J. Mol. Biol. (1994) **236**, 247–274

Local and Transmitted Conformational Changes on Complexation of an Anti-sweetener Fab

Luke W. Guddat[1], Lin Shan[1], Jerry M. Anchin[2] D. Scott Linthicum[2] and Allen B. Edmundson[1]†

[1]*Harrington Cancer Center 1500 Wallace Boulevard Amarillo, TX 79106, U.S.A.*

[2]*Department of Veterinary Pathobiology Texas A & M University College Station, TX 77843, U.S.A.*

Crystal structures of an Fab (NC6.8) from a murine IgG2b(κ) antibody and its complex with a sweet-tasting, *N*-,*N'*-,*N''*-trisubstituted guanidine compound (NC174) have been determined by X-ray analysis. Both crystal forms are produced by a microseeding technique in polyethylene glycol (PEG) 8000 but the habits and space groups are very different. The native protein crystallizes as plates in the monoclinic space group $C2$ and the complex crystallizes as prisms in the orthorhombic space group $P2_12_12$. The structures were solved by molecular replacement methods, with the Fab fragments from the 4-4-20, HyHel-5 and BV04-01 antibodies as starting models. On binding of the ligand, *N*-(*p*-cyanophenyl)-*N'*-(diphenylmethyl)-*N''*-(carboxymethyl)guanidine, the protein exhibits significant local conformational changes in the active site, particularly in the third complementarity-determining region (CDR3) of the heavy chain. The ligand enters the small crevice by end-on insertion with the cyanophenyl group in the lead and the diphenyl rings partially protruding from the entrance. No strict π-π stacking interactions are observed. However, tyrosine L32 (CDR1), tyrosine L96 (CDR3) and tryptophan H33 (CDR1) help immobilize the cyanophenyl ring and guanido group, and tyrosine H96 moves about 4·5 Å to lie between the rings of the diphenyl group. The positive charge on the guanido group is compensated by glutamic acid H50 (CDR2) while the negative charge on acetic acid is neutralized by arginine H56 (CDR2) and by hydrogen bonding with asparagine H58 (CDR2). Water molecules participate in the binding process by hydrogen bonding with the cyano and guanido groups. The mechanism of binding is a clear example of induced fit. Like hemoglobin, the NC6.8 Fab can be classified as an allosteric protein, since its overall structure is altered by the binding of a small ligand. In crystals of the native Fab the elbow bend angle is 184° while in crystals of the complex the elbow angle is 153°. There is also a reciprocal push-pull type of change where the heavy chain is flexed and the light chain is extended. The tail of the heavy chain, which would be connected to the Fc in an intact antibody, is displaced 19 Å relative to its position in the unliganded Fab. Within the limited series of sweetener-Fab complexes we have thus far examined, only the NC174 hapten has produced such results. Complexes of NC6.8 Fab with NC24 or NC90 (2-dimethyl-4-methyl-bicyclo [2.2.1] heptanyl or cyclooctanyl derivatives of guanidine) crystallize in the same space group as the native NC6.8 and have structures resembling the latter rather than the Fab liganded with NC174.

Keywords: anti-sweetener Fab; antigen-antibody complex; transmitted conformational changes; molecular signalling; crystal structure

† Author to whom all correspondence should be addressed.

0022-2836/94/060247-28 $08.00/0

© 1994 Academic Press Limited

1. Introduction

In the very early stages of this project one of us (L.S.) added ligand to a suspension of crystals of the native Fab in mother liquor. These crystals immediately began to disintegrate and crystals with different habits rapidly took their place. It was apparent that we were witnessing a modern version of Felix Haurowitz's classic experiment (Haurowitz, 1938) in which crystals of deoxyhemoglobin were shattered by exposure to air to produce oxyhemoglobin. In both examples the introduction of a small molecule led to conformational changes that were incompatible with the crystal packing of the original macromolecular species. Examination of the crystal structure of the murine NC6.8 Fab in the presence and absence of the small sweet-tasting ligand NC174 is the subject of this paper.

There are many advantages in selecting this system for X-ray analyses. The amino acid sequences of the variable (V†) domains of the NC6.8 light (L) and heavy (H) chains have been deduced from nucleotide sequences (J.M.A. & D.S.L., unpublished results) and the protein is amenable to future studies by site-directed mutagenesis. With its three planar phenyl rings and a guanido group, the NC174 ligand should be easily recognized in difference Fourier maps. The average dissociation constant for complexation of the ligand with the intact IgG2b(κ) antibody in solution is $5{\cdot}3 \times 10^{-8}$ M (J.M.A. & D.S.L., unpublished results), a value suggesting that the ligand will probably not show significant mobility when bound to the Fab.

NC174 is 200,000 times sweeter than sucrose and consequently is a good candidate for using ligand-antibody interactions as a paradigm for the binding of sweet taste ligands by biological receptors. For example, one of the NC174 components, the *p*-cyanophenyl ring, is believed to interact with residues in a putative sweet taste receptor protein (Muller *et al.*, 1992). A model for a receptor site with balanced electrostatic and hydrophobic interactions has been proposed to match the properties of low energy conformers of various sweeteners (Walters *et al.*, 1991). Compounds with optimized electrostatic potential produce a sweet taste when bound to a receptor. Ligands best suited for spatial complementarity with the receptor show the highest potencies. NC174 has zwitterionic properties, hydrogen bond donors and acceptors and three aryl groups to contribute to such requirements for sweetness.

The expected mimicry of receptor sites can be explored in depth by crystallographic analyses. In turn, the X-ray results can be used as standards for computer assisted modeling and dynamics simulations of binding interactions with related ligands.

Before the crystallographic results were made available to anyone outside the Amarillo laboratory, independent investigations were carried out at Texas A & M University and the Universi Illinois at Urbana-Champaign to predict characterize the combining site of the NC6.8 body. Computer modeling techniques (Kussie 1991; Anchin *et al.*, 1991; Bassolino-Klimas 1992; Anchin & Linthicum, 1992; Mandal Lithicum, 1992) were used concomitantly fluorescence spectroscopy data (Droupadi 1992) to predict key antibody residues involve the recognition of the sweetener ligand. The agreement between the predicted and obser results will be presented elsewhere Viswanathan, J. Anchin, P. R. Droupadi Mandal, D. S. Linthicum & S. Subraman unpublished results).

† Abbreviations used: V, variable region; L, light; H, heavy; PEG, polyethylene glycol; CDR, complementarity determining region; C, constant region; r.m.s., root-mean-square; FR, framework region.

2. Materials and Methods

(a) *Purification of the intact antibody and production the Fab by papain hydrolysis*

The murine NC6.8 IgG2b(κ) antibody was prepare Texas A & M University by hybridoma technology Amarillo the ascites fluid was dialyzed overnight at against 20 mM histidine, 1 mM sodium azide, adjuste pH 5·6. The antibody exhibited euglobulin prope (limited solubility in solutions of low ionic strength largely precipitated in the dialysis bag. Precipitate f 50 ml of ascites fluid was harvested by centrifuga washed with the dilute histidine buffer and transferred a fine suspension to a 50 ml column of carboxym cellulose (CM 52, Whatman Lab. Sales, P.O. Box 1 Hillsboro, OR). After the column was washed with 40 Tris·HCl, 1 mM sodium azide (pH 8·0), 20°C, the body was eluted by adding solid NaCl to the buffer to a final concentration of 0·15 M. By SDS-PAGE in presence and absence of 2-mercaptoethanol, the prot was found to be free of detectable contaminants and used without further purification.

The protein was concentrated in a Micro-Co apparatus fitted with cellulose ester membranes wi molecular weight cut-off of 10,000 (Spectrum, Hous TX). During the concentration process the buffer w exchanged for a pH 6·8 solution comprised of 50 Bis-Tris, 50 mM Mops (3-(*N*-morpholino)propane sulf acid), 0·15 M NaCl and 1 mM sodium azide (chem purchased from Sigma Chemical Co., St Louis, MO).

Conditions for proteolytic hydrolysis with papain w first optimized in preliminary experiments. Suffic quantities of Fab for all X-ray analyses were th prepared in a single batch process. An 8 ml solut containing 21 mg of IgG_{2b} protein/ml of the buf described above was immersed in an ice water bat Mercuripapain (4·6 mg) was activated with 2-mercapt ethanol (3 μl of 14 M) in 200 μl of 5% EDTA solut (all chemicals from Sigma Chemical Co., St Louis, MO This mixture was added to the protein solution and t hydrolysis was allowed to continue for 3·5 h. The reac was stopped by the addition of 1 μl of 3 M iodoacetate alkylate the sulfhydryl group of papain. After 0·5 h solution was transferred to a 5 cm × 25 cm column high resolution Sephacryl S-200 (Pharmacia, Piscataw NJ), equilibrated at pH 8·0 and 22°C with 25 m Tris·HCl, 0·15 M NaCl, 1 mM sodium azide. The Fab w the last of 4 components eluted from the column and w well separated from Fc and larger degradation produc Electrophoretic analyses by SDS-PAGE (Laemmli, 19 indicated that the Fab was homogeneous (>98%) wi respect to size and further purification was deemed unn

cessary. The solution was concentrated and subjected to crystallization trials.

(b) *Crystallization of the native protein and the complex with the NC174 ligand*

Crystals of the native (unliganded) Fab were grown by a microseeding technique (Shan *et al.*, 1993). Irregular crystals, 0·1 mm in the smallest dimension, appeared spontaneously in 7% (w/v) polyethylene glycol (PEG) 8000 dissolved in 20 mM NaCl. One of these crystals was washed with 50 μl of 12% PEG 8000 in water. A sample of 5 μl of the wash suspension, which contained microseeding crystals, was drawn into a Hamilton syringe. Two microseeds were transferred into the protein solution by transiently inserting the tip of the syringe into an 80-μl droplet. Typically, the droplet contained 1·9 mg of Fab in a solution of 6·6% PEG 8000, 20 mM NaCl, 0·02% (w/v) sodium azide. In 6 days at 22°C, 2 crystals grew from seeds into plates 4 mm long, 2 mm wide and 0·2 mm thick.

Many small crystals were produced at 4°C when 0·3 mg of NC174 sweetener was added to 200 μl of a solution containing 12 mg of Fab (i.e. 60 mg/ml) in 20 mM NaCl, 0·02% sodium azide. The crystals were dissolved by increasing the NaCl concentration and were then recrystallized with PEG 8000 into forms suitable for X-ray analyses. Microseeds were prepared from crystals spontaneously appearing in PEG 8000 as described above. One seed crystal was transferred into a 15 μl droplet of a solution consisting of 17 mg of ligand-Fab complex/ml, 4·8% PEG 8000, 80 mM NaCl and 0·02% sodium azide. A prismatic crystal, with a flattened face where it adjoined a glass surface, reached its mature size of 1·5 mm × 1·3 mm × 0·4 mm in about 2 weeks. Other crystals of similar size were subsequently produced by the same procedure.

(c) *Properties of the crystals*

The unliganded NC6.8 Fab crystallizes in the monoclinic space group $C2$, with $a=139{\cdot}9$ Å, $b=51{\cdot}4$ Å, $c=96{\cdot}2$ Å and $\beta=132{\cdot}8°$. In each unit cell there are 4 Fab molecules, with 1 Fab as the asymmetric unit. Protein molecules occupy 2·52 Å^3/Da and the solvent accounts for 51·4% of the volume of the unit cell (Matthews, 1968).

As emphasized in section 1. Introduction, above, addition of the ligand leads to changes incompatible with the original packing of the native protein in the crystal lattice. In cocrystals of the ligand and protein, it is therefore not surprising to find a different space group: orthorhombic $P2_12_12$, with $a=131{\cdot}0$ Å, $b=114{\cdot}1$ Å and $c=37{\cdot}3$ Å. Four molecules are present in the unit cell and there is 1 molecule/asymmetric unit. The space occupied by the ligand-protein complex is 2·78 Å^3/Da and the solvent content (55·9%) is slightly higher than in the crystals of the native protein. Values for the fractional volumes of protein and solvent in crystals of both the unliganded and liganded forms fall within the normal range exhibited by other proteins (Matthews, 1968).

(d) *Collection of X-ray diffraction data*

Data for the unliganded protein were collected with a Siemens area detector by a procedure previously described (Fan *et al.*, 1992). The rotating anode was operated at 35 kV and 80 mA and the crystal-to-detector distance was set at 12·0 cm. The oscillation width for each frame was 0·2° and the exposure time was 90 s. The total number of frames was 1800, 900 at each of 2 spindle settings. Since negligible levels of decay were detected in 45 h of crystal exposure to the X-ray beam, only 1 crystal was used to collect the entire data set (which extended to d spacings of 2·6 Å). Data were processed on a Silicon Graphics workstation with the XENGEN package of programs (Howard *et al.*, 1987).

Diffraction data for the ligand-protein complex were collected with synchrotron radiation, in combination with a 430 mm Weissenberg camera (Sakabe, 1983), at the Photon Factory in Tsukuba, Japan. Reflections were recorded on image plates for 2 orientations of the crystal. The crystal and the image plate were enclosed in a chamber filled with helium. Frames (27) were collected with the a axis of the crystal parallel to the spindle. The exposure time for each frame was 25 s, the coupling constant was 1·5°/mm and the oscillation width was 4°. With the b axis parallel to the spindle, the exposure time for each of 14 frames was 41 s, the coupling constant was 1·5°/mm and the oscillation range was 5°. Intensities were recorded with a BA100 scanner as soon as possible after removal from the chamber. Data to 2·2 Å resolution were processed with the WEIS program package (Higashi, 1989).

(e) *Determination of the crystal structures of the native and liganded forms*

The structures were solved by molecular replacement methods (Fitzgerald, 1988; Rossmann, 1990). Starting models were initially chosen after comparisons of the amino acid sequences of NC6.8 with those in crystal structures of other Fabs. For the V_L domain the greatest homology (89%) was found in the 4-4-20 Fab (Bedzyk *et al.*, 1990), while the V_H domain was most closely similar (88%) to the HyHEL-5 Fab (Sheriff *et al.*, 1987). Since the 4-4-20 structure was elucidated by our group (Herron *et al.*, 1989), we are more familiar with its properties. However, the heavy chains of NC6.8 and 4-4-20 belong to different IgG subclasses (2b *versus* 2a). The Fab from the BV04-01 IgG_{2b} (Herron *et al.*, 1991) was therefore selected as an alternative starting model. Crystals of the complex of 4-4-20 with fluorescein are isomorphous with those of the unliganded form of the BV04-01 Fab (the complex of BV04-01 with a trinucleotide crystallizes in a different space group). While these starting models are not necessarily interchangeable, we expected the initial results of the search routines to have some features in common.

Calculations for the rotation function (Crowther, 1972) and the translation function (Crowther & Blow, 1967) were performed with the MERLOT package of programs (Fitzgerald, 1988). Translation searches based on the use of correlation coefficients were conducted with the RTMAP program written by X.-M. He (Fan *et al.*, 1992).

Rotation function calculations for the unliganded protein were first carried out with the intact Fab molecules and diffraction data between 15 Å and 4·0 Å resolution. Probe molecules were then split into the V_L-V_H and C_L-C_H1 pairs of domains and the calculations were repeated. The same strategy was followed for the translation searches, but in some instances the data sets were truncated to include reflections between 8·0 Å and either 4·7 Å or 4·0 Å resolution.

Comparable calculations were performed to solve the structure of the ligand-protein complex. In the translation searches with RTMAP, however, only the V_L-V_H pairs were used as probes.

(f) *Crystallographic refinement of the structures*

The model of the unliganded Fab obtained by molecular replacement was subjected to rigid body refine-

ment with a program in the X-PLOR package (Brünger *et al.*, 1989). Sequences of the probe molecules were then replaced by those of the NC6.8 light and heavy chains, as determined by 2 of us (J.M.A. and D.S.L.). With the complementarity-determining regions (CDRs) omitted, X-PLOR was used for positional (Jack–Levitt) refinement of this model with 10 Å to 3 Å data. By alternating positional refinement with cycles of model building using TURBO-FRODO (Roussel & Cambillau, 1989), we were able to place the CDRs with confidence. Molecular dynamics refinement using X-PLOR was then applied to the model of the complete Fab. One cycle with 10 Å to 3 Å data was followed by 3 cycles with 10 Å to 2·6 Å data. Simulations involved the heating of the structure to 3000 K and subsequent slow cooling. After refinement of individual temperature (B) factors, 126 water molecules were added and the structure was again refined by the molecular dynamics protocol. Solvent atoms were assigned to electron density peaks greater than 3σ above background in $F_o - F_c$ maps. It was also necessary for the water molecules to be located within 5 Å of protein atoms and to have B factor values lower than 50 Å^2. Assignments were reexamined if the B factor values were greater than 30 Å^2.

Refinement of the ligand-protein complex followed the general outline presented for the unliganded Fab. The procedure for rigid body refinement was utilized to define the rotational and translational shifts of the 4 individual domains on ligand binding. In the initial applications of the molecular dynamics protocol with 10 Å to 2·7 Å data, both ligand and solvent molecules were omitted from the model. The data set was then expanded to include d spacings of 2·5 Å and the ligand and 125 water molecules were added. Extra electron density corresponding to the ligand was present in the earliest $F_o - F_c$ maps and the fitting of an appropriate model was straightforward. The structure of the complex was subjected to 4 cycles of molecular dynamics refinement. Individual B fac values were also refined. Finally, 5985 additional ref tions (to 2·2 Å resolution) were included, the total num of solvent molecules was increased to 296 and the m was again refined by molecular dynamics.

3. Results

(a) *Data collection*

Statistics for the collection of X-ray diffrac data for the native protein with an area detector a summarized in Table 1A. Synchrotron data for th ligand-Fab complex are presented in Table 1B. I both examples the data sets are more than 9 complete (observed/theoretical) at 2·8 Å resoluti but fall off in the higher order shells. While it is cle that the weaker reflections at higher resolution a not measured as accurately as stronger reflections other shells, the R_{merge} values are quite respectab (5·37 and 7·19%, respectively) for the entire da sets. The average redundancy in the number measurements for each ("unique") reflection is 3 for the native protein and 3·6 for the complex.

In combination the synchrotron radiation and th Weissenberg camera produce moderately hig resolution data with good statistics. Data for th native protein are less comprehensive. The ma difference in the two systems is the quality of th crystals. For example, the complex produces mor robust crystals which are easier to manipulat Differences are most noticeable when one of the th faces of a native crystal plate is rotated into th X-ray beam. In this orientation diffraction da extend only to about 3·0 Å resolution.

