NMR-Derived Model for a Peptide-Antibody Complex[†,‡]

Barbara Zilber,[§] Tali Scherf,[§] Michael Levitt,[∥] and Jacob Anglister*[,§,⊥]

Department of Polymer Research, The Weizmann Institute of Science, Rehovot 76100, Israel, and Department of Cell Biology, Stanford University School of Medicine, Stanford, California 94305

Received January 23, 1990; Revised Manuscript Received June 5, 1990

ABSTRACT: The TE34 monoclonal antibody against cholera toxin peptide 3 (CTP3; VEVPGSQHIDSQKKA) was sequenced and investigated by two-dimensional transferred NOE difference spectroscopy and molecular modeling. The V_H sequence of TE34, which does not bind cholera toxin, shares remarkable homology to that of TE32 and TE33, which are both anti-CTP3 antibodies that bind the toxin. However, due to a shortened heavy chain CDR3, TE34 assumes a radically different combining site structure. The assignment of the combining site interactions to specific peptide residues was completed by use of AcIDSQRKA, a truncated peptide analogue in which lysine-13 was substituted by arginine, specific deuteration of individual polypeptide chains of the antibody, and a computer model for the Fv fragment of TE34. NMR-derived distance restraints were then applied to the calculated model of the Fv to generate a three-dimensional structure of the TE34/CTP3 complex. The combining site was found to be a very hydrophobic cavity composed of seven aromatic residues. Charged residues are found in the periphery of the combining site. The peptide residues HIDSQKKA form a β-turn inside the combining site. The contact area between the peptide and the TE34 antibody is 388 Å^2, about half of the contact area observed in protein-antibody complexes.

Recent crystallographic studies of protein-antibody complexes have greatly contributed to our understanding of the recognition of protein antigens by antibodies (Amit et al., 1986; Sheriff et al., 1987; Padlan et al., 1989; Colman et al., 1987). The combining site for protein antigens differs in several respects from that for small ligand molecules. The contact area between antibodies and proteins is about 700 Å. All antibody CDRs[1] participate in binding, and an average of 17 antibody residues interacts with 15 protein residues. These combining sites were found to be rather flat with protuberances and depressions that fit the shape of the antigen. In contrast, antibodies against small molecules were found to contain a pocket or groove with a ligand contact area much smaller than that observed in protein-antibody complexes. Despite interest in understanding the antigenic structure of peptides, complexes between peptide and antibodies have not been studied in detail by X-ray crystallography.

The interactions of three antibodies (TE32, TE33, and TE34) with CTP3, a peptide corresponding in sequence to residues 50-64 of the B subunit of cholera toxin (VEVPGSQHIDSQKKA), have been investigated by two-dimensional transferred NOE difference spectroscopy on Fab/CTP3 complexes (Levy et al., 1989; Anglister & Zilber, 1990). We found that when TE32 or TE33 interacts with excess CTP3, the exchange between bound and free peptide is fast relative to the spin-lattice relaxation time T_1 of the Fab and peptide protons, and therefore, magnetization transferred between Fab and neighboring peptide protons is further transferred to the free peptide. As a result, the NOESY spectrum of the Fab in the presence of a large peptide excess (excess spectrum) contains extra cross peaks due to interactions between the antibody and the bound peptide and the further exchange between bound and free peptide. These extra cross peaks, the so-called transferred NOE, are not present in the NOESY spectrum when the Fab and the peptide are in a 1:1 ratio (saturated spectrum). Thus, by subtracting the saturated spectrum from the excess spectrum, we succeeded in eliminating the numerous cross peaks due to intra-Fab interactions, and the transferred NOE cross peaks due to the antibody-peptide interactions are clearly observed (Anglister et al., 1989). The assignment of the interacting peptide protons was accomplished on the basis of the COSY spectrum of the free peptide. To assign overlapping resonances from different peptide residues, we repeated the experiments with specifically deuterated peptide and followed the disappearance of cross peaks. The assignment to specific type of antibody aromatic amino acids was accomplished by specific deuteration of the antibody. This deuteration was accomplished by feeding

[†] This work was supported by grants from the Joseph and Ceil Mazer Center for Structural Biology, the Kimmelman Center for Biomolecular Structure and Assembly, and the U.S.-Israel Binational Science Foundation (88-00450).

[‡] The nucleotide sequence data reported in this paper for the TE34, TE33, and TE32 antibodies will appear in EMBL, GenBank, and DDBJ Nucleotide Sequence Databases under Accession Numbers M30458-M30459 and M30480-M30483 inclusive.

