Acta Cryst. (1996). D**52**, 315–326

Refined Structure of the Monoclonal Antibody HyHEL-5 with its Antigen Hen Egg-White Lysozyme

GERSON H. COHEN, STEVEN SHERIFF† AND DAVID R. DAVIES

Laboratory of Molecular Biology, National Institute of Diabetes, Digestive and Kidney Diseases, Bethesda, MD 20892-0560, USA

(Received 30 May 1995; accepted 27 October 1995)

Abstract

The crystal structure of the complex of the antibody Fab, HyHEL-5, with its antigen, hen egg-white lysozyme, has been refined at 2.65 Å resolution to an R value of 0.196. The resulting model has significantly better stereochemistry than the previously reported model of the complex, PDB reference 2HFL, and sufficiently improved phases, permitting the reliable location of a number of water molecules. No major conformational differences are observed between this structure and that previously reported, although small differences occur throughout the complex. 82 water molecules have been assigned, of which three are in the antibody–antigen interface involved in a hydrogen-bonding network. Three other waters are trapped within the interface between V_H and V_L and a fourth water molecule is observed near the interface but buried below the lysozyme surface as observed in crystal structures of lysozyme alone.

1. Introduction

The mechanisms of antibody specificity and affinity for proteins have been studied by a number of techniques including epitope mapping with evolutionary variants or site-directed mutants (Smith-Gill *et al.*, 1982; Grivel & Smith-Gill, 1996), hydrogen-exchange NMR spectroscopy (Paterson, Englander & Roder, 1990), calorimetry (Hibbits, Gill & Willson, 1994; Bhat *et al.*, 1994), and X-ray crystallography. The crystal structures of antibody complexes with lysozyme and a number of other protein antigens have been reported (Amit, Mariuzza, Phillips, & Poljak, 1986; Sheriff, Silverton *et al.*, 1987; Padlan *et al.*, 1989; Bhat, Bentley, Fischmann, Boulot & Poljak, 1990; Fischmann *et al.*, 1991; Chitarra *et al.*, 1993; Braden *et al.*, 1994; Chacko *et al.*, 1995; Bentley, Boulot, Riottot, & Poljak, 1990; Tulip, Varghese, Laver, Webster & Colman, 1992; Tulip, Varghese, Webster, Laver & Colman, 1992; Prasad *et al.*, 1993; Ban *et al.*, 1994; Evans, Rose, To, Young & Bundle, 1994; Malby *et al.*, 1994; Bizebard *et al.*, 1995; Bossart-Whitaker, Chang, Novotny, Benjamin & Sheriff, 1995; Fields, Goldbaum, Ysern, Poljak & Mariuzza, 1995). These studies have shown that from 13 to 20 amino acids from both antibody and antigen contact each other and a somewhat larger number are buried by the interaction. They have also shown that 10–15 hydrogen bonds are formed between antibody and antigen and that an area of 600–900 Å^2 is buried on each surface. The contact surfaces are generally complementary to one another with only a very few solvent molecules in the interface between antibody and antigen, although a large number may be located at the periphery of the interface (Bhat *et al.*, 1994).

Antibodies to lysozyme were developed because it was a structurally defined antigen (Ramanadham, Sieker & Jensen, 1990; Kurinov & Harrison, 1995; Young, Dewan, Nave & Tilton, 1993; Wilson, Malcolm & Matthews, 1992; Harata, 1994) and because a large number of avian lysozymes were readily available for epitope mapping (Smith-Gill *et al.*, 1982). The structures of five anti-lysozyme antibodies in complex with lysozyme have been reported. Three of these antibodies, HyHEL-5 (Sheriff, Silverton *et al.*, 1987), HyHEL-10 (Padlan *et al.*, 1989) and D1.3 (Amit *et al.*, 1986; Fischmann *et al.*, 1991; Bhat *et al.*, 1990, 1994) have essentially non-overlapping epitopes; D44.1 (Braden *et al.*, 1994) involves an almost identical epitope with that of the HyHEL-5–lysozyme complex; and D11.15 (Chitarra *et al.*, 1993) has extensive overlap with D1.3. For one of these antibodies, HyHEL-5, a structure in complex with a site directed lysozyme mutant, Arg68 → Lys, has been studied (Chacko *et al.*, 1995). D11.15 has been studied in complex with several different avian lysozymes that differ by two amino acid residues in the epitope region. Reported here is the completion of the refinement of the HyHEL-5 complex with native lysozyme to serve as a standard for comparison with the site-directed mutant and to provide a more accurate structure for the study of antibody–protein interactions.

2. Methods

The HyHEL-5–lysozyme complex crystals have been reported in two variations, a 'long axis' form with cell constants a = 54.79, b = 74.82, c = 78.95 Å, β = 101.82°,

† Current address: Bristol-Myers Squibb Pharmaceutical Research Institute, PO Box 4000, Princeton, NJ 08543-4000, USA.

and a 'short axis' form with cell constants a = 54.9, b = 65.2, c = 78.6 Å, β = 102.4°, both in space group $P2_1$ (Sheriff, Silverton *et al.*, 1987). This refinement centered mainly on the long axis form because of its better resolution (2.54 Å); the available data for the other form, which was only observed in one crystal, was cut off at 3.0 Å. The diffraction data used here were the same as those from the earlier work (Sheriff, Silverton *et al.*, 1987), subject to the following adjustment in resolution. Analysis of the distribution of the observed data for the long axis form showed that only 10.6% of the highest resolution data ($2.65 > d > 2.54$ Å) were observed. Consequently, the range of the reflection data used in this refinement was limited to $10.0 > d > 2.65$ Å. Table 1 shows the percentage of data observed as a function of the resolution for the resolution range used.

The starting point for refinement was the coordinate set 2HFL from the Protein Data Bank (Bernstein *et al.*, 1977). These coordinates were further refined with *PROTIN/PROLSQ* (Hendrickson, 1985) to an R value of 0.199. During this refinement some minor adjustments were made to rebuild parts of the complex, the complete side chains for chains for $Gln3_H$, $Gln5_H$ and $Tyr27_H$ were added, two residues were added to the end of the heavy chain ($Val224_H$ and $Pro225_H$), and the lysozyme sequence was corrected so that Asp103 was replaced by Asn.*

Refinement was continued with the program *X-PLOR* (Brünger, Kuriyan & Karplus, 1987), versions 2.1 and 3.0. For most of the refinement we used the *CHARMM* style (Brooks *et al.*, 1983) parameter file PARAM19X.PRO. For the final passes through the refinement we used the parameter file PARHCSDX.PRO, based on the work of Engh & Huber (1991), as suggested by an analysis of bond angles and lengths in protein molecules (Laskowski, Moss & Thornton, 1993).

The 'slow-cooling' procedure of *X-PLOR* (Brünger, 1992) was used for the first six passes of refinement. The initial temperature for the first slow cooling was 4000 K. In order not to destroy the results of the graphics rebuilding, the starting temperature was reduced in subsequent passes; the second and third passes were started at 3000 K, the fourth at 2500 K, the fifth at 2000 K and the sixth at 1500 K. In all cases the final temperature was 300 K. The B parameters for all of the atoms of the starting model were set to 15.0 $Å^2$ and kept fixed during this phase of the refinement. After each pass, the model was examined with molecular graphics and rebuilt as necessary. The regions of the molecule that required the most attention were detected by the *GEOM* program (Cohen, 1993).

Table 1. *Refinement statistics for the HyHEL-5–lysozyme complex*

R value and completeness of data by shells

	Working set		Test set		Total
Range	No. of data	R value	No. of data	R free*	% obs
4.42–10.0	3063	0.187	349	0.267	94
3.56–4.42	2759	0.168	313	0.277	86
3.13–3.56	2487	0.204	278	0.290	77
2.85–3.13	2079	0.243	249	0.350	66
2.65–2.85	1366	0.262	155	0.335	43
Total range	11754	0.196	1344	0.288	73

Number of non-H atoms	4333
Protein	4251
Solvent	82
R.m.s. deviations from ideality†	
Bond lengths (Å)	0.009
Bond angles (°)	1.9

B parameter statistics

	Average	Minimum	Maximum
Protein atoms	17.1	4.0	37.6
Solvent atoms	21.3	2.1	36.8
All atoms	17.1	2.1	37.6

* Brünger (1992). † The r.m.s. deviations have been calculated against the parameter set of Engh & Huber (1991).