Table 1

Summary of X-ray diffraction data

Shell (Å)	Unique reflections: observed	Unique reflections: theoretical	Percentage obs./theor.	No. of measurements	R_{merge} (%)
A. *Data for the native NC6.8 Fab collected with a Siemens area detector*					
20·00–5·00	2230	2230	100·0	8219	2·65†
5·00–4·06	1952	1952	100·0	6417	4·73†
4·06–3·57	1889	1906	99·1	5509	7·31†
3·57–3·26	1893	1924	98·4	4367	9·20†
3·26–3·03	1856	1913	97·0	4070	15·23†
3·03–2·86	1794	1911	93·9	4003	20·24†
2·86–2·72	1646	1914	86·0	2745	25·27†
2·72–2·60	1014	1888	53·7	1583	31·71†
20·00–2·60	14,274	15,638	91·3	36,913	5·37†
B. *Data for NC174-NC6.8 Fab complex collected with synchrotron radiation and a Weissenberg camera at the Photon Factory, Japan*					
20·00–4·30	4106	4138	99·2	15,454	5·04‡
4·30–3·46	3610	3669	98·4	16,725	5·31‡
3·46–3·03	3521	3637	96·8	16,764	6·02‡
3·03–2·76	3262	3600	90·6	20,432	7·93‡
2·76–2·57	2925	3589	81·5	20,375	10·23‡
2·57–2·42	2634	3574	73·7	16,105	16·15‡
2·42–2·30	2414	3555	67·9	12,254	18·34‡
2·30–2·20	2072	3524	58·8	8432	17·94‡
20·00–2·20	24,544	29,286	83·8	126,541	7·19‡

Reflections were considered observed if the intensity $I > \sigma(I)$ where σ is the standard deviation.
† R_{merge} is the unweighted-squared R-factor on intensity $\times 100$.
‡ $R_{merge} = \Sigma|\langle I\rangle - I_i|/\Sigma\langle I\rangle \times 100$.

Table 2

Results of rotation and translation searches

Rotation searches (Crowther, 1972)

Model	Resolution (Å)	Euler angles α°	β°	γ°	Peak height (%)	2nd peak height (%)	Rank	σ
A. Results for native NC6.8 Fab								
4420 Fab	15·0–4·0	22·5	68·0	305·0	100·0	61·2	1	5·65
4420 V_L-V_H	15·0–4·0	22·5	67·0	305·0	100·0	67·6	1	5·37
4420 C_L-C_H1	15·0–4·0	22·5	72·0	300·0	100·0	94·5	1	3·63
BV0401 Fab	15·0–4·0	20·0	68·0	305·0	100·0	52·7	1	7·77
BV0401 V_L-V_H	15·0–4·0	22·5	69·0	305·0	100·0	85·9	1	4·01
BV0401 C_L-C_H1	15·0–4·0	20·0	68·0	305·0	100·0	83·4	1	4·22

Translation searches (Crowther & Blow, 1967)

Model	Resolution (Å)	ΔX	ΔY	ΔZ	Peak height (%)	2nd peak height (%)	Rank	σ
4420 Fab	15·0–4·0	0·420	—	0·160	100·0	58·5	1	7·13
4420 V_L-V_H	15·0–4·0	0·415	—	0·155	100·0	64·7	1	4·99
4420 C_L-C_H1	15·0–4·0	0·400	—	0·195	96·8		3	2·62
BV0401 Fab	15·0–4·7	0·425	—	0·165	100·0	53·2	1	7·82
	8·0–4·7	0·425	—	0·165	100·0	47·3	1	9·26
BV0401 V_L-V_H	15·0–4·0	0·420	—	0·165	100·0	83·5	1	3·56
	8·0–4·0	0·425	—	0·165	100·0	72·5	1	5·03
BV0401 C_L-C_H1	15·0–4·0	0·420	—	0·155	83·3		3	3·02
	8·0–4·0	0·425	—	0·165	95·0		2	4·31

Rotation searches (Crowther, 1972)

Model	Resolution (Å)	Euler angles α°	β°	γ°	Peak height (%)	2nd peak height (%)	Rank	σ
B. Results for NC174-NC6.8 Fab complex								
4420 Fab	15·0–4·0	47·5	74·0	225·0	100·0	93·8	1	4·31
4420 V_L-V_H	8·0–4·0	40·0	79·0	240·0	98·8		2	4·39
4420 C_L-C_H1	8·0–4·0	30·0	70·0	210·0	79·4		7	3·57
BV0401 Fab	15·0–4·0	42·5	70·0	200·0	86·5		6	3·40
BV0401 V_L-V_H	15·0–4·0	not observed						
BV0401 C_L-C_H1	15·0–4·0	40·0	71·0	200·0	100·0	79·7	1	4·69

Translation searches (RTMAP)†

Model	Resolution (Å)	ΔX	ΔY	ΔZ	Peak height‡	2nd peak height	Rank	σ
4420 V_L-V_H	15·0–4·0	0·17	0·34	0·24	0·299	0·212	1	11·2
BV0401 C_L-C_H1	15·0–4·0	0·12	0·29	0·08	0·215	0·171	1	8·0

Rotation searches (Crowther, 1972)

Model	Resolution (Å)	Euler angles α°	β°	γ°	Peak height (%)	2nd peak height (%)	Rank	σ
C. Confirmatory tests for orientation of native NC6.8								
Complex V_L-V_H	8·0–4·0	197·5	115·0	75·0	100·0	51·8	1	8·30
Complex C_L-C_H1	8·0–4·0	170·0	96·0	70·0	100·0	84·0	1	4·97
Complex V_L	8·0–4·0	200·0	116·0	75·0	100·0	85·3	1	4·45
Complex V_H	8·0–4·0	197·5	115·0	75·0	100·0	64·1	1	7·12
Complex C_L	15·0–4·0†	170·0	96·0	70·0	100·0	91·0	1	4·18
Complex C_H1	8·0–4·0	167·0	92·0	70·0	88·0		2	3·64

†8·4 Å search did not yield the correct solution at a level >65% of the maximum.

‡The translation searches were monitored by using correlation coefficients (Fan *et al.*, 1992).

(b) *Structural determinations by molecular replacement*

The results of the rotation and translation calculations are presented in Table 2A for the native protein and in Table 2B for the complex. Search models consisting of the intact Fabs and the V_L-V_H and C_L-C_H1 pairs of domains from the 4-4-20 and BV04-01 antibodies (Herron *et al.*, 1989, 1991) all gave consistent peaks in rotation function maps for the native NC6.8 Fab. In each instance, the orientation later found to be the correct choice was represented by the highest peak in the map. The Fab (peak 7·77σ above background) and C_L-C_H1 pair (4·22σ) from BV04-01 gave stronger signals than those from 4-4-20, while the latter provided a more effective V_L-V_H combination (5·37σ). In the translation searches with BV04-01 models, more convincing results were obtained when the data sets were truncated by the omission of reflections with *d* spacings of 8 to 15 Å.

The results were less straightforward when the same starting models were used to probe the structure of the NC174-NC6.8 Fab complex. Only the 4-4-20 Fab and the BV04-01 C_L-C_H1 pair produced the highest peaks in the rotation function maps, while the BV04-01 V_L-V_H pair failed to give a solution. In the translation searches, however, the 4-4-20 V_L-V_H pair and the BV04-01 C_L-C_H1 pair provided unequivocal results (11·2 and 8·0σ peaks respectively).

Because of the conformational changes and the sizeable decrease in the elbow bend angle on complexation of NC6.8 (see below), it was surprising that the whole Fab from 4-4-20 was a suitable starting model. Both 4-4-20 and BV04-01 have elbow bend angles greater than 170° (Herron *et al.*, 1989, 1991). As illustrated in Figure 1, the V_L domains of both 4-4-20 and BV04-01 closely resemble V_L of NC6.8 in amino acid sequences (3 and 7 differences among 108 residues; Bedzyk *et al.*, 1990; Herron *et al.*, 1991)). However, the V_H domains are quite dissimilar (63 and 72 interchanges). These heavy chain differences are obviously of considerable importance in the rotation function calculations for the complex. Consistent success with the BV04-01 C pair underscores the desirability of selecting a probe with a C_H1 domain

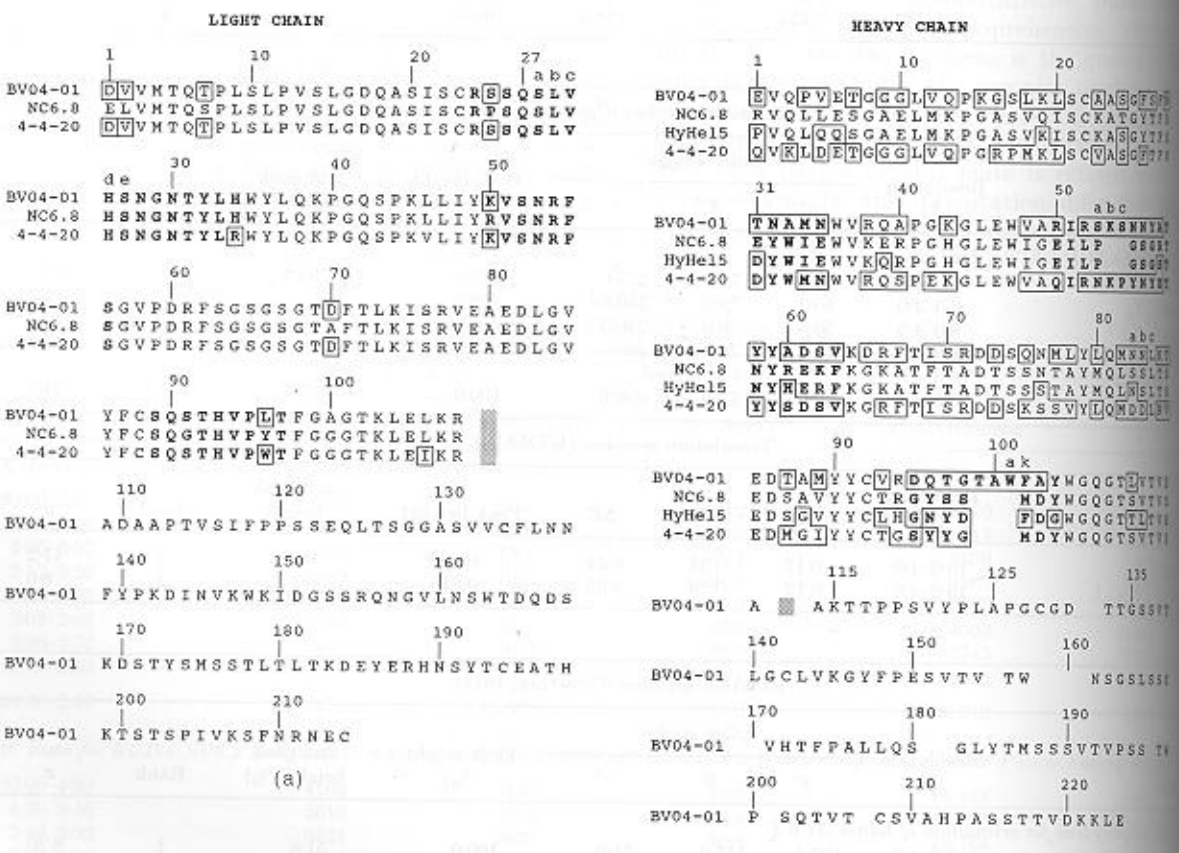

Figure 1. Amino acid sequences of the NC6.8 light chain (a) and heavy chain (b), as deduced from nucleic acid sequences by 2 of us (J.M.A. and D.S.L.). Residues are numbered as in Kabat *et al.* (1991). Constituents of the 3 CDRs are designated by boldface letters. The sequence of the V region of the light chain is sandwiched between the sequences of the κ chains of the murine antibodies BV04-01 (Herron *et al.*, 1991) and 4-4-20 (Bedzyk *et al.*, 1990). Residues different from those in NC6.8 are enclosed in boxes. Sequences of the C_L domains of the 3 proteins are assumed to be identical and only the sequence of BV04-01 is listed above. In the heavy chain V domains, the sequence of HyHEL5 (Sheriff *et al.*, 1987) is included because of its similarities to NC6.8. The latter is a member of subclass 2b and is therefore expected to have the same C_H1 sequence as BV04-01, which is cited below V_H.

Table 3
Progress of refinement

Resolution (Å), no. of reflections	*R*-factor	Stereochemistry	Comment
A. NC6.8 Fab native			
15-40, 5924	0·483 to 0·461		Rigid body refinement of 4 domains, V_L, V_H, C_L and C_H1
10-30, 9846	0·475 to 0·276	a = 0·027 b = 5·2 c = 29·5 d = 2·4	Model not including CDRs. Refinement using Jack-Levitt (positional) refinement.
10-30, 9846	0·276 to 0·255	a = 0·025 b = 4·9 c = 29·4 d = 2·1	Inclusion of CDRs. 1 cycle of molecular dynamics refinement with X-PLOR
10-26, 14,007	0·265 to 0·232	a = 0·021 b = 4·4 c = 29·7 d = 3·2	Three cycles of molecular dynamics, individual *B*-factors.
10-26, 14,007	0·232 to 0·226	a = 0·021 b = 4·5 c = 29·0 d = 2·0	Inclusion of 126 solvent molecules, 1 cycle of refinement
10-26, 13,063 ($F > 2\sigma(F)$)	0·218	a = 0·018 b = 2·6 c = 28·7 d = 2·3	Five cycles of model building and refinement to remove poor dihedral angles and improve stereochemistry.
B. NC174-NC6.8 Fab complex			
15-40, 4990	0·455 to 0·402		Rigid body refinement of 4 domains, V_L, V_H, C_L and C_H1.
10-27, 15,053	0·440 to 0·252	a = 0·021 b = 4·2 c = 29·3 d = 1·7	One cycle of molecular dynamics refinement with X-PLOR. Overall *B*-factor, no solvent, no ligand.
10-25, 18,242	0·261 to 0·209	a = 0·018 b = 3·8 c = 29·6 d = 2·4	Four cycles of molecular dynamics, individual *B*-factors, ligand and 125 solvent molecules.
10-22, 24,227	0·231 to 0·214	a = 0·016 b = 3·5 c = 28·9 d = 1·4	Three cycles of molecular dynamics, 296 solvent molecules in total.

The molecular dynamics refinements were carried out as described in the X-PLOR 3.0 manual (Brünger, 1992) using the standard protocol of heating to 3000 K and then slow cooling.

a, r.m.s. deviation in bond distance for all atoms (Å); b, r.m.s. deviation in bond angles for all atoms (°); c, r.m.s. deviation in dihedral angles for all atoms (°); d, r.m.s. deviation in improper angles for all atoms (°).

from the same subclass as the molecule under investigation (subclass 2b in this example).

As a cross-check of the molecular replacement results for the two disparate structures, the Fab model obtained for the complex was first split into its components (V and C pairs and four individual domains). To bring these components into coincidence with the native protein, each probe was subjected to rotation function calculations with the X-ray data for the native protein (see Table 2C). In five of the six cases the highest peak corresponded to the orientation of the comparable component in the unit cell of the native protein. For the sixth probe (C_H1 domain) the correct orientation was represented by the second highest peak. These assignments were monitored by displaying the models on the computer graphics screen. In summary the probes from 4-4-20, BV04-01 and the complex of NC6.8 with NC174 all give the same orientation for the unliganded form of NC6.8.

(c) *Structural refinements*

Outlines of the stages of refinement for the native and liganded forms are presented in Table 3. Cumulative *R*-factor values for different ranges of *d* spacings are giving in Table 4. Refinement statistics are summarized in Table 5.

For 13,063 reflections with *d* spacings of 10·0 to 2·6 Å the current *R*-factor value is 0·218 for the native protein. The root-mean-square (r.m.s.) deviations in bond lengths and bond angles are 0·018 Å

Table 4
R-*factor* versus *resolution*

NC6.8 Fab native			NC174-NC6.8 Fab complex		
Shell (Å)	No. of measurements	*R*-factor (cumulative)	Shell (Å)	No. of measurements	*R*-factor (cumulative)
10·00-5·00	1929	0·195	10·00–4·30	3789	0·165
5·00-4·06	1873	0·192	4·30–3·46	3610	0·169
4·06-3·57	1736	0·197	3·46–3·03	3521	0·177
3·57-3·26	1714	0·204	3·03–2·76	3262	0·187
3·26-3·03	1678	0·209	2·76–2·57	2925	0·197
3·03-2·86	1644	0·213	2·57–2·42	2634	0·204
2·86-2·72	1541	0·217	2·42–2·30	2414	0·209
2·72-2·60	948	0·218	2·30–2·20	2072	0·214

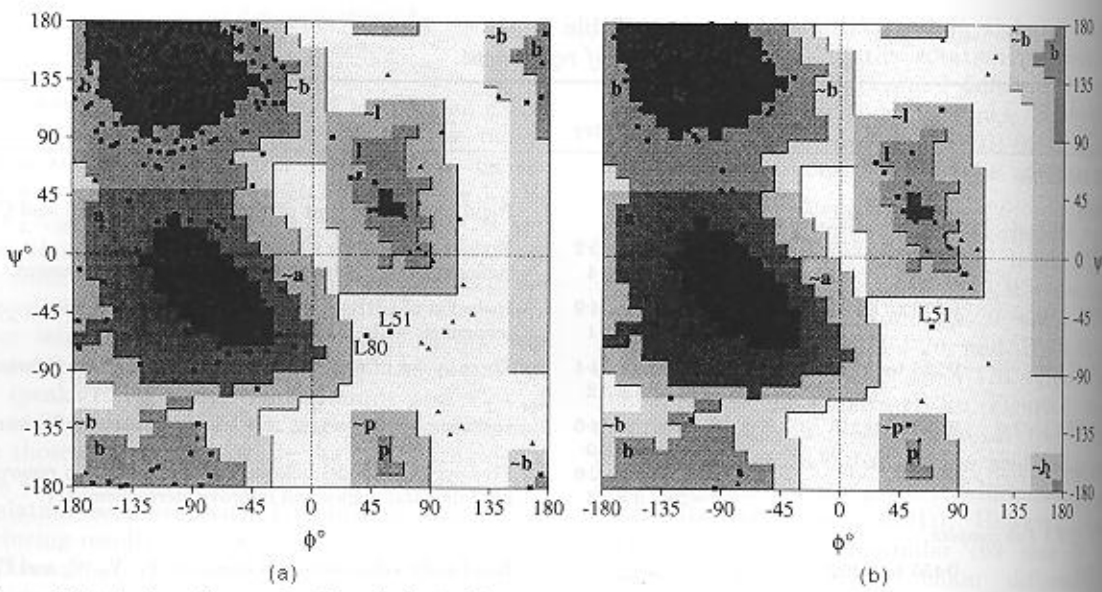

Figure 2. ϕ, ψ plots (Ramachandran & Sasisekharan, 1968) for the dihedral angles in the polypeptide chains of the native (a) and liganded (b) forms of the NC6.8 Fabs, as prepared with the program PAP (Laskowski *et al.*, 1993). Squares represent all residues except glycines, which are designated by triangles. Areas with permitted dihedral angles are shaded. The darkest shaded areas, A, B and L, correspond to the "most favored regions". Labels a, b, l and p designate "additional allowed regions" and ~a, ~b, ~l and ~p denote "generously allowed regions". Note that the clustering of values within prescribed boundaries is tighter for the ligand-Fab complex. This is a consequence both of better crystals and the use of higher resolution synchrotron data (see the text).

and 2·57°. In the complex of the N6.8 Fab with NC174 the *R*-factor value is 0·214 for 24,227 reflections with *d* spacings of 10·0 to 2·2 Å. The r.m.s. deviations in bond lengths and bond angles are 0·016 Å and 3·55°. Mean errors in coordinates (Luzatti, 1952) range from 0·30 to 0·35 Å in the native protein and 0·25 to 0·30 Å in the complex.