* Correspondence should be addressed to this author.

[§] The Weizmann Institute of Science.

[∥] Stanford University School of Medicine.

[⊥] Incumbent of the Graham and Rhona Beck Career Development Chair.

[1] Abbreviations: Ac, acetyl; CDR, *C*omplementarity *D*etermining *R*egion of the antibody molecule; COSY, 2D *J*-correlated spectroscopy; CTP3, *C*holera *T*oxin *P*eptide 3; ELISA, enzyme-linked immunosorbent assay; Fab, antibody *F*ragment made of the Fv, the light chain, and the first heavy chain constant regions; FDR, *F*rame *D*etermining *R*egion of the antibody molecule, the segments that build the conserved structure of the immunoglobulin fold; Fv, antibody *F*ragment made of the variable regions of the light and heavy chains forming a single combining site for antigen; NCA, *N*-acetyl-*C*TP3 *A*mide; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; TRNOE, transferred NOE; V_H, variable region of the heavy chain; 2D, two dimensional; 2D TRNOE difference spectrum, calculated 2D difference spectrum between the measured NOESY spectrum of the protein saturated with the ligand and that of the protein measured in the presence of a large excess of ligand. Name convention: (a) specific peptide residue, lower-case name (e.g., glutamine-7); (b) specific antibody residue, upper-case name (e.g., Tyr 32L).

0006-2960/90/0429-10032$02.50/0

deuterated amino acids to the hybridoma cells that produce the antibodies.

While TE32 and TE33 bind cholera toxin in a solid-phase immunosorbent assay (ELISA), TE34 does not bind the toxin (Anglister et al., 1988). The off-rate of CTP3 from TE34 was too slow to measure strong TRNOE cross peaks between the antibody and the peptide (Anglister & Zilber, 1990). Much faster off-rates, resulting in strong TRNOEs, were obtained for two peptide analogues: (a) CTP3 with a C-terminal amide (Ac-VEVPGSQHIDSQKKA-NH_2); (b) a truncated version of the peptide (*N*-acetyl-IDSQKKA). These modifications did not interfere significantly either with the interactions of the unmodified part of the peptide with the antibody or with intramolecular interactions occurring in the epitope recognized by the antibody (Anglister & Zilber, 1990). The combined application of these peptides allowed study of the interactions between the antibody and the whole peptide. Two tyrosine residues and one or more tryptophan and phenylalanine residues were found to interact with histidine-8, isoleucine-9, aspartate-10, lysine-13 and/or lysine-14, and alanine-15 of the peptide. In the bound peptide, we observed interactions of a lysine residue with aspartate-10 β-protons. While the peptide epitope recognized by TE34 is between histidine-8 and the negatively charged C-terminus, that recognized by both TE32 and TE33 includes residues 3–10 of CTP3.

Measurements of the intensity of intramolecular magnetization transfer between a pair of neighboring protons after a short irradiation period of one of them or after a short mixing time in a NOESY experiment are used to evaluate the distance between the protons. Such measurements are extensively used to obtain restraints on proton–proton distances subsequently used to calculate the three-dimensional structure of small proteins (Wüthrich, 1986). Clore and Gronenborn (1983) and Clore et al. (1986) extended the application of such measurements for transferred NOE in protein–ligand complexes. This analysis distinguishes four regions of chemical exchange. In the case of fast exchange between bound and free ligand, when the ligand off-rate is faster then 10 times the spin–lattice relaxation time of its protons and when the resonances of the bound and free ligand are averaged, the magnitude of the transferred NOE ($N_{H,A}$) between an antibody proton, A, and a peptide ligand proton, H, is given by $N_{H,A} \sim (1 - a)\sigma_{H,A}\tau_m$ (Clore & Gronenborn, 1983; Clore et al., 1986; Meyer et al., 1988), where $\sigma_{H,A}$ is the cross-relaxation rate between an antibody proton and a neighboring bound peptide proton and a is the mole fraction of the free peptide. This simple relationship applies to both one-dimensional and two-dimensional TRNOE experiments. The distance ratios between different pairs of protons in the complex can be determined from $\sigma_{H,A}$, which is inversely proportional to the sixth power of the distance between the protons.