Following the six passes, the refinement procedure was limited to a Jack–Levitt approach (Jack & Levitt, 1978) carried out in *X-PLOR* as a conjugate-gradient energy minimization with no heating or cooling. A *PROLSQ* weighting scheme (Hendrickson, 1985) was applied to the diffraction data and the relative weight between the diffraction data and the energy parameters (Brünger, 1992) was adjusted during refinement to keep the stereochemistry near standard geometry.

Water molecules were assigned by automatic examination of peaks in ΔF maps for each cycle using the program *ADDH2O* (Sheriff, unpublished work), that evaluated the location of the peak in the context of the current protein model. Peaks that satisfied reasonable hydrogen-bond criteria were then examined with the graphics. Near the completion of the refinement a sequence correction was discovered upon the analysis of the cloned HyHEL-5 light-chain gene (Newman & Smith-Gill, unpublished work). Examination of the electron-density maps showed that the density was consistent with the change in assignment of $Glu85_L$ to Thr. This correction was incorporated in the remaining stages of the refinement.

The last refinement pass through *X-PLOR* was used to calculate R_{free} (Brünger, 1992). In order to break any hysteresis in the 10% of the data omitted from the refinement, a brief slow-cooling protocol starting from 500 K was carried out followed by Jack–Levitt refinement, overall B-parameter refinement and individual

* The HyHEL-5–lysozyme complex consists of three polypeptide chains, each with its own numbering system. The residues of the Fab are presented in this paper according to the numbering system of Kabat, Wu, Reid-Miller, Perry & Gottesman (1987). The subscript H or L following a residue number indicates whether the residue is a member of the heavy or light chain of the antibody. Antibody domains are designated by V or C for variable or constant followed by the appropriate subscript for heavy or light. The residues in lysozyme are numbered sequentially and indicated by a Y. Solvent molecules are all identified as Wat (water) and numbered sequentially starting from 131.

isotropic B-parameter refinement. Table 1 summarizes R_{free} by resolution range.

Following the refinement of the long axis crystal form, an attempt was made to refine the short axis form. The variable pair of domains and the constant pair were independently superimposed on an earlier tentative model for the short axis form. After conservative idealization and adjustment of an isotropic overall temperature factor, the new model was subjected to slow cooling from 1000 K. While the resulting model appeared reasonable and seemed to fit its density reasonably, the associated maps were not sufficiently detailed to allow the critical adjustment of side chains. It is probable that the gross conformation of this molecule is reliable but a full set of higher resolution data are needed to confirm the detailed structure.

Molecular comparisons were carried out using *ALIGN* (Cohen in, Satow, Cohen, Padlan & Davies, 1986). Map calculation was carried out using *GHC650* (Cohen, unpublished work) and molecular modeling was performed with either *FRODO* (Jones, 1978), on an Evans & Sutherland PS390, or *O* (Jones, Zou, Cowan & Kjeldgaard, 1991), on a Silicon Graphics Iris workstation or on a Stellar GS2500. Surface calculations and the study of cavities were carried out using the *MS* program suite (Connolly, 1983) and the *MOLECULAR SURFACE PACKAGE* (Connolly, 1993). Cross interface interactions were analyzed using the program *CONTACSYM* (Sheriff, Hendrickson & Smith, 1987). Molecular diagrams included in this paper were produced using *ORTEP* (Johnson, 1965).

2.1. *Model quality*

Refinement statistics are shown in Table 1. The final R value for the model including 82 water molecules is 0.196 with tight geometry. The R_{free} is 0.288. A coordinate error of 0.30 Å is estimated for the model by the method of Luzzati (1952) as shown in Fig. 1. The stereochemistry of the final model has been improved over the published coordinate set, 2HFL (Sheriff, Silverton *et al.*, 1987). By analysis with *PROCHECK* (Laskowski, MacArthur, Moss & Thornton, 1993), the current model shows a 1.5σ deviation from the ideal values for both the ω angle and the pseudo angle ζ, used to estimate the C^{α} chirality. For the 2HFL model the corresponding deviations are 1.8σ for ω and 2.7σ for ζ. The final model presented here has six bad contacts while the 2HFL model has 50 bad contacts. Finally, the φ,ψ plot (Sasisekharan, 1962; Ramachandran & Sasisekharan, 1968) for the current model (Fig. 2) has only two residues in disallowed regions and seven residues in generously allowed regions while the model 2HFL has six and eleven.

The angles for most of the residues of the refined model are well placed in the φ,ψ plot (Fig. 2). Two residues fall in the disallowed region: $Thr51_L$ and $Ser168_L$. $Ser168_L$ is reasonably placed in its density, although the positioning is not definitive. Residue $Thr51_L$ is well placed in density with a clear choice for its side chain and the preceding and following carbonyl O atoms. As a consequence of the φ and ψ angles at $Thr51_L$, the conformation of L2 is a class 3 γ-turn (Milner-White & Poet, 1987). This conformation

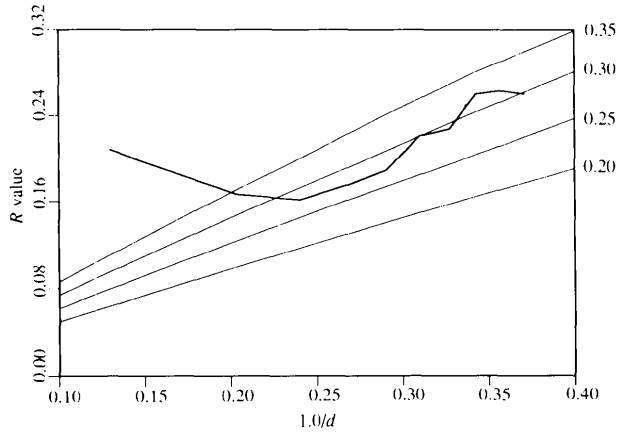

Fig. 1. Plot of observed R value as a function of $1/d$ (heavy line) for the 11 754 reflections included in the working set (Brünger, 1992). The working set of diffraction data has been divided into ten equal-volume bins. Also shown are the theoretical curves (light lines) for R value *versus* resolution (Luzzati, 1952) for assumed coordinate errors of 0.20 through 0.35 Å. The estimated error for this work is 0.30 Å.

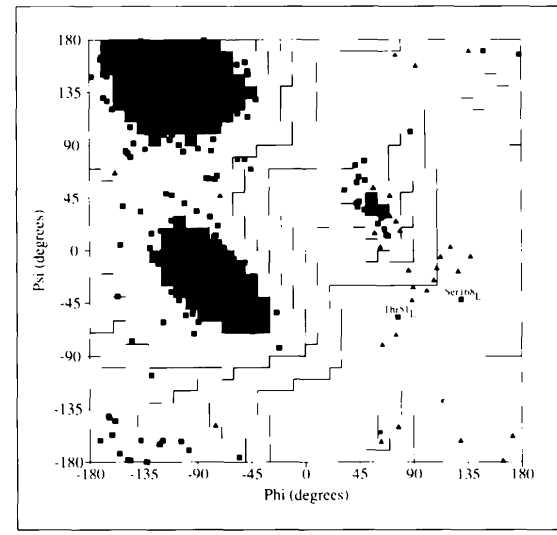

Fig. 2. Plot of φ,ψ, (Sasisekharan, 1962; Ramachandran & Sasisekharan, 1968) for the full complex as generated by the *PROCHECK* package (Laskowski *et al.*, 1993). The shaded regions are, in the order of darkest to lightest: most favored, additional allowed, generously allowed, and disallowed. These delineations were determined by an analysis (Laskowski *et al.*, 1993) of high-resolution quality structures found in the Protein Data Bank (Bernstein *et al.*, 1977). The two residues in the 'disallowed' region are labeled in the figure and discussed in the text. The seven residues in the 'generously allowed' regions are discussed in the text.

is also found in L2 of the D44.1 complex (Braden *et al.*, 1994), the NC41 Fab–neuraminidase complex (Tulip, Varghese, Laver *et al.*, 1992), and the NC10 Fab–neuraminidase complex (Malby *et al.*, 1994), as well as in L2 of other Fab light chains (Milner-White, Ross, Ismail, Belhaj-Mostefa & Poet, 1988).

Seven other residues fall in the 'generously' allowed regions of the plot. Five of these residues, $Glu154_L$, $Arg211_L$, $Ala129_H$, $Gln133_H$ and $Gln179_H$, are in broken or ambiguous density and cannot be fitted properly. The remaining two, $Ser56_L$ and $Asp170_L$ fall in reasonable to good density and appear well placed. Attempts to alter these two residues to remove them to more canonical regions of the φ,ψ plot also remove them from proper density and destroy the fit of neighboring residues in their density.

