

Table 2. *Translation-function results*

All peak heights expressed as multiples of the r.m.s.d. of the translation function.

Domain	Model	Space group	Resolution range (Å)	First peak height H1	Second peak height H2	Contrast H1/H2
C	KOL	$P3_121$	10–6	5·7*	4·55	1·25
C	KOL	$P3_121$	10–4	10·5*	6·15	1·71
C	KOL	$P3_221$	10–4	5·55†	4·85	1·14

* The solution corresponds to the true peak $x = 0{\cdot}22$, $y = 0{\cdot}92$, $z = 0{\cdot}12$.

† This peak corresponds to the translation $x = 0{\cdot}0$, $y = 0{\cdot}92$, $z = 0{\cdot}45$ (translations are expressed in fractional coordinates).

mixed-dimer models, using the VH subunit of HyHEL5, HyHEL10 and D1.3 immunoglobulins associated with KOL VL subunit.

In building 'hybrid' models we have taken into account the fact that the association of VL and VH domains has been found to be well conserved in Fab fragments of published structures, at least at the resolution involved in this work (Lascombe, 1989). The mouse VH subunit was fitted onto Fab KOL VH subunit by least-squares superposition of the Cα atoms of the framework residues (*i.e.* the structurally best conserved ones among all Fab fragment three-dimensional structures; 40 Cα atoms were used; r.m.s.d. of atomic positions between the two molecules being compared *ca* 0·7 Å). In an attempt to improve the fit of Fab KOL VL subunit in our model molecule to a λ mouse VL domain, we discarded all atoms beyond Cβ belonging to residues that, based on sequence comparisons, were very likely to be different in Fab KOL VL subunit and in a mouse λ VL subunit [sequences from Kabat, Wu, Bilofsky, Reid-Miller & Perry (1983)]. This model will be referred to as KOL 'modified' VL.

One last point needs to be mentioned. Variable domains contain hypervariable loops (three in each domain) which differ considerably from one antibody to another, and constitute the antigen combining site; all variable-domain-dimer models were tested with or without hypervariable loops (also noted complementary determining regions or CDR's), defined following Chothia, Lesk, Tramontano, Levitt, Smith-Gill, Air, Sheriff, Padlan, Davies, Tulip, Colman, Spinelli, Alzari & Poljak (1989).

Structure determination

As previously reported (Bizebard, Mauguen, Petek, Rigolet, Skehel & Knossow, 1990), the Fab fragment from clone HC19 crystallizes in space group $P3_121$ (or $P3_221$), cell constants being $a = b = 98{\cdot}9$, $c = 89{\cdot}2$ Å; data to 3·5 Å resolution have been collected using a Xentronics area detector (6147 unique reflections – 89% of the theoretical complete data – were merged from 22 929 observations; $R_{sym} = 6{\cdot}7\%$ for all data; $R_{sym} = 9{\cdot}4\%$ for data in the 4–3·5 Å resolution range). These data were used in all MR studies.

Exploration of the complete six-parameter space associated with MR would involve prohibitive computing times. Fortunately, it has been operationally shown that the problem can be separated into two steps; namely determination of the orientation and position of the molecule, each involving only three parameters. The possibility of separating these steps is associated with the fact that in each of them different interatomic vectors of the structure and of the model are compared, which correspond to different zones of the Patterson function space. They are respectively the intramolecular and intermolecular vectors, in the rotation and translation functions (Rossmann & Blow, 1962).

Parameters of the rotation-function calculations

Rotation functions were calculated by evaluating the overlap of Patterson functions of the model and of the unknown molecule within a sphere. The radius of that sphere is adjusted in an attempt to include most of the intramolecular Patterson vectors and the least of the intermolecular ones. Fast algorithms to calculate a rotation function in reciprocal space have been developed (Crowther, 1972) and allowed us to test various models, initially using computing conditions close to those already employed in the MR study of Fab fragment structures (Cygler & Anderson, 1988*a*): radius for overlap integration 24 Å, resolution range of reflections used 10–4 Å, the strongest 30% of the reflections in that range being used. The rotation function was computed in its angular asymmetric unit: $0 \le \alpha < 120°$, $0 \le \beta \le 90°$, $0 \le \gamma < 360°$ (Moss, 1985). The Eulerian angle sampling was: α and γ every 5°, β every 2·5°.