When the chemical shifts of the bound and free peptide resonances are not averaged but the off-rate is still faster than T_1 and $\sigma_{H,A}$ and the peptide is present in large molar excess, the TRNOE intensity between an antibody proton and a free peptide proton is still approximated by the above equation. In the case of medium off-rate, the efficiency of the TRNOE will depend on the off-rate (Clore & Gronenborn, 1983), and the above equation for $N_{H,A}$ can serve only as an upper limit. The off-rate of Ac-VEVPGSQHIDSQKKA-NH_2 bound to TE34 was found to be 480 s^{-1} at 37 °C (Anglister & Zilber, 1990). Although not every resonance represents the average of the bound and free peptide, this off-rate is very fast relative to the T_1 relaxation time and the cross-relaxation rates. Therefore, for a large peptide excess as used in our experiments, the above equation for the TRNOE is expected to be a very good approximation. If on the other hand the exchange between bound and free ligand is slow on the cross-relaxation and spin–lattice relaxation scale, the transferred NOE intensity is vanishingly small.

In this study we determined the amino acid sequence TE34 and assigned the specific peptide–antibody interactions to the specific residues involved. Proton–proton distances were evaluated by comparison of cross-peak intensities and subsequently used to model the three-dimensional structure of the CTP3/TE34 complex. This model provides insight into the mode of interaction between peptides and antibodies and contributes to our understanding of the antigenic structures of peptides.

Materials and Methods

The Fab labeling and sample preparation and the procedure for NOESY measurements were previously described (Anglister et al., 1989). Peptide synthesis and purification and a more detailed description of the 2D difference spectra calculations were also given (Anglister & Zilber, 1990). NMR measurements with the NCA peptide were carried out at 37 °C and pH 7.15, a 10-fold peptide excess being used. With Ac-IDSQRKA the experiment was run at 42 °C with a 2-fold peptide excess. The binding constants of TE34 for Ac-IDSQRKA and for Ac-IDSQKQA were measured at room temperature by following antibody fluorescence quenching as described previously (Anglister & Zilber, 1990).

pH Titration. The TE34 Fab solution in the presence of a 4-fold excess of the CTP3 peptide was titrated between pH 4.00 and pH 9.70. Tryptophan and phenylalanine residues of the antibody were perdeuterated while tyrosines were deuterated at the C_δ positions. Fab concentration was 2 mM. Spectra were measured at 37 °C; 256 scans were collected for each titration point with a spectral width of 6000 Hz and 3-s presaturation of the HDO line with minimal power (55L on the AM 500). The individual resonances of the histidine imidazole protons of the bound and free peptide were clearly observed in the difference spectra calculated for different pH values. pK_a values were determined from plots of the derivative of the titration curves versus the pH.

mRNA Sequencing. The mRNA sequencing was carried out as previously described (Levy et al., 1989). The universal primers for IgG1 heavy chains and κ light chains as defined by Griffiths and Milstein (1985) were used to initiate sequencing. Two additional primers were utilized for both the light chain, 5′-CTC CAC TCT GCT GAT-3′ and 5′-AGA CTG GCC TGG CCT CTG-3′, and the heavy chain, 5′-AGA GGT TTC CAA AGA-3′ and 5′-ACC CTT TCC TGG AGC CTG-3′.

Chain-Specific Labeling. Two Fab fragments were prepared: Fab[H(1)L(1)] was unlabeled, while in Fab[H(2)L(2)] all tryptophan, phenylalanine, and tyrosine residues were predeuterated. Both Fabs were reduced and alkylated, and their chains were separated under denaturing conditions on a Bio-Gel DEAE ion-exchange column at pH 8.3 in 8 M urea and 50 mM glycine–Tris buffer without salt. One of the chains bound to the column and was eluted with the above buffer containing 300 mM sodium chloride. Since the heavy chain fragment and the whole light chain that constitute the Fab share similar molecular weights, they cannot be distinguished on this basis. To differentiate between the separated chains, whole antibody reduced and alkylated was loaded under the same denaturing conditions. Again, a single chain was bound in the absence of salt and later eluted by salt addition. Polyacrylamide gel electrophoresis identified the bound and

nonbound polypeptides as the heavy and light chains, respectively. The solutions of the separated chains were adjusted to pH 7.5 with phosphoric acid and cross-recombined to produce L(1)H(2)- and L(2)H(1)-labeled Fabs. Reconstitution of the native conformation of the Fab was accomplished at 4 °C by extensive dialysis against 10 mM phosphate buffer containing 150 mM sodium chloride, pH 7.5. The reconstituted Fab was then concentrated and dialyzed extensively against phosphate-buffered deuterium oxide, pH 7.15. This procedure significantly improved the yield of reconstituted Fab relative to that previously described (Levy et al., 1989).