2.2. *Cis peptides*

The final, refined model of the complex has five *cis*-proline residues: $Pro8_L$, $Pro141_L$, $Pro149_H$, $Pro151_H$ and $Pro202_H$. The first four prolines in this list were identified in the model 2HFL. The assignment of $Pro202_H$ as a *cis* proline is clear in the current electron-density maps. This assignment was independently confirmed for the HyHEL-5 complex with the lysozyme mutant Arg68 → Lys (Chacko *et al.*, 1995; PDB reference 2IFF).

The refined coordinates for the long axis crystal form of the HyHEL-5–lysozyme complex have been deposited with the Brookhaven Protein Data Bank.*

3. Results

As a result of the refinement, the geometry of the complex is more canonical than the first model of the complex, 2HFL. In addition, the phases were sufficiently improved that we are able to locate 82 water molecules with some certainty. We have, therefore, a more reliable description of the complex and, in particular, the antibody–antigen interface.

3.1. *The Fab–lysozyme interface*

Although generally similar, the Fab–lysozyme interface differs in detail from the model that has been described before (Sheriff, Silverton *et al.*, 1987; Davies, Sheriff & Padlan, 1988, 1989; Davies, Padlan & Sheriff, 1990). Fig. 3 enumerates the cross-interface contacts between the antibody combining site and the epitope residues. The differences in the assignments of the contacting residues of this work and the earlier model are small and not significant considering the resolution of these studies and the uncertainty in the coordinates suggested by their Luzzati plots: 0.3 Å here and 0.4 Å earlier. All six CDR's of the antibody (Wu & Kabat, 1970) are involved in the interface. Contacting residues, Fig. 3, include two residues from L1, one residue from L2, three residues from L3, one residue from H1, seven residues from H2, and three residues from H3. One framework residue, $Trp47_H$, is also in contact with the lysozyme. The epitope on lysozyme, Fig. 3, presents 13 contacting residues from two strings of sequence plus one single residue.

$Arg68_Y$ also makes two sets of cross-interface hydrogen bonds through water molecules in the interface. Through its carbonyl O atom and the buried water molecule Wat138, $Arg68_Y$ makes hydrogen bonds to $Tyr34_L$ O^{η}, $Asp50_L$ $O^{\delta 2}$ and $Tyr97_H$ O^{η} (Table 2). Through its $N^{\eta 1}$ and Wat154 (2.9 Å), $Arg68_Y$ makes hydrogen bonds to $Glu35_H$ $O^{\varepsilon 1}$ (2.8 Å), $Gly95_H$ O (3.2 Å) and Wat210 (2.6 Å). Finally, Wat210 makes hydrogen bonds to three atoms on the interface surface of V_H: $Trp33_H$ O (3.0 Å), $Gly95_H$ N (2.8 Å) and $Asn96_H$ $N^{\delta 2}$ (3.1 Å). As a result of the hydrogen bonds through waters, one additional residue from each of L1 ($Tyr34_L$) and H1 ($Glu35_H$) may be included as part of the interface although it is not in direct contact with the antigen.

Lysozyme:	Gln 41	Thr 43	Asn 44	Arg 45	Asn 46	Asp 48	Gly 49	Tyr 53	Gly 67	Arg 68	Thr 69	Pro 70	Leu 84
Fab V_L:													
Asn 31	.	.	.	.	.	2	.	.	.	.	.	.	.
Tyr 32	.	.	.	.	.	.	.	.	.	.	.	9	.
Asp 50	.	.	.	.	.	.	.	.	.	.	.	2	.
Trp 91	.	.	.	7	.	.	2	.	.	6	.	.	.
Gly 92	.	.	.	3	1	.	.	.	.	.	.	.	.
Arg 93	.	.	3	14	7	.	.	.	.	.	.	.	.
Fab V_H:													
Trp 33	.	4	.	.	.	.	.	4	.	9	.	.	.
Trp 47	.	.	.	1	.	.	.	.	.	.	.	.	.
Glu 50	.	.	.	4	.	.	.	.	.	2	.	.	.
Leu 52	.	.	.	.	.	.	.	.	.	.	.	.	1
Ser 54	1	.	.	.	.	.	.	.	.	.	.	.	4
Gly 55	2	.	.	.	.	.	.	.	.	.	.	.	.
Ser 56	5	3	.	.	.	.	.	.	.	.	.	.	1
Thr 57	.	1	.	.	.	.	.	.	.	.	.	.	.
Asn 58	.	9	.	1	.	.	.	.	.	.	.	.	.
Gly 95	.	.	.	.	.	.	.	.	.	1	.	.	.
Asn 96	.	.	.	.	.	.	.	.	2	2	.	.	.
Tyr 97	.	.	.	.	.	.	.	.	5	4	4	4	.

Fig. 3. Number of contacts between residdues of lysozyme and Fab across the interface. The pairs listed here include hydrogen bonds, salt bridges and van der Waals interactions without distinction. van der Waals interactions are defined as being within 1.11 of the sum of the individual van der Waals radii (Gelin & Karplus, 1979) of the two atoms (Sheriff, Hendrickson & Smith, 1987).

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 3HFL, R3HFLSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: GR0455).

A re-examination of model 2HFL using extended van der Waals radii (Gelin & Karplus, 1979) shows that residues $Leu52_H$, $Asn96_H$ and $Thr51_Y$ are contacting residues in that model, although they were not identified as such by Sheriff, Silverton *et al.* (1987). Our current model differs from the model 2HFL in that residue $Gly55_H$ is now contacting while residues $Pro95_L$, $Glu35_H$, $Thr47_Y$ and $Thr51_Y$ are no longer contacting. Each of these five residues was or is involved in no more than three contacting atom pairs and the changes must be ascribed to the coordinate uncertainty suggested above. Residue $Glu35_H$, that was previously considered contacting, is now involved in the interface *via* hydrogen bonds through water molecules.

Several other residues not in direct contact across the interface are buried because of the interactions of the residues listed above. Using a probe of radius 1.7 Å, all of the contacting residues as well as eight additional residues from V_L, six from V_H, and ten from the lysozyme are partially or fully buried in the interface (Table 3). The buried surface area on the bound lysozyme comes to ~745 Å^2 and includes 23 residues. On the antibody, the light chain has ~350 Å^2 buried and the heavy chain buries ~415 Å^2 for a total of ~765 Å^2. This buried surface includes residues from all six CDR's as well as four framework residues for a total of 32 residues.

Table 2. *Hydrogen bonds and salt bridges in the Fab–lysozyme interface*

Lysozyme	Fab V_L	Distance (Å)
Asn44 $O^{\delta1}$	Arg93 $N^{\eta2}$	3.1
Arg45 $N^{\eta2}$	Trp91 O	3.0
Arg45 $N^{\eta2}$	Gly92 O	3.1
Arg45 N^{ε}	Gly92 O	3.1
Arg45N^{ε}	Arg93 O	3.2
Arg45 O	Arg93 N^{ε}	2.5
Arg45 O	Arg93 $N^{\eta2}$	3.2
Asn46 $O^{\delta1}$	Arg93 $N^{\eta2}$	3.0
Lysozyme	**Fab V_H**	**Distance (Å)**
Gln41 $N^{\varepsilon2}$	Gly55 O	3.2
Gln41 O	Ser56 O^{γ}	2.6
Thr43 $O^{\gamma1}$	Thr57 O	3.2
Thr43 $O^{\gamma1}$	Asn58 $N^{\delta2}$	2.6
Thr43 O	Asn58 $N^{\delta2}$	2.8
Arg45 $N^{\eta1}$	Glu50 $O^{\varepsilon2}$	3.0
Tyr53 O^{η}	Trp33 $N^{\varepsilon1}$	3.1
Gly67 O	Tyr97 N	3.0
Arg68 $N^{\eta1}$	Glu50 $O^{\varepsilon1}$	3.3
Arg68 $N^{\eta2}$	Glu50 $O^{\varepsilon2}$	3.7

Four of the residues listed above as buried are not buried in the 2HFL model: $Asp1_L$, $Ile2_L$, $Lys53_L$ and $Tyr59_H$. Two residues of 2HFL that are buried are not considered buried here: $Arg46_L$ and $Ser85_Y$. In all these cases, only a small fraction of the potential surface of the