Plots of ϕ, ψ values for the main-chain dihedral angles in the native protein and the complex (Ramachandran & Sasisekharan, 1968; Laskowski *et al.*, 1993) are presented in Figure 2. Residues (other than glycine) with unfavorable dihedral angles after model building and refinement are identified by sequence numbers. To improve the quality of the protein structure, the stereochemical parameters and topological assignments of Engh & Huber (1991) were substituted for the original set in X-PLOR. In addition the weakest reflections ($F < 2\sigma(F)$) were removed from the data set for the native protein as this seemed to improve the overall quality of the map.

In the refinement of the ligand-Fab complex higher resolution data resulted in significantly higher quality of the stereochemistry from the

Table 5
Refinement data and model geometry

	NC6.8 Fab native	NC174-NC6.8 Fab complex
R-factor	0·218	0·214
No. of reflections	13,063	24,227
r.m.s. deviation in bond distance (Å)	0·018	0·016
r.m.s. deviation in bond angles (deg.)	2·57	3·55
r.m.s. deviation in dihedral angles (deg.)	28·77	29·05
r.m.s. deviation in improper angles (deg.)	2·32	1·40
Bond distance r.m.s. deviations > 0·06 Å	29	14
Bond angle r.m.s. deviations > 10·0°	45	86
Dihedral angles r.m.s. deviations > 90·0°	0	0
Improper angles r.m.s. deviations > 20·0°	0	0
Elbow angle	184	153
Average *B*-factor, overall (Å^2)	15·8	29·6
Average *B*-factor for V_L domain (Å^2)	13·7	30·0
Average *B*-factor for C_L domain (Å^2)	17·9	29·1
Average *B*-factor for V_H domain (Å^2)	11·2	25·6
Average *B*-factor for C_H1 domain (Å^2)	19·4	34·0
No. of solvent molecules	126	296
Mean errors in coordinates (Luzatti, 1952) (Å)	0·30–0·35	0·25–0·30

earliest stages (see right panel of Fig. 2). There was only one outlier with most angles clustered near or within the boundaries characteristic of β-pleated sheets or right and left-handed α-helices.

Average B factor values for the light and heavy chains were calculated to be 16·1 and 15·4 Å^2, respectively, in the native protein and 29·5 and 29·6 Å^2 in the complex. The apparent discrepancy in the two sets of values is probably attributable to different modes of data collection and scaling. The above values for the native protein are within the expected range for data collection with an area detector in this laboratory. There are also precedents for higher B factor values when data are collected and processed by methods similar to those employed for the complex. Thermal factors ranging from 22·53 to 32·62 Å^2 were reported for the V and C domains of Fab 8F5 (Tormo *et al.*, 1992), using data collected with synchrotron radiation and image plate recording. In the 8F5 and the NC6.8 Fabs, V_H has the lowest temperature factor of the four domains (see Table 5). C_H1 has the highest B factor in NC6.8 (both forms), but not in 8F5, whereas C_L has the largest value.

After simulated annealing with X-PLOR, residues L8, L95, L141, H149 and H200 were better fitted to F_o-F_c OMIT maps (Bhat & Cohen, 1984) as *cis* proline residues in both the native and liganded proteins. A tryptophan side-chain (H199) immediately adjacent to proline H200 is exposed to solvent but is also partially protected by the proline ring. Because of the ring puckering in proline the two residues form hydrophobic rather than π-π stacking interactions. Such interactions are possible only when the proline is in the *cis* configuration.

Three-dimensional "cage" electron density maps are presented in Figure 3 for the heavy chain loop containing all of CDR3 except tyrosine H102 (from the lower right the sequence is Thr-Arg-Gly-Tyr-Ser-Ser-Met-Asp; see Fig. 1). In the native protein (upper panel of Fig. 3) the cage density is continuous for all constituents except tyrosine H96. Several alternative starting models for this segment were subjected to refinement. In the most probable orientation (Fig. 3) electron density is observed for the phenolic ring of tyrosine H96 at the $1{\cdot}5\sigma(\rho)$ level. Only a small module of electron density ($0{\cdot}8\sigma(\rho)$) is present in the alternative upright position discussed below.

As shown in the lower panel of Figure 3, the cage density for tyrosine H96 is well defined in the complex at a $2{\cdot}5\sigma(\rho)$ level. Upon complexation the tyrosine side-chain is rotated and translated to a new stable location. If measured at the base of the phenolic ring the displacement is 4·6 Å. However, the phenolic hydroxyl group moves 10·5 Å. This new orientation is maintained by interactions with the diphenyl portion of the NC174 ligand (see below). Relative positions of neighboring residues did not change appreciably, although the CDR3 loop moved as a unit during complexation.

A schematic drawing of the ligand is shown in Figure 4 with its atoms labeled. The cage density for this compound is presented as a stereo diagram in Figure 5.

In summary, the polypeptide chains and the ligand are well represented by electron density. Model building was normally conducted with electron density $\rho > 2{\cdot}5\sigma(\rho)$. In three regions (residues L180 to L188 and H159 to H164 in the native protein and H126 to H133 in the liganded form) the electron density was weaker and model fitting was performed at $\rho = 2\sigma(\rho)$. Among the side-chains, residue H31 in the native protein, and serine L67 in the complex, are not represented by detectable electron density. In the complex, electron density is clearly visible for the H31 side-chain, which was identified as phenylalanine in the sequence determination. However, the electron density module shape is more characteristic of glutamic acid than phenylalanine. This sequence assignment is currently being reexamined.

(d) *Descriptions and comparisons of the structures*

RIBBONS models (Carson & Bugg, 1986; Carson, 1987) of the native and liganded proteins are shown in Figure 6. A schematic drawing of the secondary structure of the liganded form is presented in Figure 7. The secondary structures for the two variants are nearly identical and the liganded form is chosen for display because the measurements of the hydrogen bonding distances are based on higher resolution data.

For comparison of the structures, seven segments together containing 40 residues in each V domain and 35 residues in each C domain were selected from the most regular parts of the β-pleated sheets of each polypeptide chain. In the light chain these segments include residues 3 to 7, 22 to 26, 32 to 38, 63 to 67, 70 to 74, 85 to 91 and 99 to 104 (V domain); and 114 to 118, 133 to 137, 145 to 149, 158 to 163, 176 to 180, 193 to 197 and 205 to 209 (C domain). Corresponding residues in the heavy chain are numbered 3 to 7, 21 to 25, 33 to 39, 68 to 72, 77 to 81, 88 to 95 and 104 to 109 (V domain); 120 to 123, 141 to 145, 153 to 162, 169 to 174, 188 to 192, 206 to 211 and 219 to 223 (C domain).

The r.m.s. deviation in atomic positions is only 0·53 Å when the program INSIGHT (Biosym, 1989) is used to superimpose the 40 α-carbon atoms of the V_L domain of the native Fab on the corresponding atoms in the complex. For superposition of other domains, the r.m.s. deviations are 0·45 Å for V_H, 0·50 Å for C_L and 0·56 Å for C_H1. When the V_L-V_H pairs (80 atoms) or the C_L-C_H1 pairs (70 atoms) are superimposed, the r.m.s. deviations are 0·53 and 0·54 Å, respectively. Together with the hydrogen bonding patterns shown in Figure 7, these values add to the evidence that the pleated sheet regions are quite similar in the native and liganded proteins. Moreover, the domains tend to move as pairs, rather than singly.

Rotation of the α-carbon coordinates of V_L into V_H with the program ROTMOL (kindly supplied by

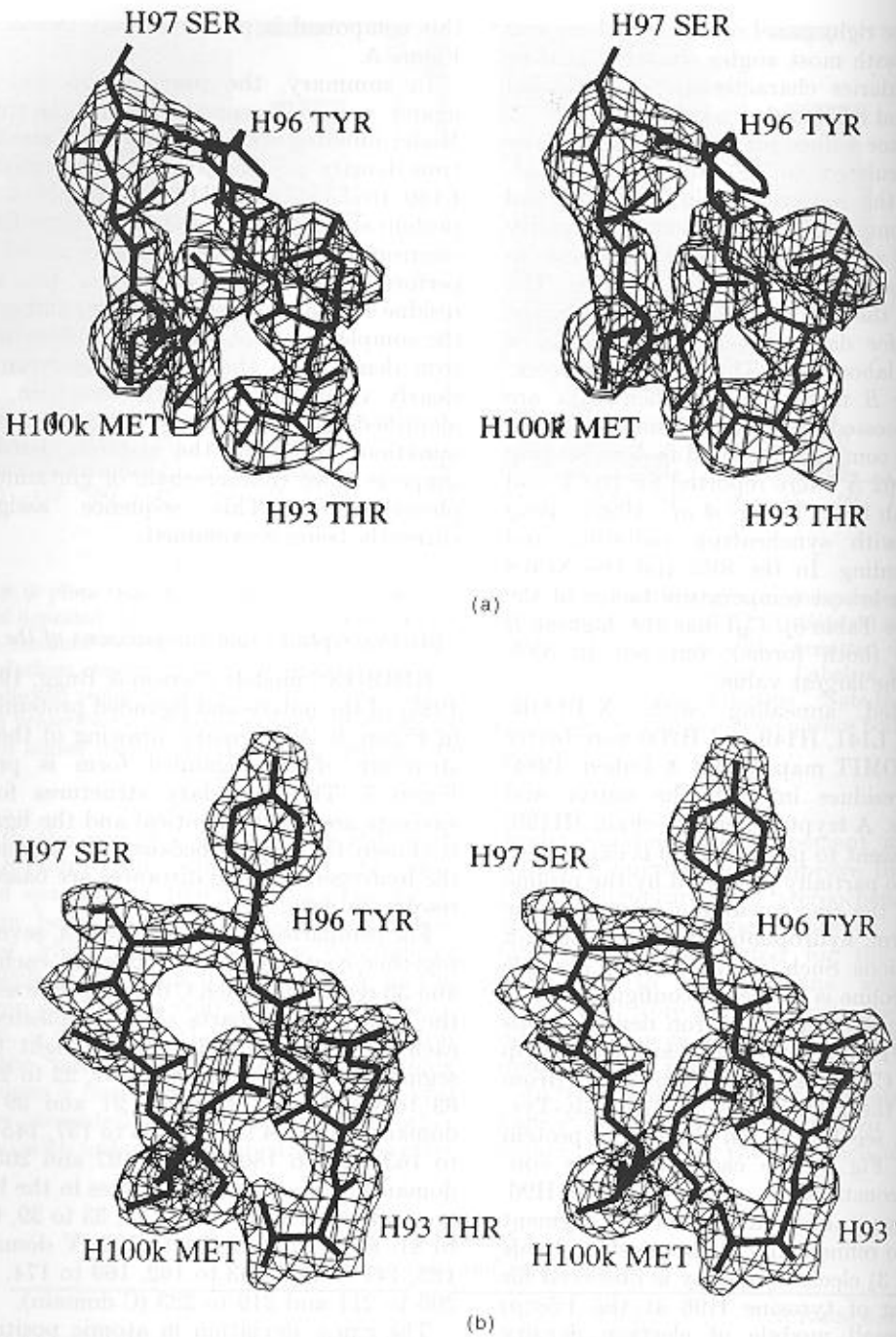

Figure 3. (a) Stereo diagram of a heavy chain segment fitted to 3-dimensional "cage" electron density for the native (unliganded) form of NC6.8. The sequence counterclockwise from the lower right is Thr-Arg-Gly-Tyr-Ser-Ser-Met-k which corresponds to the last 2 residues of FR3 and the CDR3 loop minus Tyr H102. The loop can be classified as a γ-turn, with a hydrogen bond (3·1 Å) between tyrosine H96 (backbone amide group) and serine H98 (carbonyl oxygen atom). Methionine H100k forms the floor of the binding site. (b) Same segment in the ligand-protein complex. On complexation the hydrogen bond is broken and the loop expands at the top. A new hydrogen bond is formed between the carbonyl oxygen atom of tyrosine H96 and the guanido group (N-15 in Fig. 4) of the ligand. The side-chain of tyrosine H96 moves from its position in the native protein to help lock the diphenyl moiety of the ligand into the active site of the complex. Arginine H94 and aspartic acid H101 form an ion pair (2·8 and 2·9 Å) in both the native and liganded proteins.

W. Steigemann and R. Huber) indicated that the two domains are related by an angle of 179° in the native protein and 176° in the complex (see Table 6). The pseudodyad (169°) relating the C_L and C_H1 domains remains the same after complexation. These measurements reflect only small rotations of individual V domains on complexation and no appreciable changes in the relative orientations of the domains.

By treating the domains as cylinders (Edmundson *et al.*, 1974; Herron *et al.*, 1991), we were able to estimate the angles subtended by the

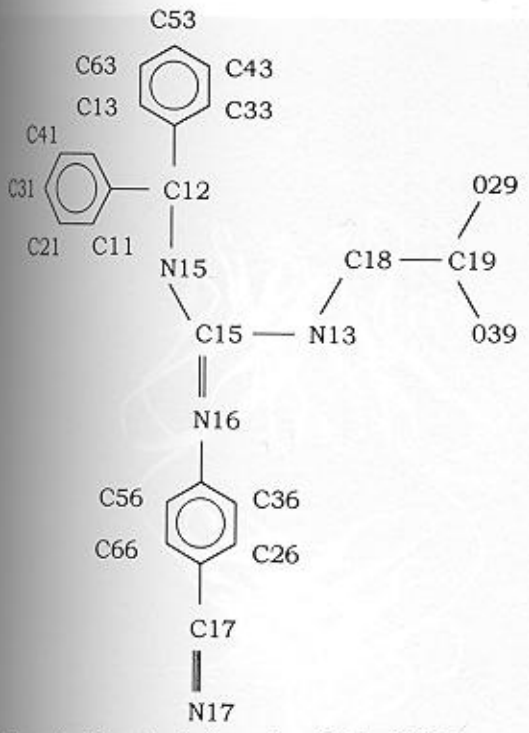

Figure 4. Structural formula of the NC174 sweet tasting compound, *N*-(*p*-cyanophenyl)-*N'*-(diphenylmethyl)-*N''*-carboxymethyl)guanidine. The atoms are labeled arbitrarily as in the crystallographic dictionary used for refinement of the structure. These labels will appear in references to atoms in the tables and in coordinate files to be made available to the scientific community.

long axes in pairs of domains. Cross-over angles of V_L and V_H were found to be 92° in the native protein and 91° in the complex. For C_L and C_H1 the corresponding angles were 109° and 108° (see Table 6). Again rigid body changes appear to be small in pairing of domains.

When other angles were considered, however, the native and liganded Fabs were significantly different. For example, the "elbow bend" angles between the V_L-V_H and C_L-C_H1 pseudodyads were calculated to be 184° and 153° for the two forms (see Table 6). The native form is the first Fab in our collection to have an elbow bend angle greater than 180°. In this conformation the angle between the long axes of V_L and C_L is 77°, whereas the corresponding angle in the heavy chain is 97°. Upon complexation the angle between V_L and C_L increases to 97° and the heavy chain angle decreases to 75°. There is thus a concerted and reciprocal change in angles within the light and heavy chains.

This type of relationship was first noted in the Mcg light chain dimer (Schiffer *et al.*, 1973; Edmundson *et al.*, 1975), where the angle between the V and C domains was flexed in monomer 1 and extended in monomer 2. Since the light chain in the Mcg IgG_1 immunoglobulin has the same sequence and assumes an extended structure (Rajan *et al.*, 1983; Guddat *et al.*, 1993), monomer 2 has been considered the light chain analog and monomer 1 the heavy chain analog in a dimer which mimics a typical Fab.

In the unliganded form of NC6.8 there is thus a reversal of roles where the heavy chain behaves like a light chain and *vice versa*. On complexation the angles revert in unison to their more familiar values. In the process the light chain is elongated by as much as 10 Å from the tip of the V domain to the distal end of the C domain and the heavy chain is shortened commensurately. This interconversion of roles can readily be appreciated by comparing the models in Figure 6.

Because of these unexpected reciprocal changes in the angles on complexation, we found it imperative to recheck the amino acid sequences against the electron density maps. This exercise convinced us that the light and heavy chains were correctly identified in the structural analyses of both crystal forms.

Most of these changes are mediated through the flexible "switch regions" between the V and C domains. As a consequence, extensive alterations

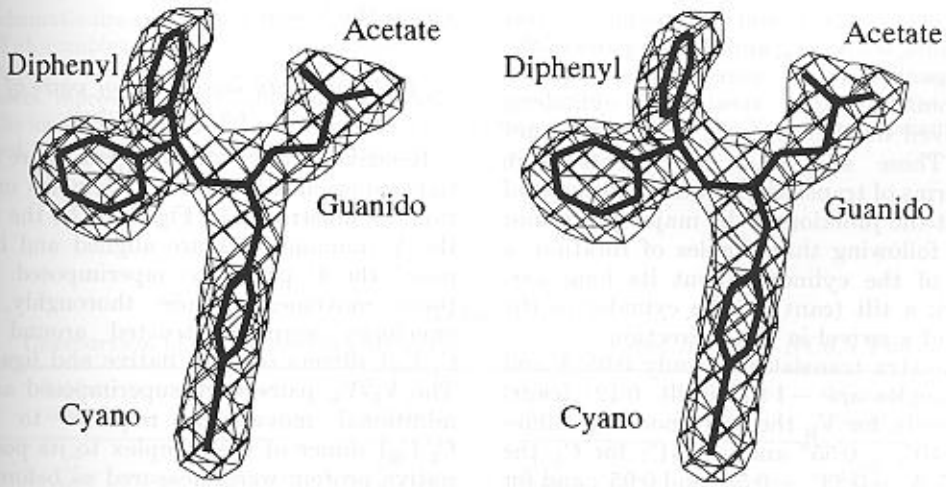

Figure 5. Stereo diagram of the cage electron density obtained for the NC174 ligand at 2·2 Å resolution. A skeletal model (see Fig. 4) is superimposed on this electron density.