Molecular Modeling. The method used to build a detailed all-atom model of the dimer formed from the light and heavy chain variable domains (Fv) of the TE34 antibody paralleled that used for TE32 and TE33 previously (Levy et al., 1989). In the previous paper, we described the modeling method in some detail; here we add some additional details.

Two computer programs, ABMOD and ENCAD, are used to calculate atomic coordinates for TE34. ABMOD finds segments of known antibodies that can be used to provide initial coordinates for TE34. ENCAD uses energy minimization to refine the coordinates. Before different antibody structures can be pieced together with ABMOD and ENCAD, all the structures must be in the same coordinate system. This is achieved by use of rigid-body rotation to superimpose all known antibody coordinates onto the structure of the KOL antibody (Marquart et al., 1980), which is taken as a standard. Only the α-carbon coordinates of the well-conserved framework are used for this superposition.

ABMOD works by first aligning sequences by a modified method that prevents gaps occurring in regions of secondary structure (Lesk et al., 1986) and then by using the atomic coordinates of the known antibody to ensure that chain gaps are between residues that are close in space. Both methods have been described before (Levy et al., 1989), and the final alignment of TE34 to McPC 603 and HyHel-5, the two proteins that ABMOD chose as the best sources of initial coordinates, is shown in Figure 2.

Initial coordinates are simply copied over from the X-ray structures of McPC 603 and HyHel-5 (after rigid-body rotation as described above). These coordinates suffer from two sources of stereochemical errors. (a) There are missing atoms when there is a longer side chain in the unknown antibody structure than in the known structure from which it is built (for example, residue 7 of TE34 is Thr that is modeled from a Ser in McPC 603; see Figure 2). (b) There are gaps in the chain when the unknown structure has a deletion (for example, at residue 31A; see Figure 2).

ENCAD, a robust energy minimization program, is used to correct both these defects and produce a low-energy structure containing all atoms including hydrogen, which is so important for NMR. This program, which has been used in the past to simulate the molecular dynamics of protein in solution (Levitt & Sharon, 1988), uses a detailed all-atom potential energy function, with terms for bond stretching, bond angle bending, bond twisting, van der Waals interactions, and electrostatic forces. Energy parameters and numerical methods are given in Levitt (1983).

Refinement of the initial TE34 model first involved building a model with only polar hydrogen atoms; the energy of the minimized structure is −1593 kcal/mol, which is comparable to that found for TE32 and TE33 (Levy et al., 1989). This minimization was done in vacuo and the effect of the missing solvent compensated for by neutralizing ionized side chains (Glu, Asp, Lys, and Arg) and by holding the C-α coordinates to the initial X-ray position by a weak restraint. Once a low-energy model was obtained with only the polar hydrogen, all other hydrogen atoms were added by ENCAD.

The same energy minimization method was used to model the peptide–antibody interactions. The total potential energy (U) was modified to include a term that restrains the distances between protons observed to interact by NOE. The form of this term is

$$U_{nmr} = \sum k(r_{ij} - r_{ij}^c)^4 \text{ kcal/mol if } r_{ij} \geq r_{ij}^c$$
$$= 0 \text{ if } r_{ij} \leq r_{ij}^c$$

where k is used as a force constant in order to translate the NOE restraints into energy values in the energy minimization calculations. It was necessary to use a fourth power term to ensure continuous first and second derivatives at $r_{ij} = r_{ij}^c$. This restraint keeps the distance between atoms i and j, r_{ij}, less than the distances deduced from the NOE value, r_{ij}^c. The initial model of the peptide was built by hand with the interactive modeling program FRODO (Jones, 1978). Because of the small size of the peptide and the strong restraint imposed by NOE interaction to the rigid antibody binding site, no attempt was made to explore the full range of peptide conformation that could fit the site.

RESULTS

Nucleic and Amino Acid Sequences. The mRNA coding for TE34 was sequenced directly. In Figure 1 the rearranged nucleotide sequences of the light and heavy chains of TE34 are compared with those of TE32 and TE33. The amino acid sequence of TE34 is given in Figure 2, again compared with those of TE32 and TE33. The aromatic residues in the light chain CDRs are Tyr 32L, His 93L, Phe 94L, and Trp 96L. The TE34 light chain sequence is typical of group II κ light chains (Kabat et al., 1987). TE34 uses a MUSJK1 J-gene while TE32 and TE33 use the MUSJK4 J-gene. The sequences of the light chain variable domains of several other antibodies are very similar to that of TE34: L-1210 and 36-60 CRI$^-$, which are anti-azophenylarsonate antibodies (Juszczak et al., 1984) with 96.5% and 95% amino acid identity, respectively; BXW-16, an anti-DNA autoantibody (Kofler et al., 1988) with 96.6% identity; 40–150, an anti-digoxin antibody (Panka et al., 1988) with 95% identity; antidextran antibodies (Akolkar et al., 1987) with 95% identity. The amino acid sequences of TE32 and TE33 are 74% and 75% identical, respectively, with that of TE34. It should be noted that Tyr 32L, although being a CDR residue, is conserved in antibodies possessing κ light chain from group I or group II.