Parameters of the translation-function calculations

For any given orientation of the model, a translation search can be performed, and several methods are available to do so. Cygler & Anderson have examined some of these in the case of a model molecule representing at least 50% of the content of

Table 3. *C-domain rotation-function results (model used: Fab KOL)*

All peak heights expressed as multiples of the r.m.s.d. of the rotation function.

Outer integration radius (Å)	Resolution range (Å)	Reflections used (%)	Correct peak peak height H1	First spurious peak height H2	Contrast H1/H2
20	10–4	30	4·05	3·35	1·21
24	10–4	30	4·65	3·75	1·24
28	10–4	30	4·35	3·6	1·21
24	10–5	30	< 2·8	2·9	< 0·95
24	10–3·5	30	4·1	3·1	1·32
24	10–4	70	4·5	3·75	1·20

the asymmetric unit. They conclude that, under such conditions, the choice of the precise algorithm used is not critical (Cygler & Anderson, 1988*b*). We have used the Crowther–Blow translation function (Crowther & Blow, 1967) as programmed by Tickle (1985). Other methods – *e.g.* searches for maxima of the correlation coefficient between observed and calculated structure factors (Fujinaga & Read, 1987) – are available.

Calculations were initially performed in the 10–4 Å resolution range and all reflections in that shell were included. For each orientation tested, a systematic translation search was performed by varying each of the three Eulerian angles every degree in the range $(-2°, +2°)$ around the rotation-function peak. The translation function was sampled approximately every 1 Å along the three cell edges, in the Cheshire group unit cell: $0 \le x < 1$, $0 \le y < 1$, $0 \le z \le \frac{1}{2}$ (Hirshfeld, 1968). Among all models used in rotation-function calculations, we selected for the translation function the one giving rise to the highest signal to noise ratio [*i.e.* peak height divided by the root-mean-square deviation (r.m.s.d.) of rotation-function values] in the rotation function for the orientation being tested. In each case, both enantiomorphic space groups, $P3_121$ and $P3_221$, were tested. As discussed later, translation-function calculations were performed independently for the variable and constant domains of the Fab fragment. Since in space group $P3_121$ (or $P3_221$) there are two possible origins, related by a $(0,0,\frac{1}{2})$ translation, the Crowther–Blow translation function leaves an ambiguity in the relative positions of solutions provided by these independent calculations; the selection of compatible origins in calculations locating the two domains was made considering the requirement for polypeptide-chain continuity in the whole Fab fragment.

Results of molecular replacement – constant-domain dimer

A contrasted solution was found using KOL constant-domain dimer as a model and conditions as described above; the likelihood of this corresponding to the orientation of the C-domain dimer was strengthened when a translation function was calculated, in space group $P3_121$, using the KOL C dimer oriented as defined by the rotation function. This translation function has a single peak of 10·5 times its r.m.s.d., whereas the next-highest peak is of 6·15 r.m.s.d. height. Since the contrast is much lower in the enantiomorphic space group, this indicates that the true space group of the structure is $P3_121$ (Table 2).

Starting from this solution, we have analyzed the influence of the various parameters of the rotation-function calculation on the contrast of its solution (Table 3). The following conclusions can be drawn from these tests:

(*a*) The radius for overlap integration does not seem to be a critical parameter as long as it remains comparable to molecular dimensions (see Table 3 lines 1 to 3). Therefore, in all subsequent calculations, we have chosen a 24 Å integration radius.