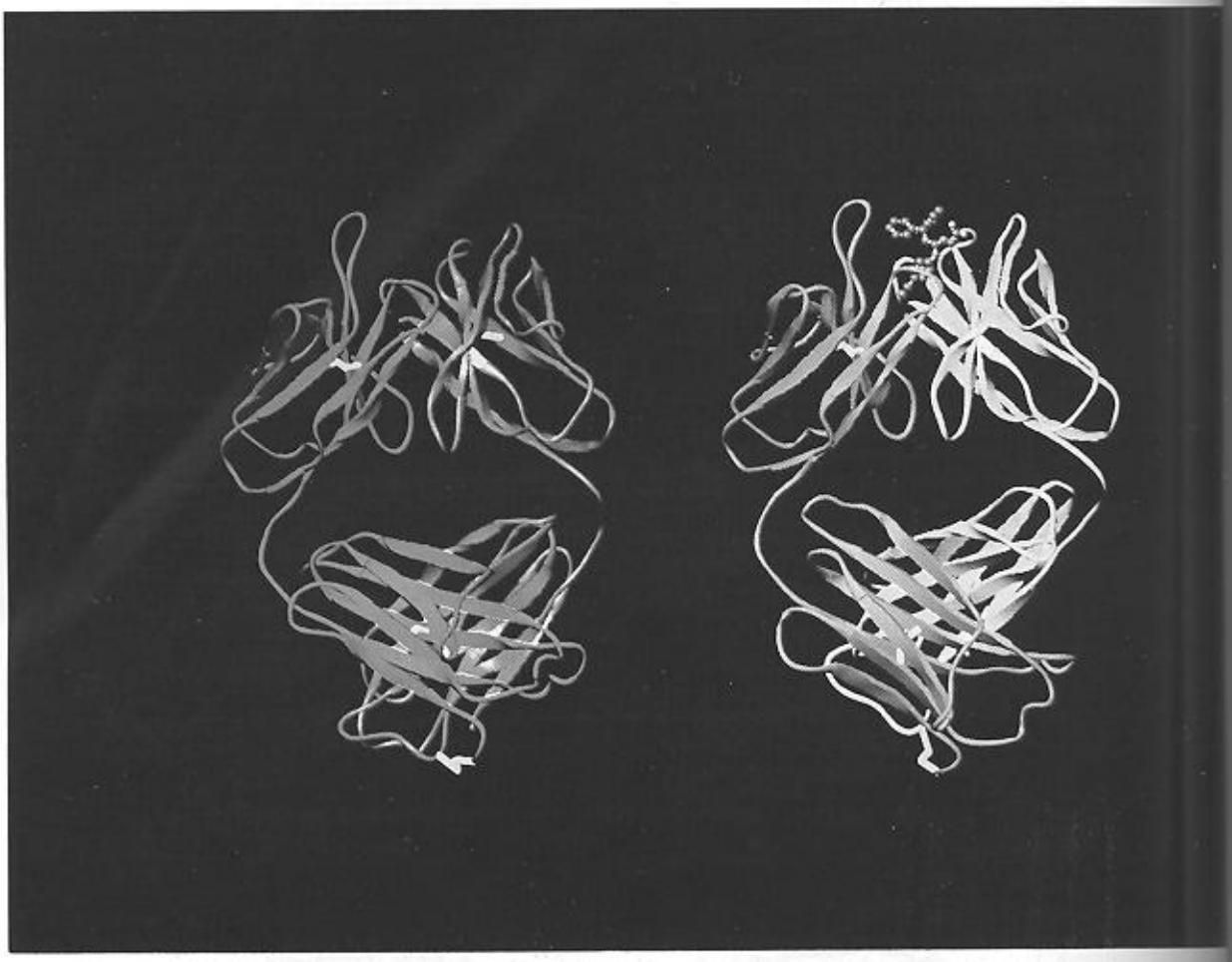

Figure 6. RIBBONS models (Carson & Bugg, 1986; Carson, 1987) of the native (left) and liganded (right) forms of the NC6.8 Fab. Light chains are colored crimson and magenta and heavy chains are green and yellow. Disulfide bonds are represented by heavy yellow tubes. The NC174 ligand (see Figs 4 and 5) is displayed as a ball-and-stick model. Ribbons correspond to segments of β-pleated sheets. This Figure was constructed by superimposing the V domains of the molecular variants and then focusing on the differences in the structures of the C domain pairs. Note that the light chain is more extended and the heavy chain is more flexed in the ligand-Fab complex (see Table 6 for reciprocal changes in the angles between the V and C domains on complexation).

should not be observed in the pairing interactions of the globular parts of the V and C domains. To test these impressions, the V_L-V_H and C_L-C_H1 pairs of the native and liganded forms were aligned in turn. Individual domains (again treated as cylinders) were then moved until they came into more exact coincidence. These additional movements were analyzed in terms of translations of the mid-point of the cylinder at the junction of the major and minor axes and the following three angles of rotation: a "barrel roll" of the cylinder about its long axis (x-axis vector); a tilt (cant) of the cylinder in the y-direction; and a swivel in the z-direction.

For V_L the extra translation is only 0·05 Å and the rotation angles are −1·84° (roll), 0·12° (cant) and 0·85° (swivel); for V_H the corresponding values are 0·05 Å, 0·40°, −0·55° and −0·31°; for C_L the values are 0·07 Å, −0·23°, −0·58° and 0·95°; and for C_H1 the values are 0·06 Å, 0·13°, −0·26° and 0·84°. As expected, these measurements indicate only small, if any, movements of individual domains within the V and C pairs.

(e) *Rigid body movements of pairs of domains on complexation*

Relative to the V domain pair, there are substantial displacements of the C domains on complexation, as illustrated in Figure 8. In the upper panel the V domain pairs are aligned and in the lower panel the C pairs are superimposed. To analyze these movements more thoroughly, cylindrical envelopes were constructed around the entire C_L-C_H1 dimers of both native and liganded forms. The V_L-V_H pairs were superimposed and then the additional movements required to restore the C_L-C_H1 dimer of the complex to its position in the native protein were measured as before. As a unit the C dimer is translated 10·9 Å, rotated −6° about the long axis (roll), tilted 35·3° (cant) and

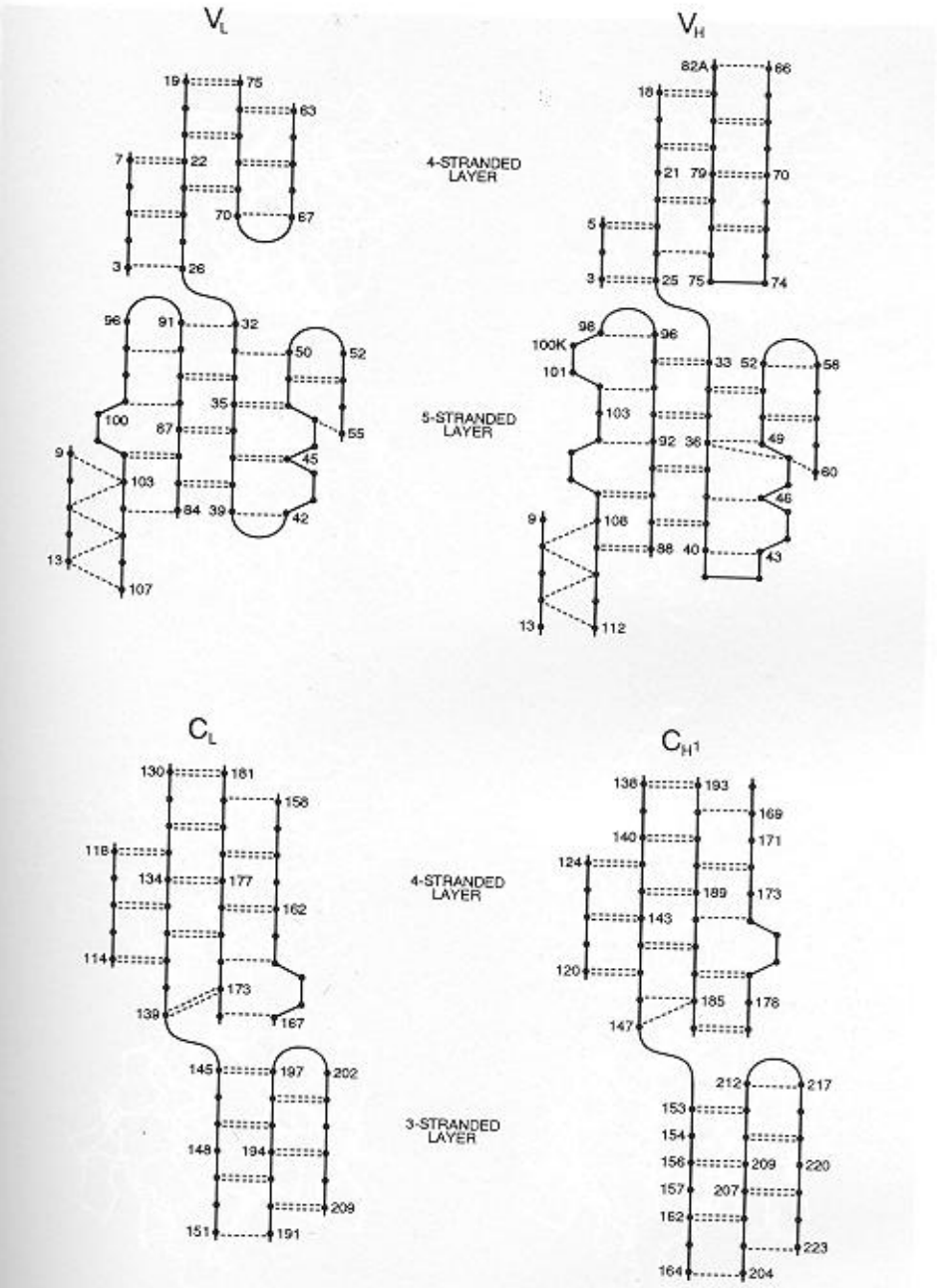

Figure 7. Schematic drawings of the hydrogen bonding patterns in the β-pleated sheets of the light and heavy chains in the ligand-Fab complex. The complex was chosen for presentation in preference to the native protein because the designated hydrogen bonds are based on X-ray data extending to higher resolution. In each domain there are 2 antiparallel sheets, represented here as 4 and 5-stranded layers in the V domains and 4 and 3-stranded sheets in the C domains. On the lower left of each V domain diagram note that residues 9 to 13 of the N-terminal segment are parallel to the last segment and form hydrogen bonds with its backbone atoms.

Table 6

Comparison of the quaternary structures of the native and liganded NC6.8 Fabs.

	Light chain	Heavy chain	V_L-V_H dimer	C_L-C_H1 dimer	V pseudo dyad	C pseudo dyad	Elbow angle
Native	77	97	92	109	179	169	184
Complex	97	75	91	108	176	169	153

All values are given in degrees.

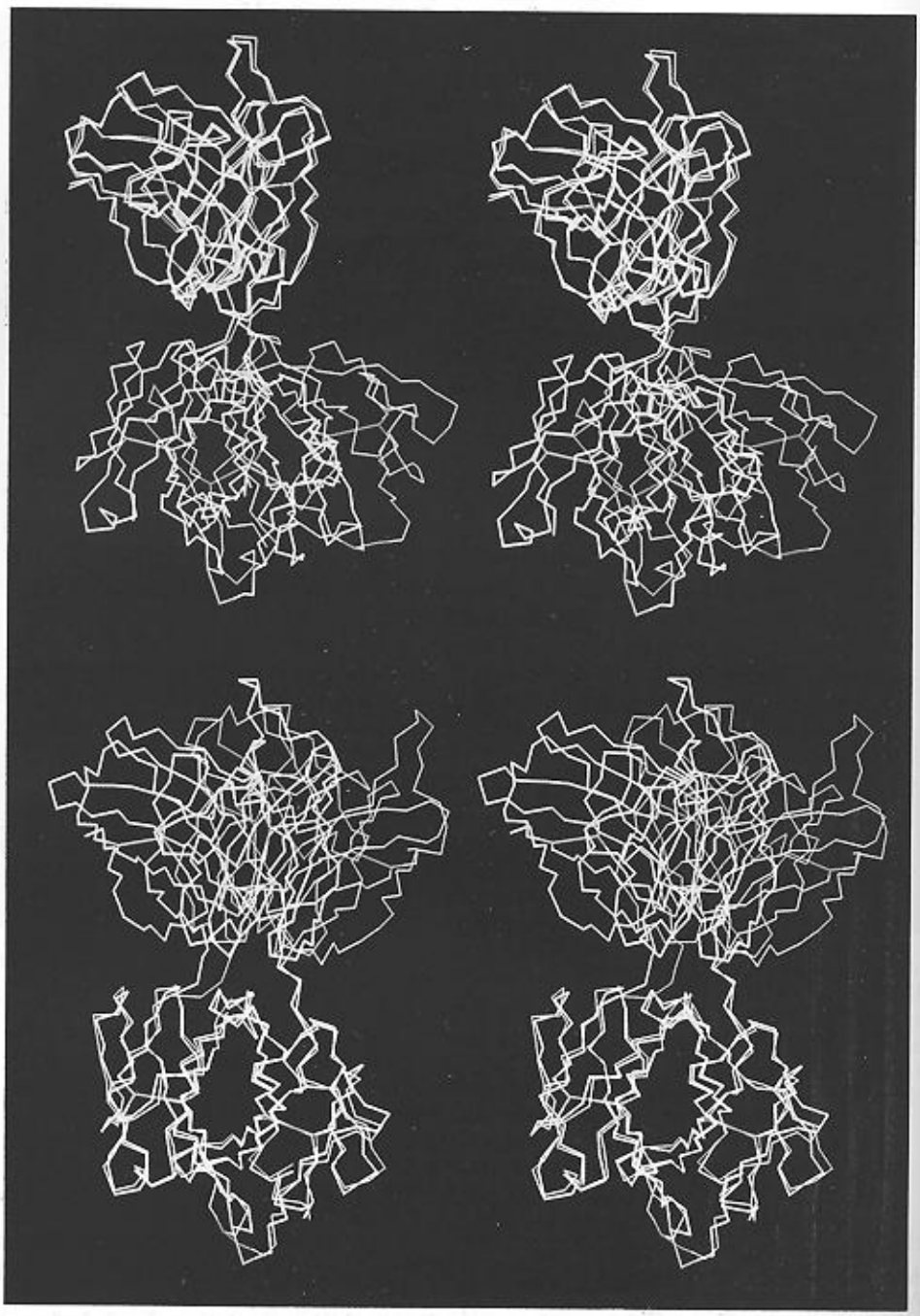

Figure 8. Comparisons of the C^{α} skeletal models of the native (purple) and liganded (yellow) forms of the N6.8 Fab. T help visualize the magnitude and directionality of the shifts of 1 form into the other, the disulfide bonds are highlighte in green in the native protein and in blue in the complex. Top, the V_L-V_H domain pairs are superimposed as closely a possible with the program INSIGHT. This view illustrates the maximum shifts in the C domains relative to the V domains on complexation. Centers of the intrachain disulfide bonds move 12·7 Å in the C_L domain and 13·6 Å in the C_H domain. The interchain disulfide at the distal end of the C pair shifts 25·2 Å. Bottom, superposition of the C_L-C_H domains. These tracings indicate that domains move mainly as paired units about the flexible "switch" regio connecting the V and C domains. Movements are analyzed in more detail in the text.

swiveled −5·0° on complexation. Thus, the 31° change in the elbow bend angle is attributable to a composite motion, of which a simple tilt is the largest component. These sizeable changes are coupled to the binding of a relatively small ligand in a relatively small active site. These interrelations will be considered in a later section, below.

(f) *Interactions between V domains in the native and liganded forms*

N-terminal glutamic acid of the light chain forms a salt bridge (3·0 Å) with arginine H60, but this is the only ion pair stabilizing the dimer in both the native and liganded forms. Overall, the V_L-V_H

contacts decrease in number from 127 to 104 on complexation. However, the losses are partly balanced by the ligand-protein interactions which bridge the light and heavy chains.

Qualitatively, the number of interdomain hydrogen bonds is roughly the same (7 *versus* 8) in the two forms. CDR1 and framework 2 (FR2) of the light chain are kept in close proximity with CDR3 of the heavy chain by hydrogen bonds between the side-chains of histidine L34 and tyrosine L36 and the main-chain atoms of serine H97, serine H98 and methionine H100k. Tyrosine L36 also makes five contacts with the side-chain of tryptophan H103 at the end of CDR3. This ladder of interactions between parallel segments is completed in the liganded form (but not in the native protein) by a hydrogen bond between the amide group ($N^{\epsilon 2}$) of glutamine L38 and a carboxyl oxygen ($O^{\epsilon 2}$) of glutamic acid H39 (both members of FR2). A similar interaction between the two glutamine 40 side-chains in the Mcg light chain dimer seals the deep binding pocket and also marks the distal end of stabilizing contacts along the V_1-V_2 interface (Edmundson *et al.*, 1974).

At the location equivalent to the end of the deep pocket in Mcg, the NC6.8 light chain turns and becomes anti-parallel to CDR3 and FR4 of the heavy chain. Side-chains of serine L43 and lysine L45 are within hydrogen bonding distances of backbone atoms at glycine H104 and aspartic acid H101 in the native protein, and serine L43 is also close to glutamine H105 in the complex. Serine L43 (2 contacts) and proline L44 (10 contacts) both touch tryptophan H103 while leucine L46 (5 contacts) impinges on methionine H100k and aspartic acid H101. In turn, backbone atoms of glutamine L42 and serine L43, as well as the serine side-chain, make six contacts with the side-chain of tyrosine H91 (end of FR3).

CDR2 of the light chain in the native protein is associated with CDR3 of the heavy chain by interactions of tyrosine L49 (1 contact) and arginine L50 (3 contacts) with the backbone and side-chain of serine H97. Conformational changes in the complex move these atoms out of range for interactions. At the distal end of the CDR2 loop in both forms, however, phenylalanine L55 interacts with the side-chains of both aspartic acid H101 (1 contact and tyrosine H102 (4 contacts).

CDR3 of the light chain is in proximity to parts of CDR2 of the heavy chain. Valine L94 interacts with asparagine H58 (1 contact) and proline L95 makes incidental contacts (2) with arginine H60 as it passes to meet the N-terminal glutamic acid of the light chain. Tyrosine L96 makes three contacts with glutamic acid H35 (CDR1) and six contacts with glutamic acid H50 (CDR2), including a hydrogen bond (3·1 Å) with a carboxylic oxygen atom.

In the most hydrophobic portion of the V_L-V_H interface, phenylalanine L87 (FR3) and CDR3 constituents valine L94, proline L95, tyrosine L96 and phenylalanine L98 are packed against leucine H45 and tryptophan H47 of FR2. In the native protein, tyrosine L96 (32 contacts) and tryptophan H47 (30 contacts) account for the greatest number of interatomic contacts between the domains. These numbers decrease to 19 and 14 in the complex and reflect subtle displacements of side-chains along the articulating surfaces.