The aromatic residues in the CDRs of the TE34 heavy chain are Tyr 27H, Phe 29H, Tyr 32H, Trp 50H, Tyr 53H, and Phe 100cH. The heavy chain sequence coded by the V_H gene in TE34 is 86% and 89% identical with the corresponding sequences of TE33 and TE32, respectively (Levy et al., 1989). This indicates that all three anti- CTP3 antibodies use V genes of the same family within group IX heavy chain genes. However, the CDR3 of the TE34 heavy chain is only four residues long, three residues shorter than that of TE32 and TE33. The sequence coded by the V_H gene is very homologous with sequences of other antibodies of family IX: DB3, an anti-progesterone antibody (Winter et al., 1985) with 85% identity; L6, an anti-carcinoma-antigen antibody (Liu et al., 1987) with 89% identity; U22B5, an anti-arsonate antibody (Meek et al., 1984), with 84% identity. All aromatic residues in the part coded by the V_H gene are conserved in those antibodies which are assumed to be coded by genes of the same family. TE34 uses the MUSJH2 J-gene in contrast to TE32

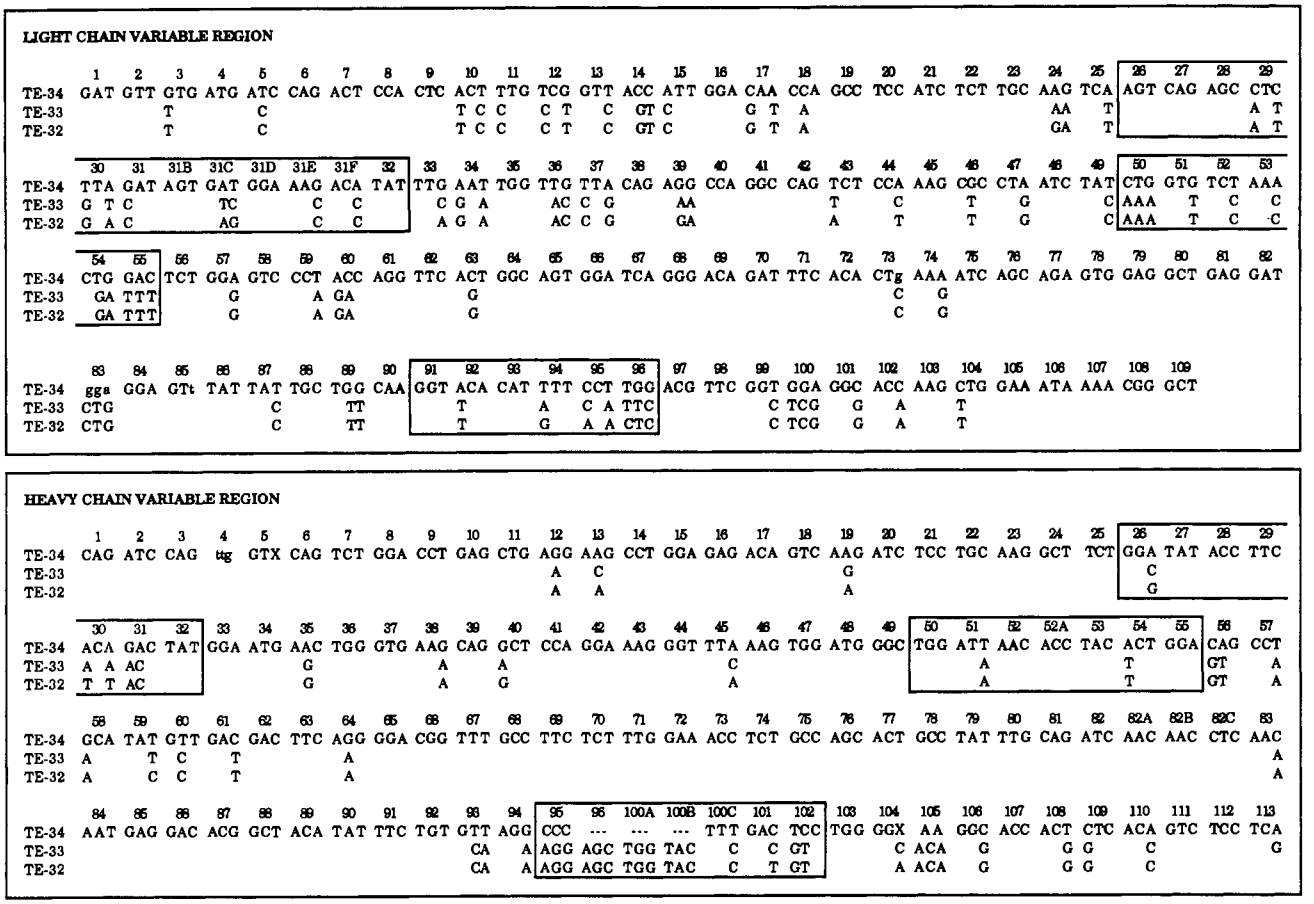

FIGURE 1: Deoxyribonucleotide sequence of the TE34 light and heavy chain variable regions as compared to those previously determined for the TE33 and TE32 antibodies. The boxes delineate boundaries of the CDRs. The full sequence is given for TE34, whereas only changes with respect to TE34 are listed for TE32 and TE33. Lower-case letters where they appear indicate some ambiguity in the sequence determination; X appears where we could not determine the nucleotide.

and TE33 which use the MUSJH1 J-gene. The third CDR varies considerably in both length and sequence in all the above antibodies.

Assignment of Intra- and Intramolecular Interactions to a Specific Lysine Residue. Two tyrosine residues and at least one phenylalanine residue from TE34 interact with protons of the two lysine residues of CTP3 (Anglister & Zilber, 1990). There is also an intramolecular interaction in the bound peptide between a lysine residue and aspartate-10. To assign these interactions to a specific lysine residue, two CTP3 analogues were prepared: Ac-IDSQRKA and Ac-IDSQKQA. These truncated peptides correspond to residues 9–15 of CTP3, and for convenience, we maintain the same residue numbering. Ac-IDSQRKA, in which lysine-13 was substituted by arginine, binds