(*b*) We have never selected reflections below 10 Å resolution, following the rationale of Blow (1985). This outlines that the very-low-order terms (spacings greater than 10 Å) are greatly influenced by the solvent continuum, which is not taken into account in the model structure. Our results confirm those of Cygler & Anderson (1988*a*) that inclusion of data in the 5–4 Å resolution shell dramatically improves the signal-to-noise ratio of the rotation function; in our case, these data bring the right orientation to the first rank of the rotation-function peaks (Table 3, lines 2 and 4). The 4–3·5 Å shell brings further, though more limited, improvement (Table 3, lines 2 and 5) but, considering that use of these data would lengthen the calculation, we ordinarily limited the resolution range to 10–4 Å.

(*c*) Results when the strongest 30% of the reflections or the strongest 70% of the reflections are used are very comparable (see Table 3, lines 2 and 6). In subsequent calculations, we generally used the 30% most intense reflections, for both the model and the unknown structure.

Using all these conditions, we have compared rotation-function results using different C-domain models. The results are summarized in Table 4. The correct rotation peak is the most frequently found as the first peak, but it is most contrasted and deviates

Table 4. *C-domain rotation-function results*

Integration radii, 4–24 Å; resolution range, 10–4 Å. All peak heights expressed as mutliples of the r.m.s.d. of the rotation function.

Model	True peak height H1	Position in list	First spurious peak height H2	Contrast first peak/ second peak
KOL	4·65	First	3·75	1·24
NEW	4·0	First	3·2	1·24
HYHEL5	3·4	Sixth	4·0	1·02
HYHEL10	< 2·8	?	4·2	1·09
MCPC603	3·75	First	3·5	1·06
J539	3·2	Third	3·4	1·01

least from our present refined orientation when Fab fragments KOL and NEW are used as models (compare lines 1 and 2 to line 5); both possess, as Fab fragment HC19, a light chain of the λ subtype and a heavy chain of the G subtype. All other model Fab fragments have a κ light chain. This result is probably a consequence of the lower sequence homology between mouse κ and λ light chains as compared to that between human and mouse λ light chains.

As far as translation-function calculations are concerned, we have found that, for a given model, the main variable influencing the contrast of the results is the resolution range of data used; when we analyzed the influence of that parameter, we found that inclusion of 6–4 Å resolution data greatly improves the signal-to-noise ratio in the calculation (Table 2). The 10–4 Å range has been used in all subsequent calculations.

Results of molecular replacement – variable-domain dimer

The higher variability of the V domain has made the rotation-function results far less easy to interpret than those obtained with the C domain. The V-domain rotation-function calculations were performed using the parameters we had found optimal for the C-domain dimer. Rotation functions were calculated using variable-domain dimers of all published structures (Table 5). None of these gave rise to a solution for which a contrasted translation function could be calculated, even though in some cases the height of the first peak of the rotation function was comparable to that obtained for the constant domain. *A posteriori* analysis shows that all orientations found were indeed wrong.

Using the hybrid models previously described, one orientation consistently gives one of the highest peaks of the rotation function and once the highest peak, albeit with a low contrast (Table 6). Nevertheless, a second orientation repeatedly gives the highest rotation-function peak (Tables 5 and 6). A translation function was calculated in space group $P3_121$ and with the KOL 'modified'-D1.3 variable-domain model using these two orientations; in the former

Table 5. *V-domain rotation-function results*

Integration radii, 4–24 Å; resolution range, 10–4 Å. For each model: first line, model with CDR's; second line, model without CDR's. All peak heights expressed as multiples of the r.m.s.d. of the rotation function.