(g) *Interactions between the C domains in the native and liganded forms*

Similar trends are seen in comparisons of the C pairs in the native and liganded forms. The number of interactions between domains decreases from 112 to 81 on complexation. It is convenient to consider these interactions in terms of the four-stranded β-pleated sheets through which they are mediated (Edmundson *et al.*, 1975).

The first strand (4-1) of this sheet encompasses residues L110 to L119 in the light chain and H116 to H125 in the heavy chain (see Fig. 7 for the residues comprising the other 3 strands). Strand 4-1 immediately follows the switch region, where conformational changes on complexation are maximal. It is therefore not surprising that the most accentuated differences in interdomain contacts involve the 4-1 strands. For example, the side-chain of phenylalanine L118 makes a total of 31 contacts with backbone atoms (proline H123, 2; leucine H124, 7; alanine H125, 12; and proline H126, 3) and the side-chain of leucine H124 (7) in the native protein but only a total of nine contacts in the complex. If considered from the view of the heavy chain, tyrosine H122, proline H123, leucine H124 and alanine H125 of strand 4-1 are collectively in contact with 42 light chain atoms in the native protein and only 20 atoms in the complex. The 4-1 strands do not appear to be dissociated from the pleated sheets seen in the native protein, since there are double hydrogen bonds between adjacent anti-parallel segments at key positions L118 and H124 in the complex (see Fig. 7). Phenylalanine L118 and leucine H124 are equivalent residues in the light and heavy chains, and as indicated above are major participants in interdomain interactions. We conclude that the 4-1 strands in C_L and C_H1 are separated by concerted movements of the whole domains when the protein binds ligand.

In the helical segment following strand 4-1 in the light chain, glutamic acid L123 and glutamine L124 participate in 18 interactions with tyrosine H122 (9), proline H123 (2) and lysine H221 (7) in the native protein and only six in the complex. Both the hydrogen bond (3·2 Å) between the hydroxyl group of serine L121 and the carbonyl oxygen atom of proline H123 and the ion pair (3·3 Å) of glutamic acid L123 and lysine H221 are disrupted during complexation.

In the heavy chain, the segment following strand 4-1 contains the half-cystine residue (H128) forming a disulfide bridge with residue L214. This region is quite different in the native and liganded proteins. In addition to shifts in the position and orientation of the disulfide bond itself, the sulfur atom in the

light chain of the complex makes two new contacts with the backbone at glycine H129. The α-carboxyl group of cystine L214 closely approaches the C^{β} and S^{γ} of cystine H128 (5 contacts).

It is interesting that the six interdomain hydrogen bonds involving strands other than 4-1 are maintained in the transition from the native protein to the complex. The van der Waals interactions are also generally conserved. For example, interdomain contacts total 23 (native) and 19 (liganded) for the 4-2 strands, 17 and 19 for the 4-3 strands and 16 and 14 for the 4-4 strands. As examples of the types of residues engaged in such interactions, serine L131, valine L133, phenylalanine L135 and asparagine L137 of strand 4-2 interact with the side-chains of leucine H143 (4-2), leucine H124 (4-1), threonine H139 (4-2), serine residues H188, H189 and H190 (4-3) and histidine H172 (4-4). In contrast to the V domains, many interface residues have polar side-chains in the C domains.

(h) *Structural features of the ligand and active site*

The orientation assumed by the ligand in the active site of the protein is shown in stereo in Figure 9 (top panel). A comparison of the orientations of these active site constituents before and after the addition of the ligand is presented in the lower panel of Figure 9. Contacts between ligand and protein constituents are listed in Table 7.

Estimates of the total surface area of the ligand ((Lee & Richards, 1971) in the X-PLOR package) vary with the choice of the probe radius but average about 700 Å^2 (663, 706 or 730 Å^2 for probe radii of 1·4, 1·6 and 1·7 Å). When the ligand is bound to the protein, the solvent accessible surface area is calculated to be 126, 122 or 117 Å^2; i.e. 81 to 84% (537, 584 or 613 Å^2) of the surface is "buried". Similar calculations for the protein with the three probe radii give values of 260, 233 or 218 Å^2 for buried surface areas. Light chain residues account for 31% of the total and include histidine L27d, asparagine L28 and tyrosine L32 of CDR1 and glycine L91, threonine L92, valine L94 and tyrosine L96 of CDR3. The remaining 69% of the buried surfaces involve the heavy chain constituents glutamic acid H31, tryptophan H33 and glutamic acid H35 of CDR1, glutamic acid H50, leucine H52, arginine H56, and asparagine H58 of CDR2 and glycine H95, tyrosine H96, serine H97, serine H98 and methionine H100k of CDR3.

The discrepancy between the buried surface areas in the ligand and protein has a simple explanation. In the concave binding site the ligand is enclosed by protein side-chains which are still partly accessible to solvent after complexation.

NC174 is inserted into a relatively narrow fissure with its cyanophenyl group anchored at the bottom of the cleft and its diphenyl rings protruding from the top. The cyanide group is wedged between the light and heavy chain backbone segments at glycine L91 and serine H97 and is stabilized by hydrogen bonds (2·8 and 3·2 Å) with two water molecules (see Fig. 9). Methionine H100k seals the bottom of the fissure and makes two van der Waals contacts with the cyanide nitrogen atom. On one side the cyano phenyl ring is stacked against tyrosine L96 (7 contacts) but the rings are not quite parallel. On the other side the cyanophenyl group is bounded by backbone atoms of tyrosine H96, serine H97 and serine H98 (total of 18 contacts) and the side-chains of glutamic acids H35 and H50 (6 contacts).

The positive charge on the ligand is partially neutralized by the formation of an ion pair (3·3 Å) between the aryl nitrogen atom (which is also a substituent of the guanido group) and the carboxyl group of glutamic acid H50. This positive charge is further delocalized by hydrogen bonding (2·9 Å) between the carbonyl oxygen atom of tyrosine H96 and the protonated guanido nitrogen atom which is linked to the biphenyl group.

Components of a continuous ligand segment including the aryl nitrogen atom, the carbon atom and the third nitrogen atom of the guanido group and the carbon atoms of the acetate moiety, are tightly packed against tryptophan H33 (total of 11 contacts). Further immobilization of the acetate occurs through multiple interactions (13) with the side-chains of arginine H56 and asparagine H58. One of the carboxyl oxygen atoms of acetate is within ion pairing or hydrogen bonding distances (3·0 and 2·8 Å) of both a guanido nitrogen atom ($N^{\eta 1}$) of arginine H56 and the amide group of asparagine H58, while the second carboxyl oxygen atom is paired (3·0 Å) with the guanido nitrogen atom $N^{\eta 2}$.

The biphenyl rings are rather loosely associated with the protein. One ring interacts primarily with

Table 7

Protein residues that are within hydrogen bond distance or van der Waals contact of the NC174

	NC6.8 Fab	
NC174	Hydrogen bond or ion pair† (atomic distance Å)	van der Waals
Cyanophenyl		Gly L91, Tyr L96, Glu L35, Glu H35, Tyr H96, Ser H97, Ser H98
Phenyl ring 1		His L27d, Tyr L32, Tyr H96, Ser H97
Phenyl ring 2		Trp H33, Tyr H96
Guanidine group	N15-H96, Tyr O, 2·9 N16-H50 Glu $O^{\epsilon 2}$, 3·2	Glu H50, Trp H33, Tyr H96
Carboxyl	O29-H56 Arg $N^{\eta 1}$†, 3·0 O29-H58 Asn $N^{\delta 2}$, 2·8 O39-H56 Arg $N^{\eta 2}$†, 3·0	Trp H33, Arg H56, Asn H58

Maximum contact distances are C-C 4·1 Å; C-N 3·8 Å; C-O 3·7 Å; N-N 3·4 Å; N-O 3·4 Å; O-O 3·3 Å. Residue numbers according to Kabat *et al.*, 1991.

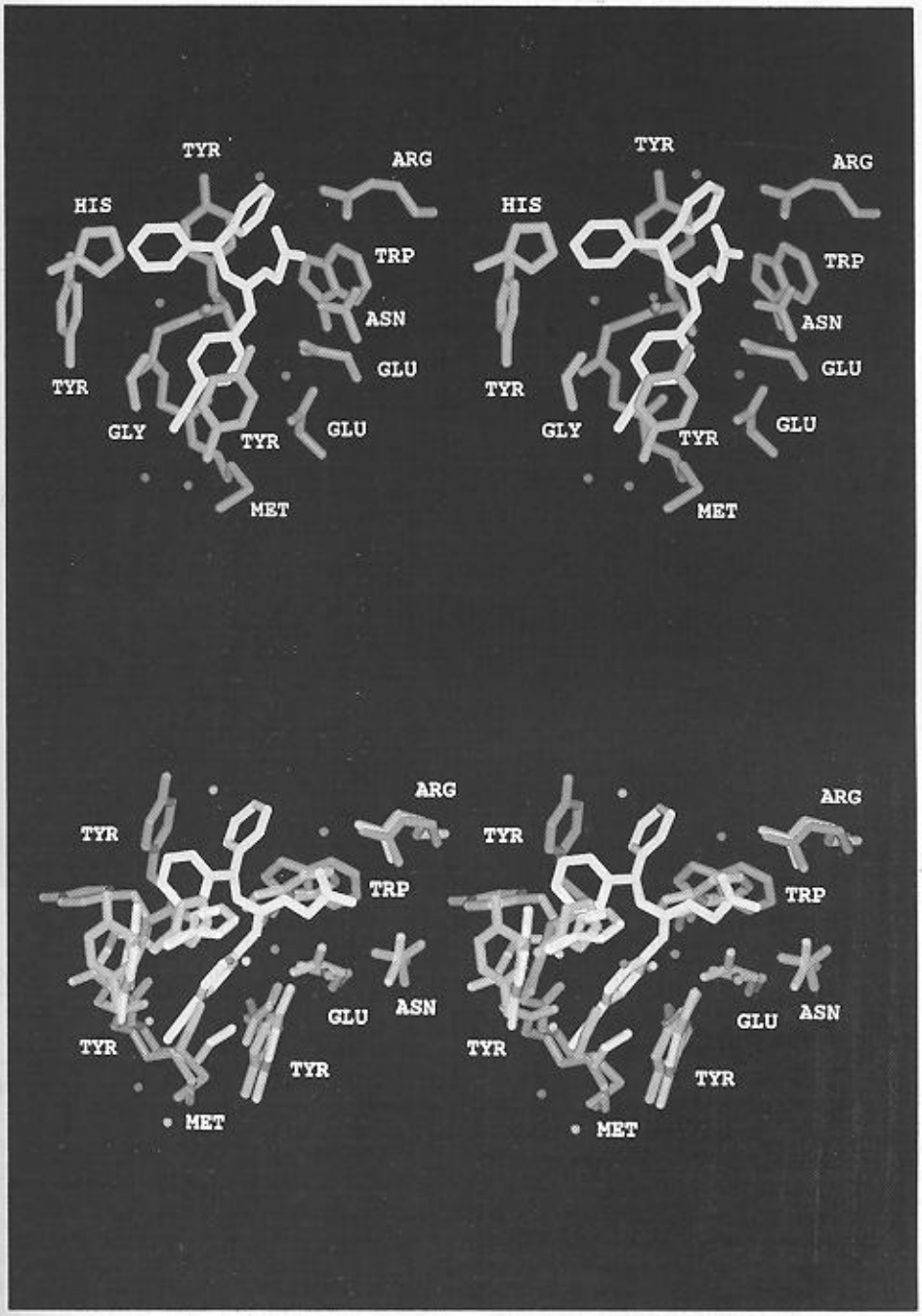

Figure 9. Top, stereo diagram of the NC174 ligand in the binding site. The light chain constituents are colored blue, the heavy chain components green, the ligand yellow and the ordered water water molecules red. Light chain contacts, in counterclockwise order, are supplied by His L27d and Tyr L32 of CDR1 and Gly L91 and Tyr L96 of CDR3. Clockwise, the heavy chain constituents are Tyr H96 of CDR3, Arg H56 of CDR2, Trp H33 of CDR1, Asn H58 and Glu H50 of CDR2, Glu H35 of CDR1, Met H100K of CDR3 and the backbone atoms of Ser H98 and Ser H97 of CDR3. The diphenyl rings of the ligand are located at the top of the crevice just in front of Tyr H96. Below the right phenyl ring is the acetic acid moiety which is neutralized by Arg H56 and Asn H58. The planar Y-shaped structure in the middle of the ligand is the positively charged guanido group, which is neutralized by Glu H50 and Tyr H96. At the lower left the cyanophenyl group is situated near Tyr L96 and the cyano group is wedged between the backbone segments of the light and heavy chains at Gly L91 and Ser H98. Two water molecules involved in the binding process (see the text) are located below the cyano group. Bottom, a slightly different view to illustrate the conformational changes in the active site on complexation. Constituents of the native Fab and its accompanying water molecules are colored magenta. Shifts in the light chain are relatively minor except for the backbone displacement at Gly L91. In the heavy chain there is a sterically dictated shift of Trp H33 and major adjustments of CDR3, including the swing of Tyr H96 to a position supporting the diphenyl rings of the ligand (also see Fig. 3).

Table 8
Protein-solvent-ligand networks in the NC174-NC6.8 Fab complex

Protein-water distance (Å)	Protein ------ solvent ---- ligand network			Ligand water distance (Å)	*B* factor (Å²) of solvent atom
2·8 H-bond	L89 O^{γ}	W14 O	N17	3·2 H-bond	52·9
2·5 H-bond	H35 $O^{\epsilon 2}$	W27 O	C66	3·1 VDW	51·2
3·0 H-bond	L96 O^{η}	W47 O	C15	3·3 VDW	56·0
3·3 H-bond	H56 $N^{\eta 2}$	W191 O	O39	2·8 H-bond	43·0
2·8 H-bond	L36 O^{η}	W192 O	N17	2·8 H-bond	23·2
2·8 H-bond	L89 O^{γ}	W192 O	N17		
2·9 H-bond	L91 O	W250 O	C51	3·6 VDW	19·3
3·0 H-bond	L92 O	W250 O	C51		
	None	W271 O	C61	3·6 VDW	47·7

Ligand atoms are numbered as in Figure 4.

the light chain (histidine L27d, 1 contact, and tyrosine L32, 5 contacts), although the side-chain of serine H97 makes two contacts with the distal side of the ring. The second ring makes contact only with tryptophan H33 (5) and tyrosine H96 (4).

There is a very attractive explanation for tyrosine H96 movements, which are coupled to the rotation and expansion of the CDR3 loop (see Figs 3 and 9). During complexation the backbone carbonyl oxygen atom of H96 is displaced 2·1 Å toward the ligand to make a hydrogen bond (2·9 Å) with a guanido nitrogen atom (N-15; see Table 7). Simultaneously, the tyrosine side-chain is recruited and stabilized by the interactions with the diphenyl group. Finally, the protein backbone at serine H97 is also directed toward the ligand and the carbonyl group makes a weak stacking interaction with the cyanophenyl ring (see Rowland *et al.*, 1990; Thomas *et al.*, 1982). These conformational shifts are consistent with an induced fit mechanism of binding (Edmundson *et al.*, 1974, 1984, 1987; Edmundson & Ely, 1985; Colman, 1988; Bhat *et al.*, 1990; Herron *et al.*, 1991; Rini *et al.*, 1992).

(i) *Solvent molecules in the active site*

Seven ordered water molecules have been detected within 4 Å of the ligand in the active site (see Table 8). They are also sufficiently close to the protein to make eight hydrogen bonds with oxygen or nitrogen atoms. Six of these molecules are visible in Figure 9. Two are within hydrogen bonding distances (3·2 and 2·8 Å) of the cyanide nitrogen atom (N-17). Another is hydrogen-bonded (2·8 Å) to a carboxyl oxygen atom (O-39) on the acetic acid moiety. Of the remaining three, one is between the cyanophenyl ring and glutamic acid H35, one is near the guanido carbon (C-15) and one is close to a diphenyl ring (C-51). These observations suggest that water molecules are included (rather than excluded) in the binding process. By bridging the ligand and protein, they provide part of the local articulating surfaces.

(j) *General distribution of ordered solvent molecules around the proteins*

The locations of ordered water molecules in the structure of the complex are shown in Figure 9. Solvent molecules forming hydrogen bonds (<3·2 Å) with protein atoms are listed in Table 8. The general distribution of water molecules among the four domains is summarized in Table 9B. Few assignments are made for the native protein because of the limited resolution (2·6 Å) of the diffraction data, but those selected follow the overall pattern observed in the complex.