to TE34 with an affinity constant of 2×10^5 M^{-1} in comparison to $K_a = 2.5 \times 10^5$ M^{-1} measured for Ac-IDSQKKA. This similarity in binding indicates that substitution of lysine-13 by arginine does not significantly interfere with the peptide–antibody interactions. The binding of the peptide Ac-IDSQKQA in which lysine-14 was replaced by glutamine was too weak to be measured (less than 10^3 M^{-1}). Thus, it appears that the positive charge on lysine-14 contributes significantly to the binding constant.

Figure 3 presents the 2D TRNOE difference spectrum showing interactions between nonaromatic protons of the TE34 Fab/Ac-IDSQRKA complex. The assignment of cross peaks in this spectrum is based on chemical shift identity with those of the free peptide protons, the COSY spectrum of the peptide, and their observed multiplicity. The resonance at 3.02 ppm has the same chemical shift for the C_ϵ protons of lysine as the unmodified peptide and therefore is assigned to lysine-14. The resonance at 3.22 ppm, which does not appear in the difference spectrum for the TE34/Ac-IDSQKKA complex (Anglister & Zilber, 1990), is assigned to the C_δ protons of arginine-13, and the cross peaks at (3.22, 1.87), (3.22, 1.78), and (3.22, 1.64) ppm are due to intraresidue interactions with the β- and γ-protons, respectively. The weak cross peaks at (3.02, 2.66) and (3.02, 2.71) ppm are assigned to intra-bound-peptide interactions between lysine-14 ϵ-protons and aspartate-10 β-protons. The strong cross peaks at (3.02, 2.56) and (3.02, 2.40) ppm are due either to interactions of peptide lysine-14 with the antibody or to exchange cross peaks. The strong cross peak at (3.22, 2.81) ppm is similar to that previously observed at (3.01, 2.81) ppm for TE34 interacting with AcIDSQKKA. It is assigned to the arginine-13 δ-protons and is due to exchange or interaction with antibody protons. Additional cross peaks which are observed in the TRNOE difference spectrum are due to interresidue interactions of the $C_\alpha H$ of serine-11 (4.38 ppm) with aspartate-10 $C_\alpha H$ (4.62 ppm) and $C_\beta H$ (2.66 and 2.71 ppm) and with isoleucine-9 $C_\alpha H$ (4.14 ppm). The interresidue interactions in the bound peptide are indicative of its conformation and provide distance restraints that are input into our modeling procedure.