Model	True peak height H1	Position in list	First spurious peak height H2
KOL	< 2·7		3·45
	3·1	4	4·3*
NEW	< 2·5		3·2
	< 2·7		2·95
HYHEL5	< 2·7		3·35
	< 2·7		3·15
HYHEL10	< 2·7		3·4*
	< 3·0		3·55*
MCPC603	< 2·7		3·35
	< 2·7		3·5
D1.3	3·75	3	4·15
	< 3·0		3·65
J539	< 2·5		3·2
	< 2·7		3·1
R19.9	< 2·7		3·05
	< 2·7		3·4*

* The first peak corresponds to a repeatedly found wrong orientation.

Table 6. *V-domain rotation-function results*

Integration radii, 4–24 Å; resolution range, 10–4 Å. Both domains without CDR's. All peak heights expressed as multiples of the r.m.s.d. of the rotation function.

Model (VL domain)	Model (VH domain)	True peak height H1	Position in list	First spurious peak height H2
KOL	KOL	3·1	Fourth	4·3*
KOL 'modified'	KOL	3·3	Third	3·9*
KOL	HYHEL5	3·3	Fourth	3·8*
KOL 'modified'	HYHEL5	3·25	Second	3·4*
KOL 'modified'	D1.3	3·85	First	3·7
KOL	HYHEL10	3·0	Fourth	3·15*
KOL 'modified'	HYHEL10	3·05	Fourth	3·15

* The first peak corresponds to a repeatedly found wrong orientation (the same as indicated in Table 5).

case, this translation function has a first peak of 11·1 r.m.s.d. height, whereas the next-highest peak is of 6·65 r.m.s.d. height; in the latter case, the first peak is of 5·05 r.m.s.d. height, whereas the next-highest peak is of 5·0 r.m.s.d. height. Thus, the translation function allows us to discriminate those two repeatedly found orientations; several tests (see below) confirm that the first orientation corresponds to that of the V domain in the structure.

Evaluation of the results of molecular replacement

Molecular replacement has provided us with tentative locations for the constant and variable domains of Fab fragment HC19, which were obtained independently one from the other (for transformations that allow positioning of the HC19 model in its unit cell, see *Appendix*). Three pieces of evidence support the validity of the resulting model.

(*a*) Conformational characteristics relating the two independently localized domains are in the range of values found for other Fab fragment structures. Two of these characteristics are the elbow angle (value found for HC19: 148°; range for other Fab fragment

structures: 115–180°) and the distance between the S atoms of residues L23 and L134 [value found for HC19: 42 Å; range for other Fab fragment structures: 38–42 Å; for the numbering of the residues see Kabat, Wu, Bilofsky, Reid-Miller & Perry (1983)].

(*b*) We calculated a difference Fourier map after molecular replacement omitting from the model S atoms of the disulfide bridges characteristic of immunoglobulin domains; this map has a maximum at 4·58 times its r.m.s.d. Disulfide bridges correspond to features of the map above the 2·9 r.m.s.d. level; two of them (Fig. 1) are among the highest peaks of the map (4·48 and 4·25 r.m.s.d.).

(*c*) The correlation coefficient between the observed structure factors and those calculated from the unrefined molecular-replacement model is 0·5 (resolution range 15–4 Å). One cycle of *X-PLOR* refinement of the model against the 3·5 Å diffraction data led to an *R* value of 17·6% (r.m.s.d. of bond lengths from their target values: 0·02 Å).