Table 9
Assignments of water molecules

Protein atoms	Complex NC174-Fab NC6.8	Native Fab NC6.8
A. *Number of hydrogen bonds (distances of 3·2 Å or less) between protein and solvent atoms*		
(i) Main chain		
C=O	92	43
NH	56	27
(ii) Side chain		
Ser O^{γ}	31	9
Thr $O^{\gamma 1}$	13	1
Tyr OH	11	5
Glu $O^{\epsilon 1}$, $O^{\epsilon 2}$	15	6
Asp $O^{\delta 1}$, $O^{\delta 2}$	13	7
Gln $O^{\epsilon 1}$, $N^{\epsilon 2}$	17	4
Asn $O^{\delta 1}$, $N^{\delta 2}$	10	1
Trp $N^{\epsilon 1}$	2	0
His $N^{\delta 1}$, $N^{\epsilon 2}$	4	5
Lys N^{ζ}	5	5
Arg $N^{\eta 1}$, $N^{\eta 2}$	7	3
Total	276	116
Domain	Complex	Native
B. *Distribution of solvent molecules within 5 Å of each domain*		
V_L	71	30
V_H	88	37
C_L	74	33
C_H1	63	26
Total	296	126

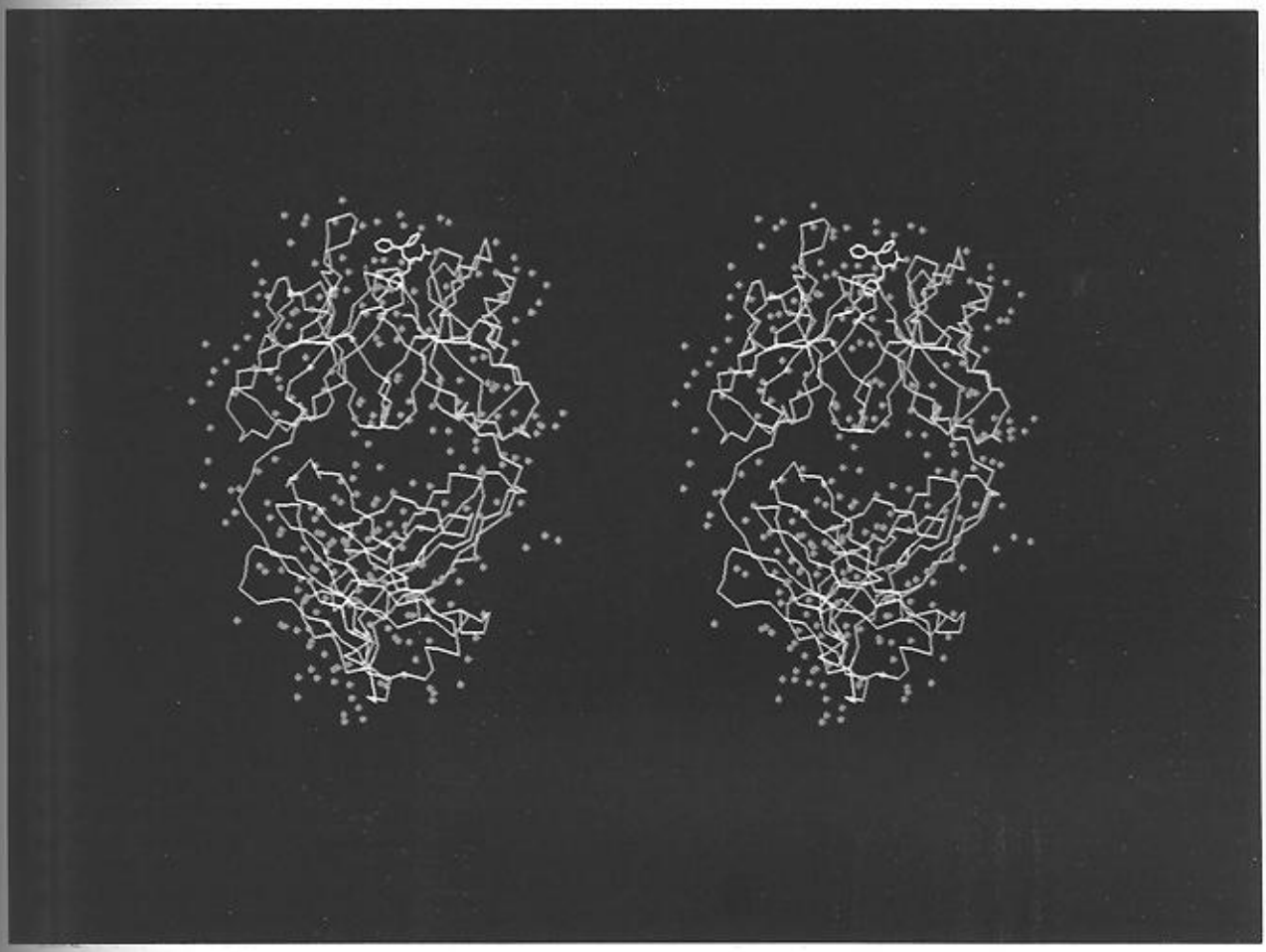

Figure 10. Stereo diagram of the hydrated C^{α} skeletal model of the ligand-Fab complex. The light chain is colored magenta, the heavy chain cyan and the ligand yellow. Ordered water molecules are represented by green dots.

In the latter, 33% of the hydrogen bonds between water and protein involve main-chain carbonyl oxygen atoms. Main-chain amide groups account for 20% and the remainder are distributed among side-chain oxygen and nitrogen atoms. In some cases water molecules are within hydrogen bonding distance of two protein atoms, particularly adjacent carbonyl and amide constituents in the main-chain. There are a few examples in which three to five protein and solvent atoms are in close contact with one water molecule.

The overall distribution of solvent molecules is not significantly different from those in other immunoglobulin fragments examined in this laboratory (Ely *et al.*, 1989; Herron *et al.*, 1989, 1991; Fan *et al.*, 1992). However, there are several structural features worthy of comment. As expected, ordered water molecules tend to accumulate in regions where there are relatively high concentrations of polar residues: e.g. CDR2 of the heavy chain (upper right hand loop in Fig. 10) and the distal ends of the C domains (surfaces at the bottom of Fig. 10). At least 13 ordered water molecules can be identified in the space between V_H and C_H1 (middle right in Fig. 10). A similar pattern of organized solvent was first noted in the space between the V and C domains of the tightly flexed heavy chain analog in the orthorhombic form of the Mcg light chain dimer (Ely *et al.*, 1989). Such clustering is not observed in the V-C interface of the light chain analog of the Mcg dimer, the light chain in the native or liganded NC6·8 Fab (middle left of Fig. 10) or in the heavy chain of the native NC6.8 Fab. The organization of bulk solvent in this region thus appears to be dependent on reduction of the available space by the flexing of the V and C domains.

(k) *Crystal packing diagrams for the native and liganded forms*

Packing diagrams are presented in stereo in Figures 11 and 12, with the native protein displayed in the top panel in each case. In Figure 11 the four Fabs in each unit cell are shown as color-coded C^{α} skeletal models. Two adjacent Fabs related by a unit cell translation along the shortest dimension of the cell (51·4 Å along *b* in the native Fab and 37·3 Å along *c* in the complex) are displayed in Figure 12.

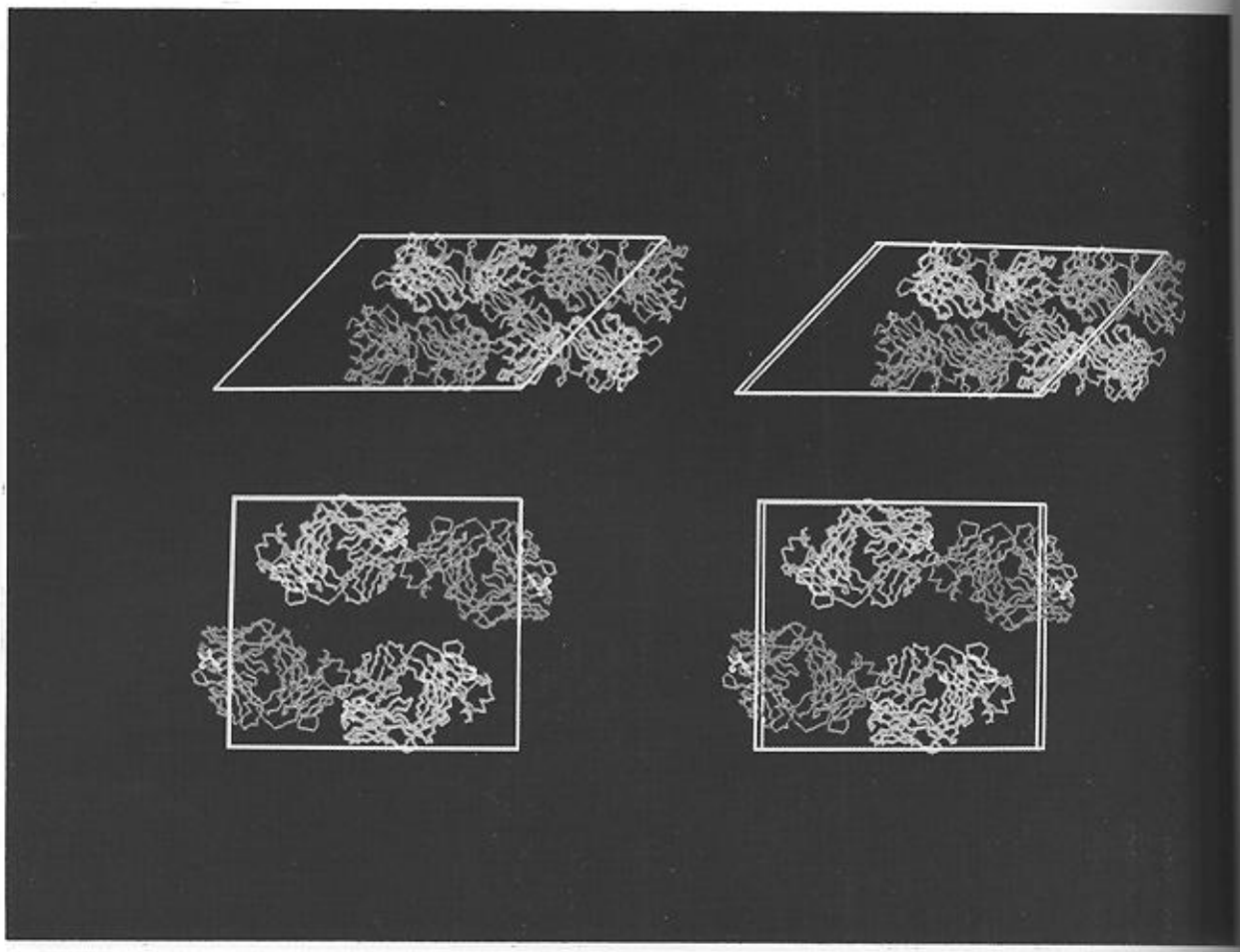

Figure 11. Comparison of the crystal packing of the native and liganded forms of the NC6.8 Fab. Top, stereo diagram of 4 native protein molecules. One monoclinic unit cell is outlined in yellow, the the x-axis horizontal and the z-axis intersecting it at an angle (β) of 132·8°. In 2 of the molecules the light chains are colored magenta and the heavy chains cyan. In the other 2, the light chains are dark blue and the heavy chains are green. V domains of the 2 molecules in the top row are on the left and the C domains are on the right; i.e. these molecules are packed in a head-to-tail arrangement. These directions are reversed in the bottom row. To aid in visualization, note that the 2 magenta C_L domains meet across the diagonal direction. Bottom, stereo diagram of 4 molecules of liganded Fab in the orthorhombic unit cell. The x-axis is horizontal and the y-axis is vertical. In this view the V domains are on the right in the top row and on the left in the bottom row. The ligands are colored yellow.

(l) *Packing interactions of the native protein in space group C2*

The native Fabs are somewhat more tightly packed than the liganded molelcules (51·5% *versus* 55·9% solvent). In Figure 11 (top panel) the view into the xz plane illustrates head-to-tail and anti-parallel side-by-side packing interactions of unliganded Fabs. Across the diagonal the C domains of light chains (magenta) are particularly close to each other.

In head-to-tail interactions, CDR1 constituents of both the light (magenta and dark blue) and heavy chains (cyan and green) and CDR2 of the heavy chain encroach on the distal C regions of the next Fab aligned along the x-axis. Arginine H56, one of the major residues in ligand binding, is in van der Waals contact with the interchain disulfide (H128) of the symmetry related molecule.

In the antiparallel series, the distal ends of the V_L domains (cyan and green) face each other in the diagonal direction in Figure 10, as do the C_L domains (magenta). There are two regions where homologous residues are juxtaposed in antiparallel Fabs. Serine H82b side-chains on adjacent V domains are within hydrogen bonding distance of each other, and the backbone carbonyl groups of two serine H165 residues in adjoining C_H1 domains are in van der Waals contact.

Between the distal ends of 2-fold-related C_L domains, as shown in Figure 11, the loops containing serine L127 or H168 are in contact with turns involving asparagine L190 or glutamine H206. These loops are sterically compatible with each other but do not contribute significantly to forces stabilizing the crystal packing.

In the yz plane side-by-side interactions occur between heavy (green) and light (magenta) chains

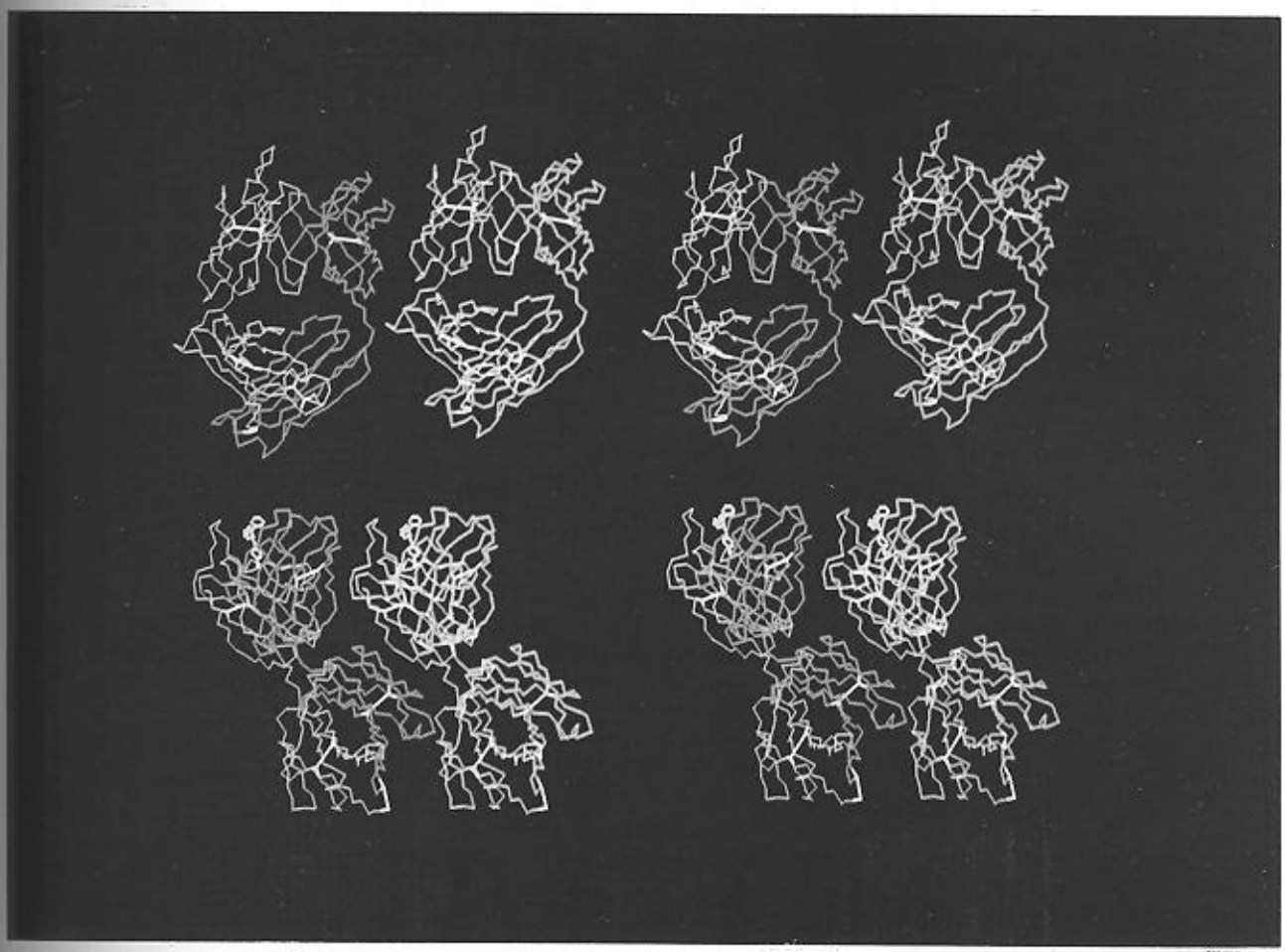

Figure 12. Comparison of the packing of the 2 species along the smallest unit cell dimension, where there is space for only 1 Fab molecule. Top, stereo diagram of 2 native Fabs in identical orientations in adjacent unit cells. The *b* cell dimension, roughly spanned by the width of each molecule in the horizontal direction, is 51·4 Å. The color scheme is the same as that in Fig. 11, with the addition of yellow disulfide bonds. Bottom, stereo diagram of 2 molecules stacked in adjacent unit cells along *c* (37·3 Å). Compare the views in Figs 11 and 12 to see how the different shapes of the 2 forms affect the crystal packing.

across the unit cell boundaries. The lower end of the first strand of V_H interacts with a turn of C_L *via* a hydrogen bond between the main-chain carbonyl oxygen of leucine H11 and the ε-amino group of lysine L199. The distal part of the switch region (residues 115 and 116) of the heavy chain is in contact with residues 202 to 204 of C_L.

Although not clearly visible in Figure 12, the N-terminal strand of the heavy chain participates in many packing interactions with the light chain of an adjacent Fab. Arginine H1 makes three contacts with the side-chain of leucine L106 and a hydrogen bond (3·1 Å) with the carbonyl oxygen of lysine L107. Glutamine H3 forms hydrogen bonds with the backbone carbonyl oxygen atom of glutamic acid L79 and the amide group of alanine L80. Other side-chain atoms of glutamine H3 are also in contact with glutamic acid L79 and alanine L80. Leucine H5 and glutamic acid L79 complete this series with three additional interactions.

(m) *Packing of NC174-Fab molecules in space group* $P2_12_12$

Unlike the native protein, the complex participates in very intricate packing. One of the most prominent features seen in Figure 11 (lower panel) is the head-to-tail packing of CDRs 1 and 2 and their adjacent framework regions in one light chain (magenta) with the distal loops of the C_L domain in the next molecule (dark blue). Of the 29 contacts of CDR1 with glycine L152, serine L153 and arginine L155, 27 involve backbone atoms, an indication of tight packing. The carbonyl oxygen atom of serine L27e forms a bifurcated hydrogen bond with two protonated nitrogen atoms of the arginine L155 side-chain, while the amide group of glycine L29 hydrogen bonds with the carbonyl oxygen atom of serine L153.

The continuation of CDR1 of the light chain, the end of FR2 and CDR2 impinge on the C-terminal

segment of C_L and a neighboring loop. Side-chains of asparagine L30 (amide group) and tyrosine L32 (hydroxyl group) form hydrogen bonds with the terminal carboxyl group of half-cystine L214. They also make seven other contacts with this residue. Tyrosine L49 (FR2) participates in 16 interactions with arginine L211 and asparagine L212, including a hydrogen bond (3·1 Å) with the backbone carbonyl oxygen atom of arginine L211. Arginine L50 (CDR2) interacts with half-cystine L214, while phenylalanine L55 extends between the C-terminal segment (asparagine L212) and the neighboring loop (asparagine L190). Serine L56 contacts backbone atoms of arginine L188, histidine L189 and asparagine L190.

CDRs 1 and 3 of the heavy chain (cyan) also encroach on the C-terminal segment of the light chain (dark blue) in the adjacent molecule. Two hydrogen bonds are formed in this series, one between the hydroxyl group of tyrosine H32 and the carboxyl oxygen of glutamic acid L213 and the second between the carboxylate of aspartic acid H101 and the side-chain amide group of asparagine L212. Tyrosine H96 and serine H97 side-chains extend to the atoms (4 contacts) of half-cystine L214, including the disulfide sulfur atom.

Interactions directly involving ligand are far too few and too weak to account for the change in space group from $C2$ to $P2_12_12$ after complexation. Only two atoms of one of the diphenyl rings of the NC174 ligand (yellow) are in van der Waals contact with the β-carbon atom of cystine L214.

In the xz plane, the c unit cell dimension (37·3 Å) is too small to accommodate even a single Fab molecule unless it is bent and stacked in the pattern shown in Figure 12, lower panel. This arrangement is stabilized by important interactions between light and heavy chains in adjacent Fabs. N-terminal glutamic acid, CDR1 and CDR3 of the light chain (magenta) are close to the N-terminal segment of the neighboring heavy chain (see end of green chain, upper right, Fig. 12). The side-chain of arginine H1 is in contact with 18 main and side-chain atoms of residues L27a, b and c of CDR1. One of the guanido nitrogen atoms is within hydrogen bonding distance of the carbonyl oxygen atoms of serine L27a and valine L27c, as well as the hydroxyl group of serine L27a. Arginine H1 also interacts with the side-chain (4 contacts) of histidine L93 of CDR3. Glutamine H3 makes single contacts with the main-chain of glutamic acid L1 and the side-chain of glutamine L27.