NOESY measurements with the modified peptide (K13 → R) proved crucial in assigning the interactions of TE34 phenylalanine and tyrosine residues to a specific lysine residue of the peptide. Figure 4A shows the interactions of phenylalanine residues of the antibody with the Ac-IDSQRKA peptide. Several cross peaks are unambiguously assigned to their interacting peptide protons. The cross peaks at (3.20, 6.63), (1.47, 6.38), and (1.32, 6.58) ppm are assigned respectively to the interactions of arginine-13 ($C_\delta H$), lysine-14

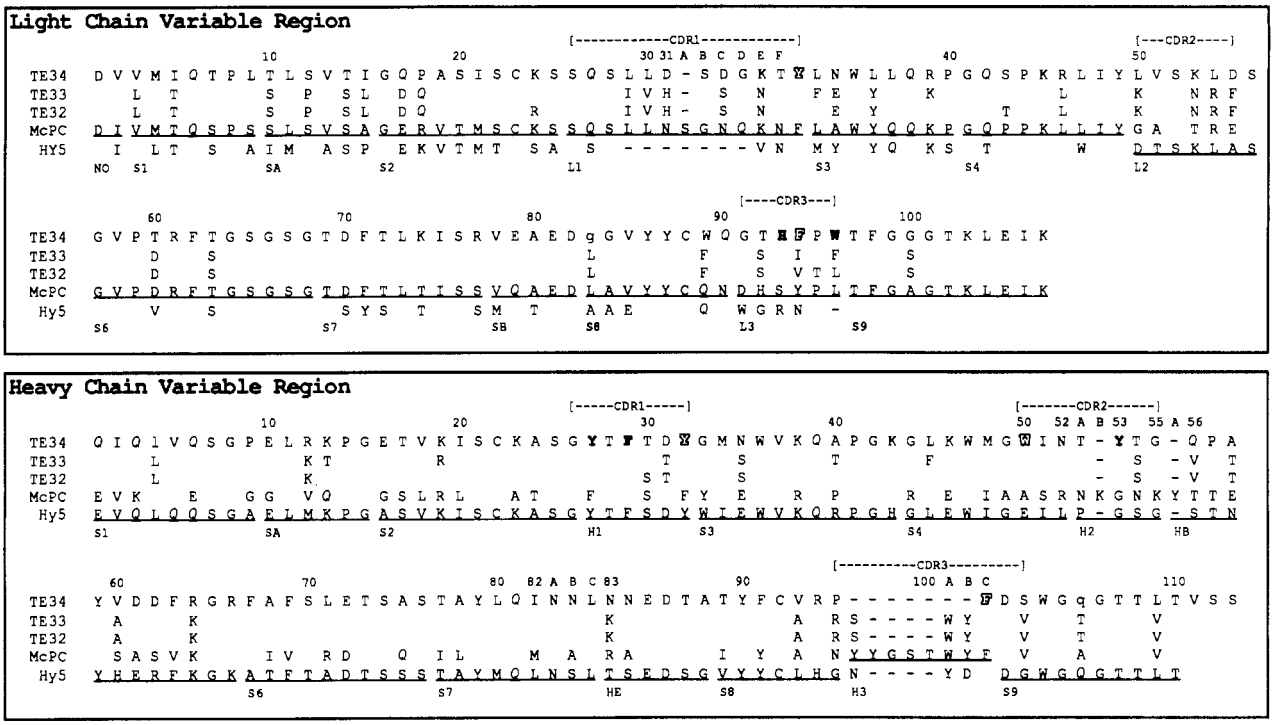

FIGURE 2: Comparison of the amino acid sequence of the variable regions of the heavy and light chains of the TE34 antibody with the known sequences for TE32 and TE33. The full sequence is given for TE34, whereas only changes with respect to TE34 are listed for TE32 and TE33. The aromatic residues in the CDRs of TE34 are typeset in boldface, while the subset of these residues found in this study to interact with the peptide antigen is outlined. Also listed are the sequences of both McPC 603 and HyHel-5, whose three-dimensional structures were used to calculate the structure of TE34 (Segal et al., 1974; Sheriff et al., 1987). McPC is a myeloma antibody that binds phosphocholine while HyHel-5 is an anti-lysozyme monoclonal antibody. The particular sequences used to provide initial coordinates for segments of the TE34 model are underlined. Each segment used for modeling is marked by a two-letter segment name placed at the start of that particular segment. Lower-case letters indicate places of ambiguity in sequence determination.