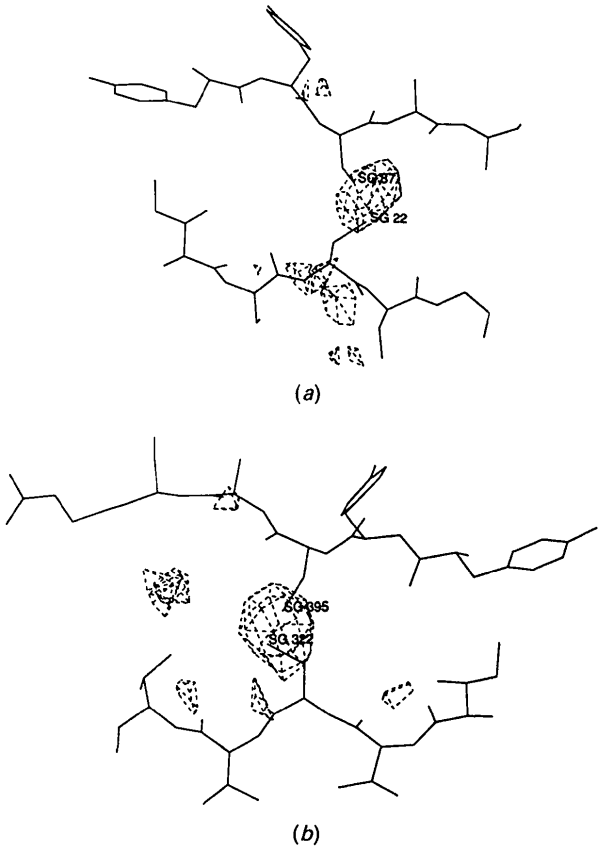

Fig. 1. (*a*) Electron density corresponding to the disulfide bridge of the VL domain, calculated at 3·5 Å resolution, with ($F_o - F_c$) as coefficients and where the F_c's were calculated with all S atoms removed from the molecular-replacement model. The contouring level is 2 r.m.s.d. above the mean of the map. (*b*) Same as above, with the disulfide bridge of the VH domain. Electron density is displayed using the program *FRODO* (Jones, 1978) on an Evans & Sutherland graphics system.

Discussion

Molecular replacement allowed us to solve the structure of a Fab fragment belonging to a class for which no structure has yet been published. Several points which arose during this work seem to be worth mentioning; the first is a confirmation of previous results on a Fab fragment (Cygler & Anderson, 1988*a*) concerning the limit in resolution of useful data in MR calculations. We found that the shell at 4 Å resolution improves the signal-to-noise ratio of the rotation function. Importance of the 4 Å data may be related, as suggested by Blundell & Tickle (1985), to the all β-sheet nature of the Fab fragment structures; it may also be related to the expected deviation between the model and the structure. When one defines the common core of two molecules as the atoms of these molecules which can be superposed with less than 3 Å deviation, the proportion of atoms in that core and the r.m.s.d.'s of their positions can be estimated from the sequence homology between the two molecules (Chothia & Lesk, 1986). In the case of the variable-domain dimer, we expect a sequence homology of 35%; the corresponding average proportion of atoms in the core is 85% and their estimated r.m.s.d. is 1·3 Å; 1·6 Å is then an expected lower estimate of the r.m.s.d. of all the atoms of the variable dimers. Data at a resolution lower than 2–3 times that value should still be sufficiently consistent between the model and the part of the structure it represents such as to enhance the overlap of their Patterson functions.

An often recurring problem in MR is the detection of a rotation-function solution corresponding to the actual orientation of the model in the unknown structure. Our results demonstrate that rotation functions do not necessarily reliably indicate the orientation of the model as the highest peak in a given rotation function or even as a consistently found high-ranking orientation in a set of rotation functions calculated under different conditions (see Tables 5 and 6). A better indicator of the right orientation is the achievement of a high contrast in the associated translation function. In our calculations, a wrong orientation never gave rise to a contrast in the translation function comparable to that obtained for the right orientation. This was true both for the V- and C-dimer searches, where one half of the content of the asymmetric unit is included in the model.

Finally, and perhaps most important, our results emphasize the influence of the choice of a model on the MR results. Building of a hybrid-model dimer for the variable part of the Fab fragment was

decisive in our MR calculations. Other approaches, involving a systematic analysis of the first peaks of the rotation function could have led to the exploration of peaks corresponding to an orientation close to the right one, at least in the case of some of the models of the variable-domain dimer (Table 5, lines 2 and 11). This would have enabled us to detect the solution; such might also have been the case had we systematically refined other models we have been using (Brünger, 1990). Our approach allowed us to avoid these computer-demanding steps.