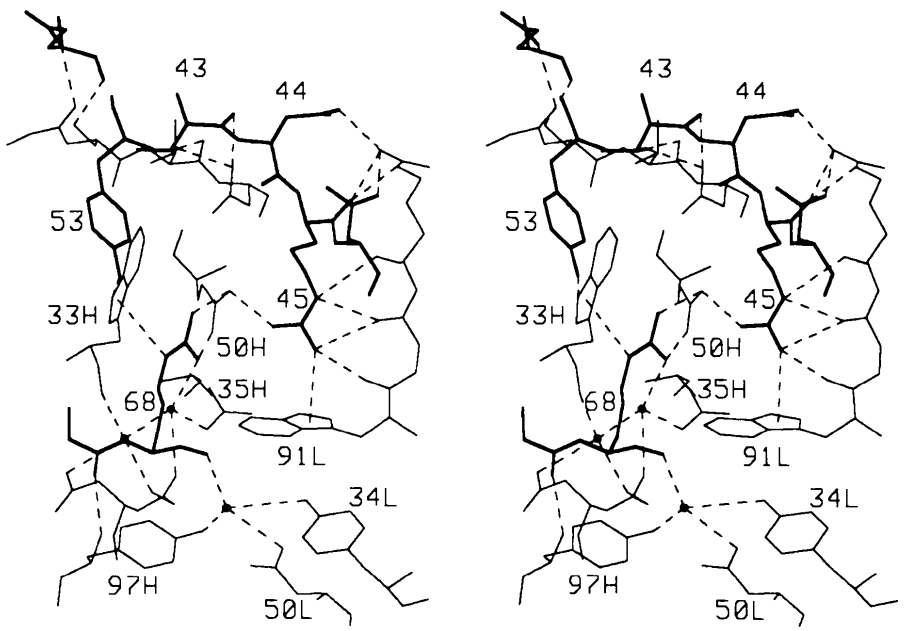

Fig. 4. *ORTEP* (Johnson, 1965) stereoview showing residues of the Fab–lysozyme interface that participate in hydrogen bonding and/or salt bridges. The buried water molecule, Wat138, is included to show the hydrogen bonds from $Arg68_Y$ through it to $Tyr34_L$, $Glu50_L$ and $Tyr97_H$. Water molecules Wat154 and Wat210 are included to show the hydrogen bonds from $Arg68_Y$ through it to $Glu35_H$, $Gly95_H$ and Wat210 and from Wat210 to $Trp33_H$, $Gly95_H$ and $Asn96_H$. Thin lines show residues from the V_L and V_H domains; heavy lines show residues from the lysozyme. In the figure, the bound lysozyme is closest to the reader and V_L is mostly on the right. The dashed lines show the interactions enumerated in Table 3, the Arg/Tyr interactions discussed in the text and the hydrogen bonds with Wat138, Wat154 and Wat210.

Table 3. *Additional residues that are partially or fully buried in the Fab–lysozyme interface*

V_L	V_H	Lysozyme
$Asp1_L$	$Ser30_H$	$Thr47_Y$
$Ile2_L$	$Asp31_H$	$Thr51_Y$
$Ser29_L$	$Tyr32_H$	$Asp52_Y$
$Val30_L$	$Glu35_H$	$Arg61_Y$
$Tyr34_L$	$Tyr59_H$	$Asn65_Y$
$Lys53_L$	$Asp98_H$	$Asp66_Y$
$Gln90_L$		$Gly71_Y$
$Pro95_L$		$Ser72_Y$
		$Pro79_Y$
		$Ser81_Y$

residue was or is buried. These differences must also be ascribed to the coordinate uncertainty suggested above.

Eight of the residues in the lysozyme epitope form hydrogen bonds or salt bridges with the Fab (Fig. 4, Table 2). These residues are involved in at least 18 such interactions with three residues of CDR3 of the V_L domain of the Fab and seven residues distributed in all three CDR's of the V_H domain. Residues $Arg45_Y$ and $Arg93_L$ each make several hydrogen bonds and salt bridges.

Residues $Arg68_Y$ and $Arg45_Y$ interact with the side chains of two tryptophan residues of the Fab in addition to the salt bridges with $Glu50_H$ (Fig. 5). $Arg68_Y$ is stacked approximately parallel to the side chain of $Trp33_H$ (Fig. 5); the plane of the four atoms at the end of the $Arg68_Y$ side chain makes a 24° angle with the plane of the $Trp33_H$ ring system. The two planes are approximately 3.5 Å from each other. Atom N^{ε} of $Arg68_Y$ is 3.9 Å from the center of the five-membered ring of $Trp33_H$ along a line 37° off the perpendicular to the ring. An analysis of the interactions of arginine side-chains with aromatic side chains (Flocco & Mowbray, 1994) supports the importance of the interaction found here between $Arg68_Y$ and $Trp33_H$. Braden *et al.* (1994) also note this stacking interaction in the D44.1–lysozyme complex.

Atom $N^{\eta 2}$ of $Arg45_Y$ is 3.4 Å from the center of the five-membered ring of $Trp91_L$ along a line 2.0° off the ring perpendicular (Fig. 5); in projection, the $N^{\eta 2}$ falls 0.12 Å from the center point of the ring. The C^{ζ} to $N^{\eta 2}$ bond makes an angle of 53° with the perpendicular to the ring system. This is in reasonable agreement with theoretical calculations (Levitt & Perutz, 1988) that show that the optimum energy benefit occurs at an N to ring separation of 3.40 Å. While the angle from perpendicular is somewhat removed from the optimum orientation of 0°, the theoretical calculations indicate that there is still a significant attractive energy at 53°.

The only peripheral water molecules involved in cross-interface interactions are Wat139 and Wat196. The surface residues $Asn44_Y$ and $Arg45_Y$ interact with $Arg93_L$ and $Tyr59_H$ through this two-water-molecule chain. Wat139 is 2.9 Å from $Asn44_Y$ $N^{\delta 2}$, 3.1 Å from $Arg45_Y$ N, and 3.2 Å from Wat196. Wat196 is, in turn, 3.2 Å from $Arg93_L$ O and 2.7 Å from $Tyr59_H$ O.

3.2. *Buried water molecules and cavities*

Two water molecules, Wat138 and Wat154, are located in the antibody–lysozyme interface making contact with the Fab and $Arg68_Y$, and a third, Wat210, is a member of a hydrogen-bonding chain directed to V_H in the interface, as noted above. Two of these, Wat154 and Wat210, are located approximately 5 Å from the nearest surface of the complex at the end of a pore that extends into the interface. The other molecule, Wat138, is at the top of an extensive cavity extending about 10 Å from the vicinity of $Arg68_Y$ into the interface between the V_L and V_H domains of the antibody (Fig. 6). This cavity is approximately 250 $Å^3$ in volume and contains another three water molecules, Wat155, Wat160 and Wat193, that are strung out in a chain (Fig. 7). They are separated by 3.0 and 2.9 Å, respectively, situated about 8 Å from the antibody–lysozyme interface, and make hydrogen bonds with residues in V_L and V_H (Table 4). The cavity is lined by 19 residues from the three components of the complex (Fig. 7; Table 5). The top of the cavity is almost cut off from the lower part by a narrow passage only slightly larger than the diameter of a water molecule. No waters have been located in this region of the cavity perhaps because two of the residues defining the walls of the cavity at this neck are aromatic: $Tyr34_L$ and $Trp91_L$. The complex with the mutant lysozyme R68K (Chacko *et al.*, 1995; PDB reference 2IFF) shows the same cavity with essentially the same dimensions. The contents of the cavity are different, however, including only two

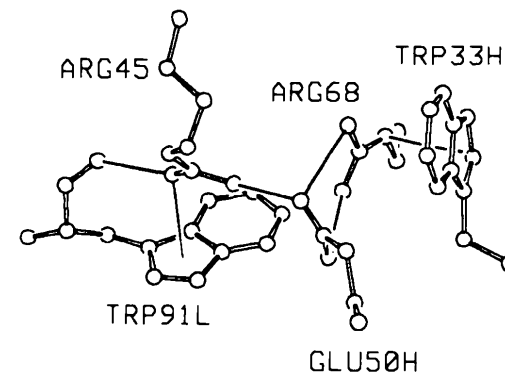

Fig. 5. *ORTEP* (Johnson, 1965) view of the three salt bridges made between $Gln50_H$ and the two arginine residues from the bound lysozyme: $Arg45_Y$ and $Arg68_Y$. Also indicated are the interactions of the arginine residues with $Trp91_L$ and $Trp33_H$. The interaction of $Arg45_Y$ and $Trp91_L$ shows the 3.37 Å distance from the center of the five-membered ring of $Trp91_L$ to the $N^{\eta 2}$ atom of $Arg45_Y$. The end of the arginine side chain makes an angle of 63° with the plane of the tryptophan ring system and the angle between the C^{ζ} to $N^{\eta 2}$ vector with the perpendicular to the plane is 53°. The angle between the planes of the side chains of $Arg68_Y$ and $Trp33_H$ is 24° and the plane separation is ~ 3.5 Å.

water molecules. Water molecule HOH748 in the R68K structure is 0.7 Å from Wat160 and HOH756 is about 3 Å above HOH748.