The molecules along z are tilted in such a way that the top outer part of C_L (residues L142 to L146) is situtated opposite one of the most distal loops of C_H1 (residues H196 to H201). Asparagine L145 contacts five heavy-chain residues and participates in 17 of the 25 interactions. These include a bifurcated hydrogen bond between the side-chain amide group and the carbonyl oxygen atoms of serine residues H195 and H196 and another hydrogen bond between the side-chain oxygen atom and the hydroxyl group of serine H202.

The overall complementarity of the stacked molecules, the relatively large number of interactions and the special anchoring effects of the side-chains of asparagine L145 and arginine H1 are features not found in the crystal packing of the unliganded Fab molecules. Such interactions are only possible when the molecules are bent into different shapes from the extended conformations of the native forms.

(n) *Attempt to fit molecules of the antigen-Fab complex into the unit cell of the native protein*

Molecules of the native protein were replaced in the $C2$ unit cell by comparable models of the complex. In this procedure the V domains of the complex were first superimposed on those of unliganded Fabs in the $C2$ unit cell on the graphics screen and then the models of the latter were removed. Two transposed molecules of the complex are shown in stereo in Figure 13. Note that the distal portions of the C domains overlap severely. It is clear that the structure of the complex is incompatible with the crystal packing of the native Fab and *vice versa*.

4. Discussion

The NC6.8 protein appears to have numerous features in common with a sweet taste receptor. Affinity for the guanidine based NC174 ligand is reasonably high and all components deemed essential for the sweet taste are matched by complementary constituents of the protein. Binding occurs by end-on insertion with the functionally most important group in the lead (i.e. the cyanophenyl moiety). If binding were dictated strictly by a hydrophobic effect with the displacement of water, we would have expected the biphenyl rings to enter the crevice first. Instead, these rings are accommodated at the entrance of the site rather non-specifically like handles on a stopper. They are exposed to bulk solvent on their outer surfaces and make only two interatomic contacts with ordered water molecules technically "inside" the active site. This internal hydration of the ligand increases in the vicinity of the zwitterionic portion (7 contacts with water). At the base of the active site, where the polar cyano group is wedged between the light and heavy chains, water is actually included rather than excluded in the binding interactions (9 contacts).

Overall interactions with the protein increase with the penetration of the ligand into the active site. The biphenyl, zwitterionic and cyanophenyl groups account for 22 (23%), 33 (34%) and 41 (43%) of the total of 96 interatomic contacts. In all three groups interactions are predominantly with heavy chain constituents (68, 97 and 69%; 77% overall). Such disproportionate favoring of the heavy chain has been noted in many other antibody-antigen complexes (Davies *et al.*, 1990; Tulip *et al.*, 1992*a*,*b*; Sheriff, 1993). The light chain plays more prominent roles in providing key contact

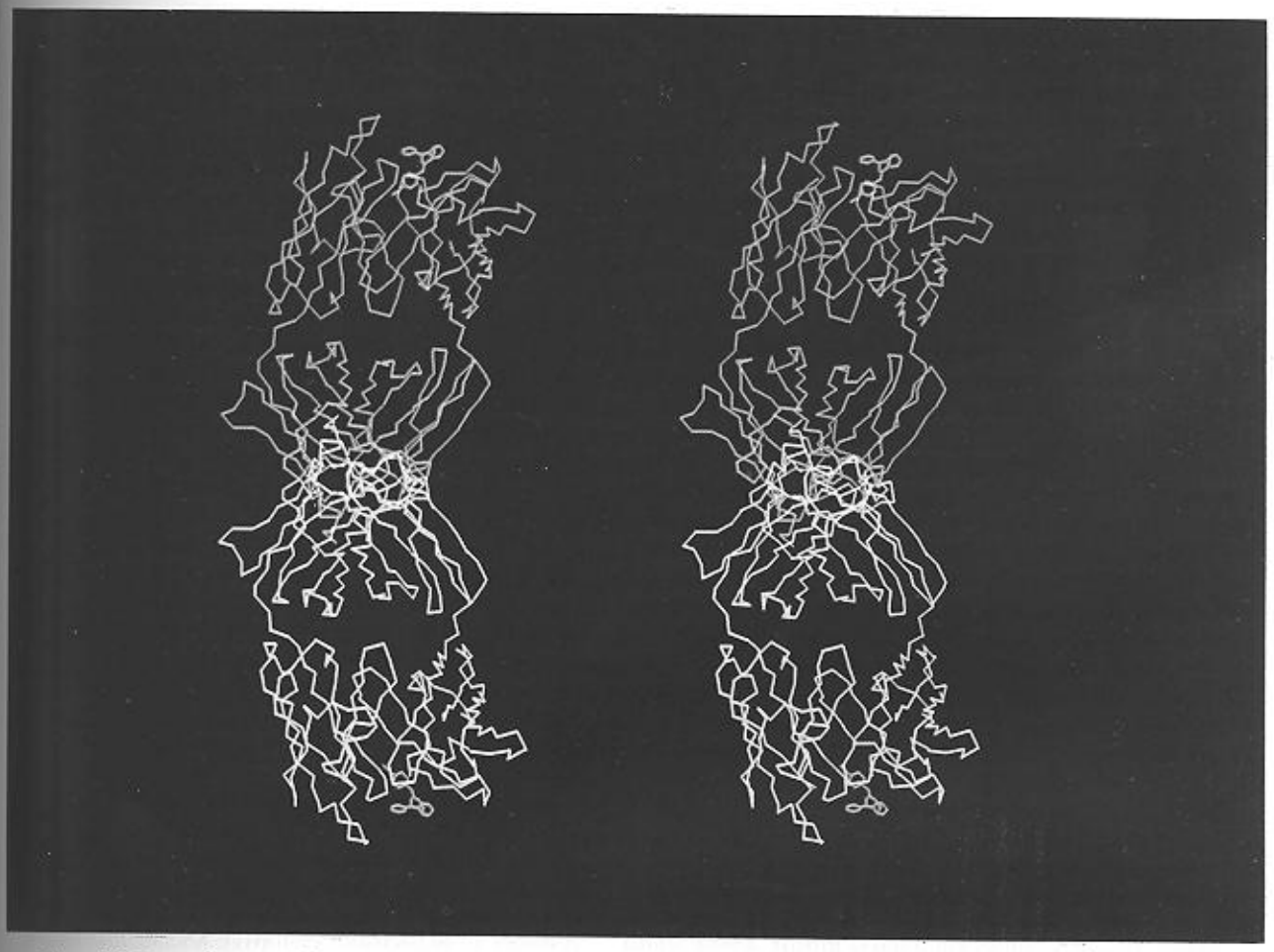

Figure 13. Stereo diagram of 2 molecules of the liganded form packed into a crystal lattice identical to that occupied by native Fabs. To produce this photograph, the V domains of the liganded structures were first superimposed on those of native Fabs placed in unit cells like that shown in the top panel of Fig. 11. When the molecules of native Fabs were removed from the graphics screen, it was clear that the substitution of liganded molecules led to unacceptable steric clashes of the type seen here.

residues or higher proportions of the interacting surfaces in fewer cases, such as the 4-4-20-fluorescein, AN02-2,2,6,6-tetramethyl-1-piperidinyloxy-dinitrophenyl group and the D1.3-E225 complexes (Herron *et al.*, 1989; Bentley *et al.*, 1990; Brünger *et al.*, 1991).

Both the diphenyl and zwitterionic groups of NC174 interact mainly with protein side-chains rather than backbone atoms (19 of 22 and 30 of 33 contacts in the 2 groups). Interactions of the diphenyl rings are primarily with aromatic residues (20 of 22 contacts). While guanido and acetate groups of the zwitterion are sterically restrained by the tryptophan H33 side-chain, the functionally most prominent interactions are polar in nature. The positive charge on the guanido group is neutralized by ion pairing with glutamic acid H50 and hydrogen bonding with the tyrosine H96 carbonyl oxygen atom. The negative charge on the acetate is balanced by ion pairing with arginine H56 and hydrogen bonding with asparagine H58.

It seems clear that the formation of a stable ligand-protein complex is dependent on conformational changes in CDR3 of the heavy chain, particularly in the backbone and side-chain of tyrosine H96. The key interaction is probably the backbone displacement of the H96 carbonyl group to make the hydrogen bond with the ligand's guanido group as mentioned above. After these changes the carbonyl oxygen atom remains in van der Waals contact with 12 atoms of the ligand. The upward rotation and translation of the tyrosine side-chain to stabilize one of the diphenyl rings appears to be a finishing touch rather than a crucial step in the initial binding events.

To understand the latter, we looked for clues in the mode of binding of the cyanophenyl group. In contrast to the zwitterion and the biphenyl group, 28 of 43 (65%) of the contacts of the cyanophenyl moiety are with backbone atoms (6 with the light and 22 with the heavy chain). Interactions with the light chain backbone are all with the C^{α} atom and

carbonyl group of glycine L91. In the heavy chain the backbone interactions are spread over the tetrapeptide sequence of Tyr-Ser-Ser-Met (residues H96 to H100k of CDR3). The cogent feature is the nearly comprehensive set of atoms involved: the carbonyl oxygen atom of tyrosine H96; the C^{α}, carbonyl carbon and carbonyl oxygen atoms of serine H97; the amide nitrogen atom, the C^{α}, carbonyl carbon and carbonyl oxygen atoms of serine H98; and the amide nitrogen atom of methionine H100k. Because of the congestion, final adjustments of atomic positions had to be made under program control (X-PLOR) to minimize steric clashes.

After complexation, the C–C≡N linear array of atoms and the lower part of the phenyl ring are thus found to be closely sandwiched between the two chains. Remaining (upper) surfaces of this ring are covered by the side-chains of tyrosine L96, glutamic acid H35 and glutamic acid H50. Ordered water molecules occupy the small volume of space remaining in the subsite. Opposing surfaces of the cyanophenyl group, water molecules and protein substituents are too tightly packed in the complex to be compatible with the unaltered native structure. We concluded that penetration of the cyanophenyl group into the narrow space between glycine L91 and CDR3 of the heavy chain is the probable initiating step in events leading to dissolution of crystals of the native protein.

Locally, the wedge serves to tighten some interactions that are sterically possible without conformational changes and to shift CDR3 segments to make room for the cyanophenyl group. These small changes are amplified into large global movements. The heavy chain is pushed by the ligand and behaves as if it is in compression. Its overall length is shortened as the angle between the V and C domains decreases. At the same time, the light chain acts as if it is in tension. The angle between V and C domains increases and the light chain becomes more elongated. These concerted and reciprocal changes in the overall conformations are facilitated because the chains are linked covalently by a disulfide bond at the distal ends of the C domains and noncovalently by sets of pairing interactions in the V and C domains.

The effects of these changes on crystal packing are profound. As outlined earlier, the elbow bend angle is decreased by 31° and the C domain pair is rotated, swiveled and tilted relative to the V pair. These changes alter the shape of the Fab envelope and the relative locations of surface features important for crystal packing of the native protein. When NC174 is added to a crystal of the native protein, the stabilizing elements are rearranged and unacceptable steric clashes occur between the C domains of neighboring molecules. The original crystal is shattered and replaced by a new form which has a different space group and unit cell dimensions.

Adjacent to the site of papain cleavage of the heavy chain to produce the Fab (glutamic acid H224), the polypeptide backbone is locked into the globular part of C_H1 by hydrogen bonding (see Fig. 7). Consequently, this segment accompanies the C_H1 domain in the rigid body movement described above; i.e. it is rotated, swiveled and tilted. The α-carbon atom of the C-terminal glutamic acid H224 is displaced 19 Å from its position in the native protein. Because of the contraction in length of the heavy chain, it is also 5 Å closer to V_H. In an intact IgG molecule these changes in position would result in relayed tension on the strand connecting the Fab to the Fc.

Collectively, these observations suggest a mechanism for intramolecular signaling between the V and C domains of the Fab in the NC174-Fab complex. If this transmission continues beyond the Fab regions, it could well influence the relative orientation of the Fc, where complement activation and other effector functions are performed.

In the crystal structure of an intact murine antibody (Mab 231) of the same subclass as NC6 (IgG_{2b}), the Fab-Fc connecting segments act as elongated tethers rather than a more restrictive "hinge" (Harris *et al.*, 1992). The Fc region of the IgG_{2b} molecule is apparently easy to displace, since it assumes an asymmetrical orientation relative to the two Fab arms. Mobility of the Fc was previously surmised in crystal studies of the Kol human IgG_1 molecule (Colman *et al.*, 1976; Marquart *et al.*, 1980) and the Zie human IgG_2 protein (Ely *et al.*, 1978) where the Fc regions were too disordered to contribute to the X-ray diffraction patterns.

In the mid-1970's examination of a limited number of structures of native and liganded forms (Sarma *et al.*, 1971; Poljak *et al.*, 1973; Schiffer *et al.*, 1973; Segal *et al.*, 1974; Epp *et al.*, 1974; Amzel *et al.*, 1974; Edmundson *et al.*, 1974, 1975; Fehlhammer *et al.*, 1975; Padlan *et al.*, 1976; Colman *et al.*, 1976) led to a suggestion that antigen binding is accompanied by a stiffening of a flexible antibody molecule through the formation of longitudinal contacts between domains (Huber *et al.*, 1976). While the hypothesis did not prove valid in detail, the concept of partially closed interdomain contacts influenced the thinking of our laboratory from the outset. We still do not discount the possibility of increased rigidity after complexation. The NC174-Fab complex is definitely more ordered in its crystal structure than the native protein.

The heavy chain analog of the Mcg light chain dimer exemplifies the facility with which the V-C interface is opened and closed. A comparison of the structures of the orthorhombic (water) and trigonal (ammonium sulfate) forms of the dimer (Ely *et al.*, 1989; Schiffer *et al.*, 1973; Edmundson *et al.*, 1975) reveals that the slightly open V-C interface in the heavy chain analog of the orthorhombic form closes significantly when the crystallization medium is changed to ammonium sulfate. The V-C bend angle decreases by about 11° from 53° to 42° and at least nine water molecules are expelled from the interface. While the V-C bend angle in the light chain analog does not change (108·0° *versus* 107·8°), the overall elbow bend of the dimer decreases from 132°

to 115°. Moreover, the center of mass of the C domain shifts laterally by 6 Å relative to its position in the orthorhombic form. As in the NC174-Fab complex, an increase in the flexion of one of the two chains leads to C domain movements, which are amplified into large swings of the segments at the distal ends.

The authentic heavy chain in the Mcg intact IgG_1 immunoglobulin has an even more closed and more complicated conformation at the V_H-C_H1 junction (V-C bend angle of 30°; Guddat *et al.*, 1993). In interacting with the heavy chain to produce the Fab, the light chain retains an open conformation (V-C bend angle of 108°) but the center of mass of the C_L domain is translated laterally by 16 Å relative to that in the orthorhombic form of the light chain dimer. A crystal structure of a liganded-IgG_1 complex has not as yet been determined for comparison, but it seems likely that flexion at the V-C junction is again coupled with sizeable movements of the C domains. These crystallographic studies reveal the ease with which the C domains can be moved with respect to the V domains and suggest that this property can be utilized for intramolecular signaling.

Two other murine antibodies in our collection were compared with the Mcg proteins because they are very dissimilar in their quaternary structures. Fabs from the 4-4-20 IgG_{2a} (Gibson *et al.*, 1985; Herron, *et al.*, 1989) and the BV04-01 IgG_{2b} (Herron *et al.*, 1991) antibodies are extended structures with elbow bend angles >170°. The liganded form of 4-4-20 (anti-fluorescein), crystallized in 2-methyl-2,4-pentanediol is nearly isomorphous with the unliganded (native) Fab of BV04-01 (anti-single-stranded-DNA), crystallized in ammonium sulfate. Unliganded 4-4-20 has not been crystallized. The V-C bend angles for the light and heavy chains are 83° and 96° (4-4-20, liganded) and 81° and 96° (BV04-01, native), values very similar to those of native NC6.8. In all three proteins the light chain adopts a flexed conformation while the heavy chain is more extended. This is a reversal of the situation in the Mcg series and reinforces the conclusion that the heavy chain can mimic the light chain in its range of conformational flexibility.

On complexation of BV04-01 with a trinucleotide of deoxythymidylate, the crystal habit and the space group are both altered. The active site of BV04-01 is enlarged by local conformational changes and by rigid body movements of V_H relative to V_L (Herron *et al.*, 1991). These changes are accompanied by concerted shifts in the V-C bend angles of the light (81° to 84°) and heavy chains (96° to 88°). Although the movements are smaller, the reciprocal changes on complexation are in the same directions as in NC6.8 (i.e. extension of the light chain, flexion of the heavy chain).

We subjected other molecules to similar calculations. The human Kol IgG_1 (unliganded) and its Fab fragment (Matsushima *et al.*, 1978; Marquart *et al.*, 1980) show only small differences in the V-C bend angles of the light (84° *versus* 86°) and heavy chains (90° *versus* 88°). In the murine B13I2 IgG_1 molecule (activity against a peptide from myohemerythrin; Stanfield *et al.*, 1990), the light chain is extended (bend angles of 93° and 92°) and the heavy chain is flexed (74° and 77°) before and after binding of the ligand. The 17/9 IgG_{2a} antibody (activity against a peptide from influenza hemagglutinin; Rini *et al.*, 1992) behaves in the opposite way from NC6.8. On complexation the bend angle decreases for the light chain (88° to 82°) and increases for the heavy chain (82° to 92°).

It is well known that the binding of antigen is a prerequisite for the aggregation of IgG antibodies and attachment of the first component (C1q) of the complement cascade to the Fc region. This aggregation process is generally associated with larger (often polyvalent) antigens rather than haptens (Metzger, 1983).

We agree with the reviewer that the present results are very different from those of many other hapten studies, including our own. If NC174 is unusual, we should be able to devise a set of experiments to evaluate what features are responsible for its special effects.

Preliminary analyses indicate that these effects are not attributable to the protein *per se* or to peculiarities in the crystal packing. Among other sweet tasting compounds binding to the NC6.8 antibody, two compounds lacking the cyanophenyl moiety (NC24, a 2-dimethyl-4-methyl-bicyclo [2.2.1] heptanyl derivative of guanidine and NC90, a cyclooctanyl derivative) fail to dissolve crystals of unliganded NC6.8 Fab. Cocrystals of these ligands with NC6.8 Fab resemble those of the native Fab rather than the crystals of the NC174-Fab complex. For example, both complexes crystallize in space group $C2$, with unit cell dimensions indistinguishable from those of the unliganded protein within the errors of the methods.