It is a pleasure to acknowledge Dr T. N. Bhat's help with the use of the program *X-PLOR*.

APPENDIX

Coordinates of the unrefined molecular-replacement model of Fab fragment HC19

This model is obtained by applying the following rotations and translations:

(*a*) To the coordinates of Fab-fragment KOL (Protein Data Bank File Reference 2FB4)

$$x_{\text{model}} = \Omega.x_{\text{kol}} + t.$$

Light-chain variable domain (VL)

$$\Omega = \begin{matrix} -0{\cdot}60 & -0{\cdot}50 & -0{\cdot}63 \\ 0{\cdot}28 & -0{\cdot}86 & 0{\cdot}43 \\ -0{\cdot}75 & 0{\cdot}08 & 0{\cdot}65 \end{matrix} \quad t = \begin{matrix} 101{\cdot}1\ \text{Å} \\ 51{\cdot}4\ \text{Å}. \\ 21{\cdot}8\ \text{Å} \end{matrix}$$

Constant-domain dimer (CL + CH1)

$$\Omega = \begin{matrix} -0{\cdot}55 & -0{\cdot}40 & -0{\cdot}73 \\ 0{\cdot}50 & -0{\cdot}86 & 0{\cdot}09 \\ -0{\cdot}67 & -0{\cdot}32 & 0{\cdot}67 \end{matrix} \quad t = \begin{matrix} 97{\cdot}0\ \text{Å} \\ 53{\cdot}1\ \text{Å}. \\ 45{\cdot}0\ \text{Å} \end{matrix}$$

(*b*) To the coordinates of Fab-fragment D1.3 (Protein Data Bank File Reference 1FDL)

$$x_{\text{model}} = \Omega.x_{\text{d1.3}} + t.$$

Heavy-chain variable domain (VH)

$$\Omega = \begin{matrix} 0{\cdot}99 & -0{\cdot}11 & 0{\cdot}01 \\ -0{\cdot}10 & -0{\cdot}90 & 0{\cdot}41 \\ -0{\cdot}04 & -0{\cdot}41 & -0{\cdot}91 \end{matrix} \quad t = \begin{matrix} 19{\cdot}3\ \text{Å} \\ 43{\cdot}1\ \text{Å}. \\ 38{\cdot}4\ \text{Å} \end{matrix}$$

The coordinates of the model x_{model} are expressed in the orthonormal frame (**i**, **j**, **k**) with **i**//**a**, **j**//**c**$\wedge$**a**, **k**//**c**.

References

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F., BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BIZEBARD, T., MAUGUEN, Y., PETEK, F., RIGOLET, P., SKEHEL, J. J. & KNOSSOW, M. (1990). *J. Mol. Biol.* **216**, 513–514.

BLOW, D. M. (1985). *Molecular Replacement. Proceedings of the Daresbury Study Weekend*, 15–16 *February* 1985, edited by P. A. MACHIN, pp. 2–7. Warrington: SERC Daresbury Laboratory.

BLUNDELL, T. L. & TICKLE, I. J. (1985). *Molecular Replacement. Proceedings of the Daresbury Study Weekend*, 15–16 *February* 1985, edited by P. A. MACHIN, pp. 89–91. Warrington: SERC Daresbury Laboratory.

BRÜNGER, A. T. (1990). *Acta Cryst.* A**46**, 46–57.

CHOTHIA, C. & LESK, A. M. (1986). *EMBO J.* **5**, 823–826.

CHOTHIA, C., LESK, A. M., TRAMONTANO, A., LEVITT, M., SMITH-GILL, S. J., AIR, G., SHERIFF, S., PADLAN, E. A., DAVIES, D., TULIP, W. R., COLMAN, P. M., SPINELLI, S., ALZARI, P. M. & POLJAK, R. J. (1989). *Nature* (*London*), **342**, 877–883.

CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173–178. New York: Gordon & Breach.

CROWTHER, R. A. & BLOW, D. M. (1967). *Acta Cryst.* A**23**, 544–548.

CYGLER, M. & ANDERSON, W. F. (1988*a*). *Acta Cryst.* A**44**, 38–45.

CYGLER, M. & ANDERSON, W. F. (1988*b*). *Acta Cryst.* A**44**, 300–308.

DANIELS, R. S., DOUGLAS, A. R., SKEHEL, J. J. & WILEY, D. C. (1983). *J. Gen. Virol.* **64**, 1657–1662.

FRIGUET, B., DJAVADI-OHANIANCE, L. & GOLDBERG, M. E. (1989). *Protein Structure – A Practical Approach*, edited by T. E. CREIGHTON, pp. 287–310. Oxford: IRL Press.

FUJINAGA, M. & READ, R. J. (1987). *J. Appl. Cryst.* **20**, 517–521.

HIRSHFELD, F. L. (1968). *Acta Cryst.* A**24**, 301–311.

HUBER, R. & BENNETT, W. S. (1987). *Nature* (*London*), **326**, 334–335.

JONES, A. T. (1978). *J. Appl. Cryst.* **11**, 614–617.

KABAT, E. A., WU, T. T., BILOFSKY, H., REID-MILLER, M. & PERRY, H. (1983). *Sequences of Proteins of Immunological Interest*. US Department of Health and Human Services, Public Health Service, National Institute of Health.

LASCOMBE, M. B. (1989). PhD Thesis. Université Paris VII, France.

LASCOMBE, M. B., ALZARI, P. M., BOULOT, G., SALUDJIAN, P., TOUGARD, P., BEREK, C., HABA, S., ROSEN, E. M., NISONOFF, A. & POLJAK, R. J. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 607–611.

MARQUART, M., DEISENHOFER, J., HUBER, R. & PALM, W. (1980). *J. Mol. Biol.* **141**, 369–391.

MOSS, D. S. (1985). *Acta Cryst.* A**41**, 470–475.

PADLAN, E. A., SILVERTON, E. W., SHERIFF, S., COHEN, G. H., SMITH-GILL, S. J. & DAVIES, D. R. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 5938–5942.

POLJAK, R. J., AMZEL, L. M., AVEY, H. P., CHEN, B. L., PHIZACKERLEY, R. P. & SAUL, F. A. (1973). *Proc. Natl Acad. Sci. USA*, **70**, 3305–3310.

ROSSMANN, M. G. (1990). *Acta Cryst.* A**46**, 73–82.

ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24–31.

SAUL, F. A., AMZEL, L. M. & POLJAK, R. J. (1978). *J. Biol. Chem.* **253**, 585–597.

SEGAL, D. M., PADLAN, E. A., COHEN, G. H., RUDIKOFF, S., POTTER, M. & DAVIES, D. R. (1974). *Proc. Natl Acad. Sci. USA*, **71**, 4298–4302.

SHERIFF, S., SILVERTON, E. W., PADLAN, E. A., COHEN, G. H., SMITH-GILL, S. J., FINZEL, B. C. & DAVIES, D. R. (1987). *Proc. Natl Acad. Sci. USA*, **84**, 8075–8079.

SUH, S. W., BHAT, T. N., NAIRA, M. A., COHEN, G. H., RUDIKOFF, S., RAO, D. N. & DAVIES, D. R. (1986). *Proteins Struct. Funct. Genet.* **1**, 74–80.

TICKLE, I. J. (1985). *Molecular Replacement. Proceedings of the Daresbury Study Weekend*, 15–16 *February* 1985, edited by P. A. MACHIN, pp. 22–26. Warrington: SERC Daresbury Laboratory.

WILEY, D. C. & SKEHEL, J. J. (1987). *Annu. Rev. Biochem.* **56**, 365–394.