In the D44.1–lysozyme complex, that has some resemblance to the HyHEL-5–lysozyme complex in the V_H domain and in the epitope on the lysozyme, Braden *et al.* (1994) note two cavities in the interface between V_H and the antigen. No water molecules could be located in these cavities which the authors ascribe to the fact that they are lined by hydrophobic residues. Based on the enumeration of the V_H residues by which they are defined these cavities do not appear to correspond to the cavity described above for the HyHEL-5–lysozyme complex. In the Fv D1.3–lysozyme complex (PDB reference 1VFB) a water containing cavity is found that is approximately one-half the volume of

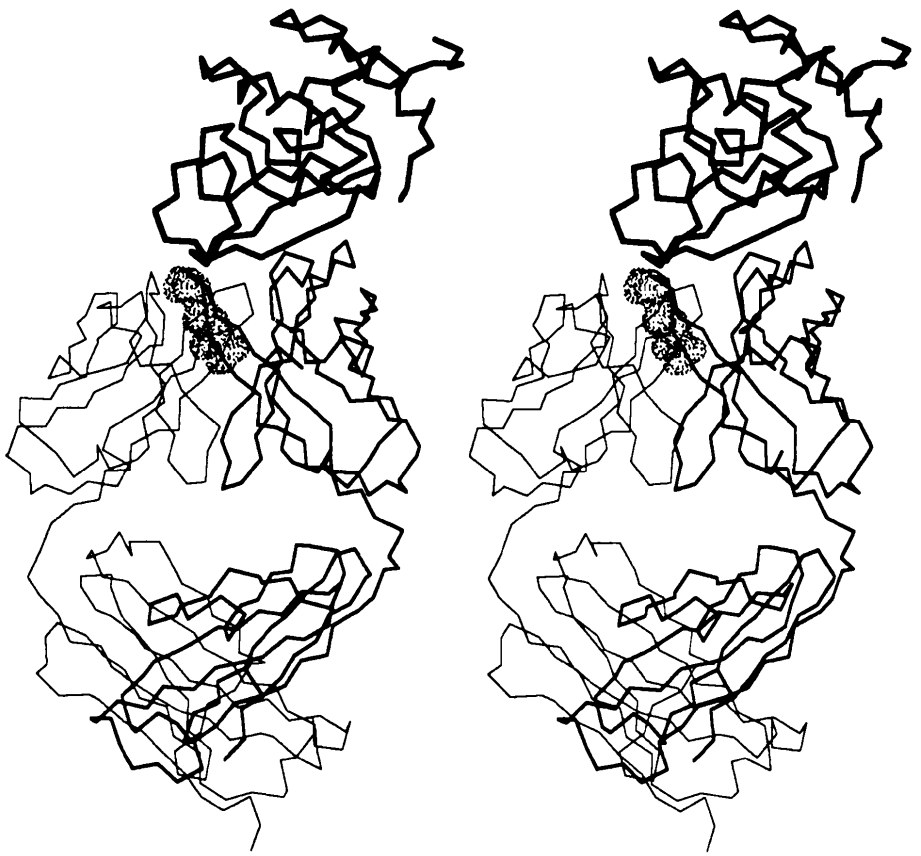

Fig. 6. *ORTEP* (Johnson, 1965) stereoview of the C^{α} trace of the complex showing the extensive cavity located between the V_L and V_H domains that extends just up to the lysozyme. The light chain of the Fab is drawn with a light line and is mainly on the left side of the figure and in front of the cavity. The heavy chain is drawn with medium lines and is mainly to the right of the figure and behind the cavity. The bound lysozyme is drawn with heavy lines and is at the top of the figure.

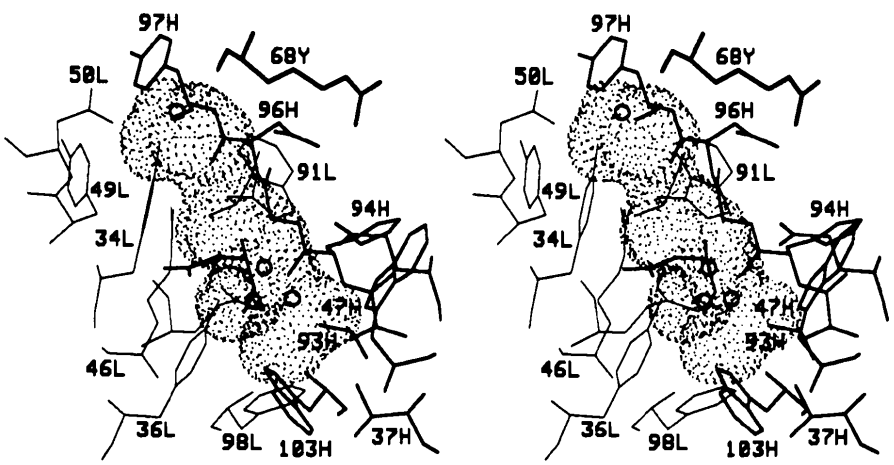

Fig. 7. *ORTEP* (Johnson, 1965) stereoview of a closeup of the cavity between V_L, V_H and the bound lysozyme. The four water molecules buried in the cavity are shown as open circles. The residues that line the surface of the cavity are shown. Residue $Arg68_Y$ is shown in bold.

Table 4. *Hydrogen bonds between buried water molecules and protein atoms*

Water	Protein	Distance (Å)
Wat135	$Gly49_Y$ O	2.7
	$Thr51_Y$ $O^{\gamma 1}$	2.8
	$Asp66_Y$ $O^{\delta 2}$	2.8
	$Arg68_Y$ N^{ϵ}	3.3
	$Thr69_Y$ $O^{\gamma 1}$	2.9
Wat138	$Tyr34_L$ O^{η}	3.5
	$Asp50_L$ $O^{\delta 1}$	3.2
	$Tyr97_H$ O^{η}	3.3
	$Arg68_Y$ O	2.6
Wat152	$Tyr53_Y$ O	2.9
	$Lys56_Y$ N	3.0
	$Ser91_Y$ O^{γ}	2.8
Wat155	$Tyr36_L$ O^{η}	3.3
	$Gln89_L$ $N^{\eta 2}$	2.9
	$Glu35_H$ $O^{\epsilon 2}$	3.2
Wat160	$Tyr36_L$ O^{η}	2.7
Wat169	$Trp35_L$ $N^{\epsilon 1}$	3.0
	$Thr51_L$ O	2.8
	$Ser65_L$ N	3.2
	$Ser65_L$ O	2.9
Wat193	$His94_H$ O	2.4
	$Trp103_H$ $N^{\eta 1}$	3.0

Table 5. *Residues that line the cavity between V_L and V_H*

V_L	V_H	Lysozyme
$Tyr34_L$	$Glu35_H$	$Arg68_Y$
$Tyr36_L$	$Val37_H$	
$Arg46_L$	$Trp47_H$	
$Tyr49_L$	$Leu93_H$	
$Asp50_L$	$His94_H$	
$Gln89_L$	$Gly95_H$	
$Trp91_L$	$Asn96_H$	
$Phe98_L$	$Tyr97_H$	
	$Asp101_H$	
	$Trp103_H$	

the cavity in the HyHEL-5–lysozyme complex. When the Fv domains of the HyHEL-5–lysozyme and the D1.3–lysozyme complexes are superimposed the two cavities overlap by only a small amount. The cavity in the Fv D1.3–lysozyme complex is located more in the antigen–antibody interface and less in the V_HV_L interface and contains four water molecules (Bhat *et al.*, 1994).

Three additional buried water molecules are located in the HyHEL-5–lysozyme complex (Table 4) in positions unlikely to influence antigen binding. Wat135 is buried within the bound lysozyme (Fig. 8) in the same location as Wat326 of tetragonal lysozyme (Kurinov & Harrison, 1995; PDB reference 1LSE). These water molecules are located 0.2 Å from each other when the lysozymes are superimposed using their C^{α} coordinates. Both water molecules make the same set of hydrogen bonds in their respective structures. It is not likely that this water contributes to the antigen–antibody interface in as much as the water molecule is not accessible to bulk solvent in the free lysozyme. The other two buried water molecules in the complex are far removed from the antigen–antibody interface. Wat152 is located in the bottom of the active-site cleft of the bound lysozyme. Wat169 is located in the V_L domain near the side of the antibody, about 14 Å from the axis of the complex.