While the structural studies of the complexes are still in progress, it is fair to say that the Fab moieties adopt the native type of conformation in the complexes with either NC24 or NC90. NC24 and NC90 have been outlined in 3·0 Å electron density maps. In each complex the ligand fits securely in the active site but the penetrating end (a carboxyl group) is not wedged between the light and heavy chains like the cyanophenyl group of NC174. Moreover, the mobile tyrosine H96 residue remains in the "down" position similar to the orientation in the native protein. It should be emphasized that the two alternative ligands bind to NC6.8 with affinity constants two orders of magnitude lower than that of NC174.

It is possible to prepare crystals of a double derivative of NC6.8 Fab with NC174 and NC24. This complex crystallizes in orthorhombic space group $P2_12_12_1$ rather than $P2_12_12$ as in the single derivative with NC174. Again the results are preliminary but the Fab moiety has the now familiar conformation associated with the binding of NC174 alone.

In summary, the facile interconversion of light

and heavy chain conformations suggests one possible mechanism for amplification and transmission of conformational changes associated with ligand binding. For NC6.8 and NC174 the signaling is dependent on reciprocal flexion and extension of the two chains. In the Mcg IgG1 or light chain dimer, flexion of the heavy chain alone can lead to substantial lateral displacements of the C_H1-C_L or C_L1-C_L2 module because the two C domains are tightly coupled.

NC174 departs from the more familiar patterns associated with the binding of less intrusive haptens to the NC6.8 Fab. It will be interesting to see if it mimics the long distance actions of larger antigens.

Direct evidence remains to be gathered to test whether the sizeable movements of the last segment in the C_H1 domain continue into the Fc. The latter has also been shown to be susceptible to large shifts in orientation, particularly in the IgG_{2b} series of antibodies. At the present time we can only speculate whether any such transmitted motions are associated with the conformational changes and aggregation required for complement lysis. If we restrict the arguments to functions distal to the active site in the Fab alone, the NC6.8 molecule can be classified as an allosteric protein. Its overall structure can clearly be altered by the binding of a small ligand like NC174.

We are grateful to Kim Andersen for her many contributions in the computing, graphics and bibliographic aspects of this project. We thank Professor Noriyoshi Sakabe for allowing us to use his Weissenberg camera and for his kind hospitality; Dr Bret Church for his help in collecting the X-ray diffraction data; Angela Lampkin for help in processing the data; Leif Hanson and Zhao-Chang Fan for numerous discussions. Part of this study (D.S.L.) was facilitated by a collaborative research agreement with the NutraSweet Co. (Mt. Prospect, IL) and we thank Drs Grant DuBois, Mike Kellog, Srinivasan Nagarajan, Jeff Carter, Chris Culberson and Eric Walters for their helpful discussions and gifts of ligands used in the study. This work was supported by grant CA 19616 from the National Cancer Institute to A.B.E., grant GM46535 from the National Institutes of Health to D.S.L. and The Harrington Cancer Center to A.B.E.

References

Amzel, L. M., Poljak, R. J., Saul, F., Varga, J. M. & Richards, F. F. (1974). The three-dimensional structure of a combining region-ligand complex of immunoglobulin NEW at 3·5 Å resolution. *Proc. Nat. Acad. Sci., U.S.A.* **71**, 1427–1430.

Anchin, J. M. & Linthicum, D. S. (1992). Molecular and computational techniques for modelling antibody combining sites . *J. Clin. Immunoassay*, **15**, 35–41.

Anchin, J., Subramaniam, S. & Linthicum, D. S. (1991). Binding of the neuroleptic drug haloperidol to a monoclonal antibody: refinement of the binding site molecular model using canonical structures. *J. Mol. Recog.* **4**, 7–15.

Bassolino-Klimas, D., Bruccoleri, R. E. & Subramaniam, S. (1992). Modeling the antigen combining site of an antidinitrophenyl antibody, ANO2. *Prot. Sci.* **1**, 1465–1476.

Bedzyk, W. D., Herron, J. N., Edmundson, A. B. & E. W., Jr (1990). Active site structure and binding properties of idiotypically cross-reactive antifluorescein monoclonal antibodies. *J. Biol.* **265**, 133–138.

Bentley, G. A., Boulot, G., Riottot, M. M. & Poljak (1990). Three-dimensional structure of an idiotope antiidiotope complex. *Nature (London)* 254–257.

Bhat, T. N. & Cohen, G. H. (1984). OMITMAP: electron density map suitable for examination errors in a macromolecular model. *Crystallogr.* **17**, 244–248.

Bhat, T. N., Bentley, G. A., Fischmann, T. O., & Poljak, R. J. (1990). Small rearrangements structures of Fv and Fab fragments of antibody on antigen binding. *Nature (London)*, 347 485.

Biosym (1989). *Insight User Manual, Version 2.5*. Technologies, San Diego, CA.

Brünger, A. T., Karplus, M. & Petsko, G. A. Crystallographic refinement by simulated anneal application to crambin. *Acta Crystallogr. sect. A* 50–61.

Brünger, A. T., Leahy, D. J., Hynes, T. R. & Fox (1991). 2·9 Å resolution structure of an dinitrophenyl-spin-label monoclonal antibody fragment with bound hapten. *J. Mol. Biol.* 239–256.

Carson, M. (1987). Ribbon models of macromolecules *J. Mol. Graph.* **5**, 103–106.

Carson, M. & Bugg, C. (1986). Algorithm for ribbon models of proteins. *J. Mol. Graph.* **4**, 121–122.

Colman, P. M. (1988). Structure of antibody-antigen complexes: implications for immune recognition *Advan. Immunol.* **43**, 99-132.

Colman, P. M., Deisenhofer, J., Huber, R. & Palm (1976). Structure of the human antibody molecule Kol (immunoglobulin G1): an electron density map 5 Å resolution. *J. Mol. Biol.* **100**, 257–282.

Crowther, R. A. (1972). Fast rotation function. In *Molecular Replacement Method: A Collection of Papers on the Use of Non-crystallographic Symmetry* (Rossmann, M. G., ed.), pp. 173–178, Gordon Breach, New York.

Crowther, R. A. & Blow, D. M. (1967). A method positioning a known molecule in an unknown crystal structure. *Acta Crystallogr.* **23**, 544–548.

Davies, D. R., Padlan, E. A. & Sheriff, S. (1987). Antibody-antigen complexes. *Annu. Rev. Biochem.* **59**, 439–473.

Droupadi, P. R., Anchin, J. M., Meyers, E. A. Linthicum, D. S. (1992). Spectrofluorimetric study the intermolecular complexation of monoclonal antibodies with the high potency sweetener N-(1-cyanophenyl)-N'-(diphenylmethyl) guanidinoacetic acid. *J. Mol. Recog.* **5**, 173–179.

Edmundson, A. B. & Ely, K. R. (1985). Binding of N-formylated chemotactic peptides in crystals of the Mcg light chain dimer: similarities with neutrophil receptors. *Mol. Immunol.* **22**, 463–475.

Edmundson, A. B., Ely, K. R., Girling, R. L., Abola, E. E., Schiffer, M., Westholm, F. A., Fausch, M. D. & Deutsch, H. F. (1974). Binding of 2,4-dinitrophenyl compounds and other small molecules to a crystalline λ-type Bence-Jones dimer. *Biochemistry* 3816–3827.

Edmundson, A. B., Ely, K. R., Abola, E. E., Schiffer, & Panagiotopoulos, N. (1975). Rotational allomerism and divergent evolution of domains in

immunoglobulin light chains. *Biochemistry*, **14**, 3953–3961.

Edmundson, A. B., Ely, K. R. & Herron, J. N. (1984). A search for site-filling ligands in the Mcg Bence-Jones dimer: crystal binding studies of fluorescent compounds. *Mol. Immunol.* **21**, 561–576.

Edmundson, A. B., Ely, K. R., Herron, J. N. & Cheson, B. D. (1987). The binding of opioid peptides to the Mcg light chain dimer: flexible keys and adjustable locks. *Mol. Immunol.* **24**, 915–935.

Ely, K. R., Firca, J. R., Williams, K. J., Abola, E. E., Fenton, J. M., Schiffer, M., Panagiotopoulos, N. C. & Edmundson, A. B. (1978). Crystal properties as indicators of conformational changes during ligand binding or interconversion of Mcg light chain dimers. *Biochemistry*, **17**, 158–167.

Ely, K. R., Herron, J. N., Harker, M. & Edmundson, A. B. (1989). Three-dimensional structure of a light chain dimer crystallized in water: conformational flexibility of a molecule in two crystal forms. *J. Mol. Biol.* **210**, 601–615.

Engh, R. A. & Huber, R. (1991). Accurate bond and angle parameters for X-ray protein structure refinement. *Acta Crystallogr. sect. A*, **47**, 392–400.

Epp, O., Colman, P., Fehlhammer, H., Bode, W., Schiffer, M., Huber, R. & Palm, W. (1974). Crystal and molecular structure of a dimer composed of the variable portions of the Bence-Jones protein REI. *Eur. J. Biochem.* **45**, 513–524.

Fan, Z.-C., Shan, L., Guddat, L. W., He, X. M., Gray, W. R., Raison, R. L. & Edmundson, A. B. (1992). Three-dimensional structure of an Fv from a human IgM immunoglobulin. *J. Mol. Biol.* **228**, 188–207.

Fehlhammer, H., Schiffer, M., Epp, O., Colman, P. M., Lattman, E. E., Schwager, P. & Steigemann, W. (1975). The structure determination of the variable portion of the Bence-Jones protein Au. *Biophys. Struct. Mech.* **1**, 139–146.

Fitzgerald, P. M. D. (1988). MERLOT, an integrated package of computer programs for the determination of crystal structures by molecular replacement. *J. Appl. Crystallogr.* **21**, 273–278.

Gibson, A. L., Herron, J. N., Ballard, D. W., Voss, E. W., Jr, He, X. M., Patrick, V. A. & Edmundson, A. B. (1985). Crystallographic characterization of the Fab fragment of a monoclonal anti-ss-DNA antibody. *Mol. Immunol.* **22**, 499–502.

Guddat, L. W., Herron, J. N. & Edmundson, A. B. (1993). Three-dimensional structure of a human immunoglobulin with a hinge deletion. *Proc. Nat. Acad. Sci., U.S.A.* **90**, 4271–4275.

Harris, L. J., Larson, S. B., Hassel, K. W., Day, J., Greenwood, A. & McPherson, A. (1992). The three-dimensional structure of an intact monoclonal antibody for canine lymphoma. *Nature (London)*, **360**, 369–372.

Haurowitz, F. (1938). Das gleichgewicht zwischen hämoglobin und sauerstoff. *Hoppe-Seyler's Z. Physiol. Chem.* **254**, 266–274.

Herron, J. N., He, X. M., Mason, M. L., Voss, E. W., Jr & Edmundson, A. B. (1989). Three-dimensional structure of a fluorescein-Fab complex crystallized in 2-methyl-2,4-pentanediol. *Proteins: Struct. Funct. Genet.* **5**, 271–280.

Herron, J. N., He, X. M., Ballard, D. W., Blier, P. R., Pace, P. E., Bothwell, A. L. M., Voss, E. W., Jr & Edmundson, A. B. (1991). An autoantibody to single-stranded DNA: comparison of the three-dimensional structures of the unliganded Fab and a deoxynucleotide-Fab complex. *Proteins: Struct. Funct. Genet.* **11**, 159–175.

Higashi, T. (1989). The processing of diffraction data taken on a screenless Weissenberg camera for macromolecular crystallography. *J. Appl. Crystallogr.* **22**, 9–18.

Howard, A. J., Gilliland, G. L., Finzel, B. C., Poulos, T. L., Ohlendorf, D. H. & Salemme, F. R. (1987). The use of an imaging proportional counter in macromolecular crystallography. *J. Appl. Crystallogr.* **20**, 383–387.

Huber, R., Deisenhofer, J., Colman, P. M., Matsushima, M. & Palm, W. (1976). Crystallographic structure studies of an IgG molecule and an Fc fragment. *Nature (London)*, **264**, 415–420.

Kabat, E. A., Wu, T. T., Perry, H. M., Gottesman, K. S. & Foeller, C. (1991). *Sequences of Proteins of Immunological Interest*, 5th edit., Public Health Service, U.S. Department of Health and Human Services, National Institutes of Health, Bethesda, MD.

Kussie, P. H., Anchin, J., Subramaniam, S., Glasel, J. A. & Linthicum, D. S. (1991). Analysis of the binding site architecture of monoclonal antibodies to morphine using competitive ligand binding and molecular modelling. *J. Immunol.* **146**, 4248–4257.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)*, **227**, 680–685.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cyrstallogr.* **26**, 283–291.

Lee, B. & Richards, F. M. (1971). The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55**, 379–400.

Luzatti, P. V. (1952). Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallogr.* **5**, 802–810.

Mandal, C. & Linthicum, D. S. (1992). Computer-aided modelling of complete antibody Fab structures using alpha carbon atom coordinates. *J. Clin. Immunoassay*, **15**, 42–50.

Marquart, M., Deisenhofer, J., Huber, R. & Palm, W. (1980). Crystallographic refinement and atomic models of the intact immunoglobulin molecule Kol and its antigen-binding fragment at 3·0 Å and 1·9 Å resolution. *J. Mol. Biol.* **141**, 369–391.

Matsushima, M., Marquart, M., Jones, T. A., Colman, P. M., Bartels, K., Huber, R. & Palm, W. (1978). Crystal structure of the human Fab fragment Kol and its comparison with the intact Kol molecule. *J. Mol. Biol.* **121**, 441–459.

Matthews, B. W. (1968). Solvent content of protein crystals. *J. Mol. Biol.* **33**, 491–497.

Metzger, H. (1983). The effect of antigen on antibodies: recent studies. *Contemp. Topics Mol. Immunol.* **7**, 119–152.

Muller, G. W., Walters, D. E. & DuBois, G. E. (1992). *N,N'* disubstituted guanidine high-potency sweeteners. *J. Med. Chem.* **35**, 740–743.

Padlan, E. A., Davies, D. R., Rudikoff, S. & Potter, M. (1976). Structural basis for the specificity of phosphorylcholine-binding immunoglobulins. *Immunochemistry*, **13**, 945–949.

Poljak, R. J., Amzel, L. M., Avey, H. P., Chen, B. L., Phizackerly, R. P. & Saul, F. (1973). Three-dimensional structure of the Fab' fragment of a human immunoglobulin at 2·8 Å resolution. *Proc. Nat. Acad. Sci., U.S.A.* **70**, 3305–3310.

Rajan, S. S., Ely, K. R., Abola, E. E., Wood, M. K., Colman, P. M., Athay, R. J. & Edmundson, A. B. (1983). Three-dimensional structure of the Meg IgG1 immunoglobulin. *Mol. Immunol.* **20**, 787–799.

Ramachandran, G. N. & Sasisekharan, V. (1968). Conformation of polypeptides and proteins. *Advan. Protein Chem.* **23**, 283–437.

Rini, J. M., Schulze-Gahmen, U. & Wilson, I. A. (1992). Structural evidence for induced fit as a mechanism for antibody-antigen recognition. *Science*, **255**, 959–965.

Rossmann, M. G. (1990). The molecular replacement method. *Acta Crystallogr. sect. A*, **46**, 73–82.

Roussel, A. & Cambillau, C. (1989). TURBO-FRODO molecular graphics program. In *Silicon Graphics Geometry Partner Directory*, pp. 77-78, Silicon Graphics, Mountain View, CA.

Rowland, R. S., Carson, W. M. & Bugg, C. E. (1990). Applications of crystallographic databases in molecular design. In *Use of X-ray Crystallography in the Design of Antiviral Agents* (Laver, W. G. & Air, G. M., eds), pp. 261–281, Academic Press, San Diego, CA.

Sakabe, N. (1983). A focusing Weissenberg camera with multilayer-line screens for macromolecular crystallography. *J. Appl. Crystallogr.* **16**, 542–547.

Sarma, V. R., Silverton, E. W., Davies, D. R. & Terry, W. D. (1971). The three-dimensional structure at 6 Å resolution of a human γG1 immunoglobulin molecule. *J. Biol. Chem.* **246**, 3753–3759.

Schiffer, M., Girling, R. L., Ely, K. R. & Edmundson, A. B. (1973). Structure of a λ-type Bence-Jones protein at 3·5 Å resolution. *Biochemistry*, **12**, 4620–4631.

Segal, D. M., Padlan, E. A., Cohen, G. H., Rudikoff, S., Potter, M. & Davies, D. R. (1974). The three-dimensional structure of a phosphorylcholine-binding mouse immunoglobulin Fab and the nature of the antigen binding site. *Proc. Nat. Acad. Sci., U.S.A.* **71**, 4298–4302.

Shan, L., Guddat, L. W., Raison, R. L. & Edmundson, A. B. (1993). Crystallization of an Fv fragment from a human IgM cryoglobulin by a microseeding technique. *J. Crystl. Growth*, **126**, 229–244.

Sheriff, S. (1993). Some methods to examine the interaction between two molecules. *Immunomethods* in the press.

Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C. & Davies, D. R. (198?). Three-dimensional structure of an antibody-antigen complex. *Proc. Nat. Acad. Sci., U.S.A.* **8**?, 8075–8079.

Stanfield, R. L., Fieser, T. M., Lerner, R. A. & Wilson, I. A. (1990). Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2·8 Å. *Science*, **248**, 712–719.

Thomas, K. A., Smith, G. M., Thomas, T. B. & Feldmann, R. J. (1982). Electronic distributions within protein phenylalanine aromatic rings are reflected by the three-dimensional oxygen atom environments. *Proc. Nat. Acad. Sci., U.S.A.* **79**, 4843–4847.

Tormo, J., Stadler, E., Skern, T., Auer, H., Kamzer, B., Betzel, C., Blaas, D. & Fita, I. (199?). Three-dimensional structure of the Fab fragment of a neutralizing antibody to human rhinovirus serotype 2. *Protein Sci.* **1**, 1154–1161.

Tulip, W. R., Varghese, J. N., Laver, W. G., Webster, R. G. & Colman, P. M. (1992*a*). Refined crystal structure of the influenza virus N9 neuraminidase-NC41 Fab complex. *J. Mol. Biol.* **227**, 122–148.

Tulip, W. R., Varghese, J. N., Webster, R. G., Laver, W. G. & Colman, P. M. (1992*b*). Crystal structures of two mutant neuraminidase-antibody complexes with amino acid substitutions in the interface. *J. Mol. Biol.* **227**, 149–159.

Walters, D. E., Orthoefer, F. T. & Dubois, G. E. (199?). Three-dimensional model for the sweet taste receptor. In *Sweeteners: Discovery, Molecular Design and Chemoreception (ACS Symposia 450)* p. 2?, ACS, Washington, DC.

Edited by R. Huber

(Received 27 April 1993; accepted 22 September 1993)