3.3. *Comparison of current structure with 2HFL*

The most significant changes in the main chain of the model resulting from the refinement are at the carboxyl end of the C_L domain, the beginning of the V_H domain and in a loop of C_{H1} centered at residue $Gly127_H$. None of these points is internal. The end of C_L and the beginning of V_H are both poorly defined. The loop in C_{H1} is relatively near the end of C_L, both being at the end of the Fab opposite to the antigen binding site. No significant main-chain differences have been found in the lysozyme. For the entire complex, the r.m.s. difference in the 554 common C^{α} positions is 0.820 Å. Omitting the six C^{α} positions that differ by more than 3.5 Å ($Asn212_L$, $Glu213_L$, $Pca1_H$, $Gly127_H$, $Ala129_H$, $Thr134_H$) decreases the difference to 0.642 Å. Of these six residues, the first two are at the carboxyl end of C_L and the third is at the amino terminus of V_H. $Gly127_H$, $Ala129_H$ and $Thr134_H$ are in a long loop found in weak density in most Fab's and interpreted differently than in 2HFL.

Some side-chain differences are large, but none alter the basic structure of the complex. In a comparison of the current model with 2HFL, the 29 residues showing a change greater than 3.0 Å in the position of their side chains are all on the surface of the molecule. 12 of the residues are found in weak and ambiguous density; six residues are found at the beginning or end of a polypeptide chain and are different due to a different interpretation of the chain trace. 12 residues are in good or reasonable density and are materially better placed than in the 2HFL model. One residue, $Gln105_H$, might be better modeled as discrete disordered with one conformer as in 2HFL and one as current. 14 of the residues are in very weak density but would not be placed as in 2HFL.

The residues of the Fab–lysozyme interface show very little difference in the two models. Angle χ_1 of $Thr43_Y$ is turned by 80°, now agreeing with its conformation in the 1.7 Å model of tetragonal lysozyme (Kurinov & Harrison, 1995; PDB reference 1LSE). Residue $Asn44_Y$, a surface residue, might be modeled as discretely disordered with one rotamer corresponding to the current model and one to 2HFL. No solvent molecules were defined in the 2HFL model.

The elbow bends previously reported for the two crystal forms (Sheriff, Silverton *et al.*, 1987) are essentially unchanged. The final values resulting from this refinement are 162° for the long axis form and 155° for

the short axis form, both differing by only 1° from the previously published values.

3.4. *Comparison of HyHEL-5–lysozyme (WT) with HyHEL-5–lysozyme (R68K)*

The HyHEL-5–lysozyme complex has been independently determined in two forms: the complex with wild-type hen egg-white lysozyme refined in this report and the complex with a mutant HEL having Arg68 replaced by a lysine (Chacko *et al.*, 1995; PDB reference 2IFF). The two structures, refined by different investigators using the same refinement package and data of equivalent resolution, provide a means of testing the reliability of the coordinate sets. Aside from the point mutation at residue 68_Y, the sequences are identical and the structures are essentially the same. Differences observed between the two structures should reflect the actual coordinate errors in the structures. The coordinate uncertainty estimated from a Luzzati (1952) plot for the WT structure is 0.3 Å. The coordinate uncertainty in the 2IFF structure is estimated from the published Luzzati plot as 0.25 Å. If these error estimates are reliable, the r.m.s. difference in coordinates between the two structures would be 0.4 Å. Comparison of the two structures, examining only atoms that are common to the two structures, yields an r.m.s. difference of 0.35 Å in the coordinates for the C^α atoms, 0.4 Å for all main-chain atoms, 0.7 Å for all atoms, and 0.9 Å for all atoms other than main-chain atoms. Contrary to the findings of Daopin, Davies, Schlunegger & Grütter (1994) in their comparison of two, higher resolution TGF-β2 structures, the r.m.s. coordinate differences calculated from the individual Luzzati plots is somewhat smaller than the observed r.m.s. differences for all atoms. Chambers & Stroud (1979) have pointed out that estimates derived from Luzzati's (1952) formalism are dominated by atoms having the lowest B parameters. We must, therefore, assume that the observations here regarding the r.m.s. coordinate differences predicted based on the Luzzati plots either are low for the entire set of coordinates or represent best the main-chain atoms that are likely to be the most reliably determined in a protein structure. With this consideration, the predicted r.m.s. coordinate difference is reasonable.

No statistically significant coordinate differences are found between the WT model and the R68K model in the antibody–antigen interface. With one exception, all large differences are located in regions remote from the interface. This exception involves $Thr57_H$ that makes a cross-interface contact (Fig. 3) *via* its carbonyl O atom and has two atoms that are partially buried in the interface: the N and O atoms. In the R68K model, the side chain of that residue, which points away from the interface, has the angle χ_1 turned by ~180° from its conformation in the WT model. At this resolution, the density for the side chain of this residue cannot select the proper value for this angle. The largest r.m.s. discrepancy in main-chain coordinates is 2.0 Å at $Val2_H$. Six of the most significant main-chain and C^α-atom differences between the two molecules derive from interpretation of weak electron density. The five remaining instances involve the flipping of the peptide plane by approximately 180° following residues $Thr108_L$, $Ser167_L$, $Ser215_H$,

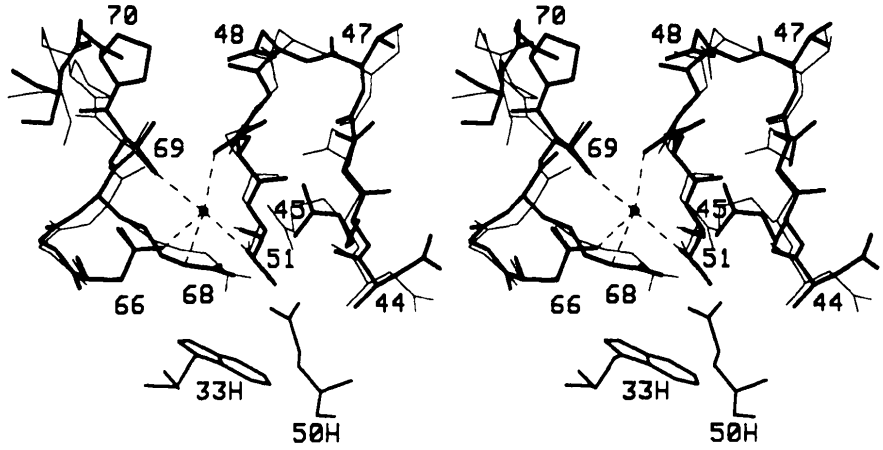

Fig. 8. *ORTEP* (Johnson, 1965) stereoview showing the comparison of the structure of HyHEL-5 bound lysozyme with that of tetragonal lysozyme (1LSE; Kurinov & Harrison, 1995) in the vicinity of $Arg45_Y$ and $Arg68_Y$. Also included is the water molecule trapped wholly within the lysozyme and common to the two structures. The five hydrogen bonds made with the water molecule are shown for the current work. The water molecule of the native structure has the same hydrogen bonds as are formed in the complex and is coincident with the water molecule of the complex in this view. The difference in conformation at $Pro70_Y$ is also shown as well as the different orientations of the two surface asparagine residues, 44 and 46. The 180° difference in χ_1 at $Thr47_Y$ puts $Thr47_Y$ within 30° of a canonical rotamer (Ponder & Richards, 1987). The bound lysozyme is drawn with heavy lines and the tetragonal lysozyme with light lines. Residues $Trp33_H$ and $Glu50_H$ are included to show the relationship of the lysozyme molecule to the antibody–antigen interface and are drawn with a medium weight line.

Phe3$_Y$ and Arg21$_Y$. The largest r.m.s. shift in side-chain coordinates is 2.65 Å found at Asn135$_H$. This residue and five others that show an r.m.s. shift of side-chain atoms exceeding 2 Å are located in regions of weak or poor density. The side chains of two other residues that are located in good density, Ser180$_H$ and Trp92$_Y$, also show r.m.s. shifts of side-chain atoms exceeding 2 Å due to differences in interpretation.

The refined model of the WT complex includes 82 water molecules. The model of the mutant R68K complex includes 80 water molecules. Seven pairs of these water molecules are located within 1.1 Å from each other upon superposition and may represent the same water molecule in the two models. The most certain assignment is a buried water molecule in the bound lysozyme, Wat135, which falls only 0.10 Å from HOH805 in 2IFF. The remaining pairs include Wat194 (0.5 Å to HOH782), Wat190 (0.6 Å to HOH732), Wat186 (0.6 Å to HOH730), Wat160 (0.7 Å to HOH748), Wat179 (0.9 Å to HOH804) and Wat211 (1.1 Å to HOH750). Wat160/HOH748 is located in a cavity between V_L and V_H that is inaccessible to bulk solvent. The other five sites listed are all on the surface of the Fab.

3.5. *Comparison of bound and free lysozyme*

Comparison of the four models of free hen egg-white lysozyme available from three 2 Å or better structure determinations (triclinic, 2.0 Å, PDB reference 2LZT, Ramanadham *et al.*, 1990; monoclinic, 1.7 Å, PDB reference 1LYS, two independent molecules, Harata, 1994; tetragonal, 1.7 Å, PDB reference 1LSE, Kurinov, & Harrison, 1995) with each other shows that the lysozyme molecule is somewhat flexible at several points. The bound lysozyme in the HyHEL-5–lysozyme complex is most similar to the tetragonal form (Fig. 9*d*) with an r.m.s. mismatch of 0.51 Å at the C^α positions. The most significant difference in the main-chain trace occurs at Pro70$_Y$–Gly71$_Y$ where we find 1.55 Å and 1.87 Å mismatches. From the monoclinic form of lysozyme, 1LYS, we see that the region of the molecule around Pro70–Gly71 is very variable between the two molecules in the crystal. In the 1LYS molecules, the distance between the corresponding Gly71 C^α positions after superposition is 3.7 Å and appears as a sharp peak in the plot of the distances between the C^α positions (Harata, 1994). This is also apparent in Fig. 9 in which we see a notable difference of 3.9 Å at Pro70–Gly71 for the comparison of the bound lysozyme with molecule 2 of 1LYS (Fig. 9*c*) but little significant difference in the same region of the comparison with molecule 1 of 1LYS (Fig. 9*b*). In the complex, residues Pro70$_Y$ and Gly71$_Y$ are both buried in the interface with the Fab but in the crystals of the free HEL, these residues are not in contact with other molecules. The notable difference around Arg45$_Y$–Asp48$_Y$, with a maximum of 3.16 at Thr47$_Y$, in the comparison with 2LZT (Fig. 9*a*) is due to a rearrangement of that segment of chain of 2LZT with respect to all of the other molecules. The third region of apparently significant mobility in the lysozyme molecule is centered on Gly102$_Y$ with a maximum mismatch at C^α of 4.35 Å in the comparison with molecule 1 of 1LYS (Fig. 9*b*). This residue corresponds to a peak of size 3.4 Å in the plot comparing the two molecules of 1LYS (Harata, 1994) and is located in a loop on the surface of the molecule. These results are fully consistent with the study of Artymiuk *et al.* (1979) where the *B* parameters of chicken and human lysozymes showed prominent peaks at residues 47, 70–73 and 101–102.

The plane of the peptide link Leu25$_Y$–Gly26$_Y$ is flipped by approximately 180° in the bound lysozyme compared with the free lysozyme models. The planes of three other links agree with three of the four models for free lysozyme: Gly102$_Y$–Asn103$_Y$, for which 2LZT is in disagreement, Arg125$_Y$–Cys126$_Y$, 1LSE disagrees, and Arg128$_Y$–Leu129$_Y$, where molecule 2 of 1LYS disagrees.

The side chains of 17 residues of the bound lysozyme have significantly different conformations in comparison with one or more of the four free HEL molecules. Three

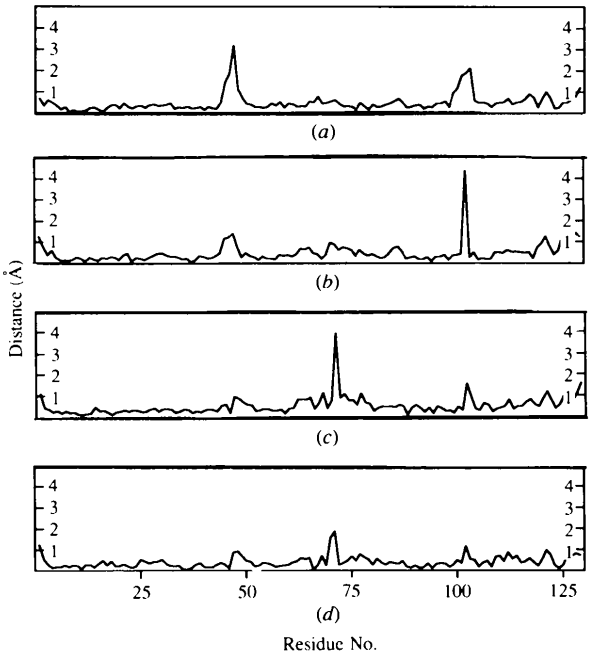

Fig. 9. Distances between the C^α atoms of the bound lysozyme and four high-resolution structures of free lysozyme. The distance scale for each section, (*a*)–(*d*), of the figure runs from 0 to 5 Å. (*a*) The triclinic crystal form, 2LZT. The r.m.s. distance between pairs of C^α atoms is 1.51 Å. (*b*) Molecule 1 of the monoclinic crystal form, 1LYS. The r.m.s. distance between pairs of C^α atoms is 1.54 Å. (*c*) Molecule 2 of the monoclinic crystal form, 1LYS. The r.m.s. distance between pairs of C^α atoms is 2.49 Å. (*d*) The tetragonal crystal form, 1LSE. The r.m.s. distance between pairs of C^α atoms is 0.51 Å.

residues that are in the epitope, $Arg45_Y$, $Asp48_Y$ and $Arg68_Y$, have their side chains in somewhat different conformations in the molecules of 1LYS and 2LZT. All the other residues are surface residues that are generally not near the interface. $Trp62_Y$ and $Arg128_Y$ are involved in packing contacts in the 1LSE crystals and are differently oriented with respect to the bound lysozyme. In all, the differences observed in the residues other than those involved in the interface are not considered of consequence.

The final $2F_o - F_c$ map for the model indicates some disorder for the indole rings of residues $Trp62_Y$ and $Trp63_Y$. The disorder in this region led to the original structure report in which the side chain of $Trp63_Y$ was turned by 180° at χ_1 with respect to its conformation in the native lysozyme structure (Sheriff, Silverton *et al.*, 1987). The structure reported here adopts a conformation consistent with the native structure. The four examples of native lysozyme molecules show a range of conformers for $Trp62_Y$ and general agreement at $Trp63_Y$. At Trp62 the angle χ_1 varies in the range −32 to −59° and the χ_2 value ranges from −79 to +98°. The corresponding ranges for Trp63 are χ_1 = −52 to −59° and χ_2 = 103 to 110°. The refinement of the complex sets the angles for $Trp62_Y$ at (χ_1, χ_2) = (−58, 50°) and for $Trp63_Y$ at (χ_1, χ_2) = (−45, 102°). Crystal structures of lysozyme with bound oligosaccharides [Blake *et al.* (1967); Cheetham, Artymiuk & Phillips (1992); Strynadka & James (1991); Ford, Johnson, Machin, Phillips & Tjian (1974)] indicate that there is a stacking interaction between the indole ring of Trp62 and the non-polar surface of the NAM residue in site *B* of the substrate. This role in substrate binding may require enhanced mobility in the side chain of Trp62.

4. Conclusions

The structure of the refined complex presented here confirms and elaborates on the major findings previously presented (Sheriff, Silverton *et al.*, 1987). The interface of the complex is composed of two highly complementary protein units; two surfaces, each of more than 700 Å^2, fit fairly tightly with the incorporation of only three waters. While the residues involved in the interface have been identified with greater certainty, their original identification is mostly confirmed. The refined structure reveals a large cavity between V_H and V_L containing four water molecules. The waters participate in hydrogen bonding with the neighboring protein residues.

The authors thank Drs Susan Chacko and Eduardo Padlan for their critical reading of the manuscript.

References

Amit, A. G., Mariuzza, R. A., Phillips, S. E. V. & Poljak, R. J. (1986). *Science*, **233**, 747–753.

Artymiuk, P. J., Blake, C. C. F., Grace, D. E. P., Oatley, S. J., Phillips, D. C. & Sternberg, M. J. E. (1979). *Nature (London)*, **280**, 563–568.

Ban, N., Escobar, C., Garcia, R., Hasel, K, Day, J., Greenwood, A., McPherson, A. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 1604–1608.

Bentley, G. A., Boulot, G., Riottot, M. M. & Poljak, R. J. (1990). *Nature (London)*, **348**, 254–257.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr, Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Bhat, T. N., Bentley, G. A., Boulot, G., Greene, M. I., Tello, D., Dall'Acqua, W., Souchon, H., Schwartz, F. P., Mariuzza, R. A. & Poljak, R. J. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 1089–1093.

Bhat, T. N., Bentley, G. A., Fischmann, T. O., Boulot, G. & Poljak, R. J. (1990). *Nature (London)*, **347**, 483–485.

Bizebard, T., Gigant, B., Rigolet, P., Rasmussen, B., Diat, O., Bvsecke, P., Wharton, S. A., Skehel, J. A. & Knossow, M. (1995). *Nature (London)*, **376**, 92–94.

Blake, C. C. F., Johnson, L. N., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1967). *Proc. R. Soc. London Ser. B*, **167**, 365–377.

Bossart-Whitaker, P., Chang, C. Y., Novotny, J., Benjamin, D. C. & Sheriff, S. (1995). *J. Mol. Biol.* **253**, 559–575.

Braden, B. C., Souchon, H., Eiselé, J.-L., Bentley, G. A., Bhat, T. N., Navaza, J. & Poljak, R. J. (1994). *J. Mol. Biol.* **243**, 767–781.

Brooks, B., Bruccoleri, R., Olafson, B., States, D., Swaminathan, S. & Karplus, M. (1983). *J. Comput. Chem.* **4**, 187–217.

Brünger, A. T. (1992). *X-PLOR Manual*. New Haven, CT, USA: Yale University Press.

Brünger, A. T., Kuriyan, J. & Karplus, M. (1987). *Science*, **235**, 458–460.

Chacko, S., Silverton, E., Kam-Morgan, L., Smith-Gill, S., Cohen, G. & Davies, D. (1995). *J. Mol. Biol.* **245**, 261–274.

Chambers, J. L. & Stroud, R. M. (1979). *Acta Cryst.* B**35**, 1861–1874.

Cheetham, J. C., Artymiuk, P. J. & Phillips, D. C. (1992). *J. Mol. Biol.* **224**, 613–628.

Chitarra, V., Alzari, P. M., Bentley, G. A., Bhat, T. N., Eiselé, J.-L., Houdusse, A., Lescar, J., Souchon, H. & Poljak, R. J. (1993). *Proc. Natl Acad. Sci. USA*, **90**, 7711–7715.

Cohen, G. H. (1993). *J. Appl. Cryst.* **26**, 495–496.

Connolly, M. L. (1983). *J. Appl. Cryst.* **16**, 548–558.

Connolly, M. L. (1993). *J. Mol. Graphics*, **11**, 139–141.

Daopin, S., Davies, D. R., Schlunegger, M. P. & Grütter, M. G. (1994). *Acta Cryst.* D**50**, 85–92.

Davies, D. R., Padlan, E. A. & Sheriff, S. (1990). *Annu. Rev. Biochem.* **59**, 439–473.

Davies, D. R., Sheriff, S. & Padlan, E. A. (1988). *J. Biol. Chem.* **263**, 10541–10544.

Davies, D. R., Sheriff, S. & Padlan, E. (1989). *Cold Spring Harbor Symp. Quant. Biol.* **LIV**, 233–238.

Engh, R. A. & Huber, R. (1991). *Acta Cryst.* A**47**, 392–400.

Evans, S. V., Rose, D. R., To, R., Young, N. M. & Bundle, D. R. (1994). *J. Mol. Biol.* **241**, 691–705.

Fields, B. A., Goldbaum, F. A., Ysern, X., Poljak, R. J. & Mariuzza, R. A. (1995). *Nature (London)*, **374**, 739–742.

Fischmann, T. O., Bentley, G. A., Bhat, T. N., Boulot, G., Mariuzza, R. A., Phillips, S. E. V., Tello, D. & Poljak, R. J. (1991). *J. Biol. Chem.* **266**, 12915–12920.

Flocco, M. M. & Mowbray, S. L. (1994). *J. Mol. Biol.* **235**, 709–717.
Ford, L. O., Johnson, L. N., Machin, P. A., Phillips, D. C. & Tjian, R. (1974). *J. Mol. Biol.* **88**, 349–371.
Gelin, B. R. & Karplus, M. (1979). *Biochemistry*, **18**, 1256–1268.
Grivel, J.-C. & Smith-Gill, S. J. (1996). *Structure of Antigens*, Vol. 3, edited by M. H. Van Regenmotel, pp. 91–154. Boca Raton, Florida: CRC Press.
Harata, K. (1994). *Acta Cryst.* D**50**, 250–257.
Hendrickson, W. A. (1985). *Methods Enzymol.* **115**, 252–270.
Hibbits, K. A., Gill, D. S. & Willson, R. C. (1994). *Biochemistry*, **33**, 3584–3590.
Jack, A. & Levitt, M. (1978). *Acta Cryst.* A**34**, 931–935.
Johnson, C. K. (1995). *ORTEP: A Fortran Thermal-Ellipsoid Plot Program for Crystal Structure Illustrations*. Oak Ridge National Laboratory, Report ORNL-3794. Revised.
Jones, T. A. (1978). *J. Appl. Cryst.* **11**, 268–272.
Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.
Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M. & Gottesman, K. S. (1987). *Sequences of Immunological Interest*, 4th ed. Bethesda, Maryland, USA: National Institutes of Health.
Kurinov, I. V. & Harrison, R. W. (1995). *Acta Cryst.* D**51**, 98–109.
Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.
Laskowski, R. A., Moss, D. S. & Thornton, J. M. (1993). *J. Mol. Biol.* **231**, 1049–1067.
Levitt, M. & Perutz, M. F. (1988). *J. Mol. Biol.* **201**, 751–754.
Luzzati, V. (1952). *Acta Cryst.* **5**, 802–810.
Malby, R. L., Tulip, W. R., Harley, V. R., McKim-Breschkin, J. L., Laver, W. G., Webster, R. G. & Colman, P. M. (1994). *Structure*, **2**, 733–746.
Milner-White, E. J. & Poet, R. (1987). *Trends Biochem. Sci.* **12**, 198–192.
Milner-White, E. J., Ross, B. M., Ismail, R., Belhaj-Mostefa, K. & Poet, R. (1988). *J. Mol. Biol.* **204**, 777–782.
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.
Paterson, Y., Englander, S. W. & Roder, H. (1990). *Science*, **249**, 755–759.
Ponder, J. W. & Richards, F. M. (1987). *J. Mol. Biol.* **193**, 775–791.
Prasad, L., Sharma, S., Vandonselaar, M., Quail, J. W., Lee, J. S., Waygood, E. B., Wilson, K. S., Dauter, Z. & Delbaere, L. T. J. (1993). *J. Biol. Chem.* **268**, 10705–10708.
Ramachandran, G. N. & Sasisekharan, V. (1968). *Adv. Protein Chem.* **28**, 283–437.
Ramanadham, M., Sieker, L. C. & Jensen, L. H. (1990). *Acta Cryst.* B**46**, 63–69.
Sasisekharan, V. (1962). *Collagen*, edited by N. Ramanathan, pp. 39–78. New York: Interscience Publishers.
Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. (1986). *J. Mol. Biol.* **190**, 593–604.
Sheriff, S., Hendrickson, W. A. & Smith, J. L. (1987). *J. Mol. Biol.* **197**, 273–296.
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.
Smith-Gill, S. J., Wilson, A. C., Potter, M., Prager, E. M., Feldmann, R. J. & Mainhart, C. R. (1982). *J. Immunol.* **128**, 314–322.
Strynadka, N. C. & James, M. N. G. (1991). *J. Mol. Biol.* **220**, 401–424.
Tulip, W. R., Varghese, J. N., Laver, W. G., Webster, R. G. & Colman, P. M. (1992). *J. Mol. Biol.* **227**, 122–148.
Tulip, W. R., Varghese, J. N., Webster, R. G., Laver, W. G. & Colman, P. M. (1992). *J. Mol. Biol.* **227**, 149–159.
Wilson, K. P., Malcolm, B. A. & Matthews, B. W. (1992). *J. Biol. Chem.* **267**, 10842–10849.
Wu, T. T. & Kabat, E. A. (1970). *J. Exp. Med.* **132**, 211–250.
Young, A. C. M., Dewan, J. C., Nave, C. & Tilton, R. F. (1993). *J. Appl. Cryst.* **26**, 309–319.