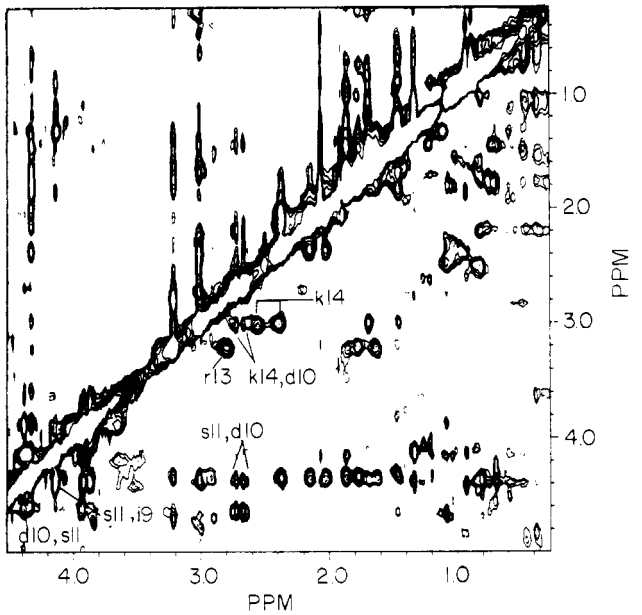

FIGURE 3: Difference between the NOESY spectrum of TE34 Fab with a 2-fold excess of acetyl-IDSQRKA and that of the peptide-saturated Fab. The spectrum contains cross peaks due to intramolecular interactions in the bound peptide and possibly intermolecular interactions between the peptide and the antibody. Interacting peptide protons are marked by small letters and numbers denoting their location in the peptide sequences.

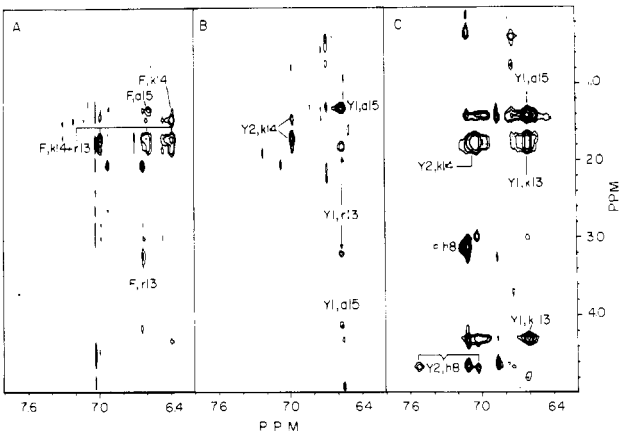

FIGURE 4: 2D TRNOE difference spectrum showing interactions of aromatic protons of the antibody (TE34) with Ac-IDSQRKA and with Ac-VEVPGSQHIDSQKKA-NH_2 (NCA). (A) Interactions of antibody phenylalanine with Ac-IDSQRKA. Antibody tyrosine and tryptophan are perdeuterated while phenylalanine is unlabeled. (B) Interactions of $C_\epsilon H$ of TE34 tyrosine residues with Ac-IDSQRKA. Antibody tryptophan and phenylalanine are perdeuterated while tyrosine residues are deuterated at the C_δ positions. (C) Interactions of tyrosine $C_\beta H$ and $C_\delta H$ of the antibody with the NCA peptide. Antibody tryptophan and phenylalanine residues are perdeuterated while tyrosine residues are unlabeled. Assigned peptide residues are marked by small letters and their location in the sequence; antibody residues are marked by capital letters denoting amino acid type and arbitrary numbers.

($C_\gamma H$), and alanine-15 (CH_3) with antibody phenylalanine. The strong cross peaks at (1.72, 6.39), (1.72, 6.60), and (1.72, 6.98) ppm are assigned to interactions of lysine-14 δ-protons with antibody phenylalanine protons. The cross peaks at (1.89, 6.38), (1.84, 6.58), and (1.84, 6.98) ppm could be assigned to interactions of β-protons of either arginine-13 or lysine-14 with antibody phenylalanine. Figure 4B shows the interaction of $C_\epsilon 1$ and $C_\epsilon 2$ protons of antibody tyrosine with the Ac-IDSQRKA peptide. One tyrosine residue (Y1) interacts with arginine-13 β- and δ-protons while a second tyrosine (Y2) interacts with lysine-14 β, γ, and δ-protons. The arginine and lysine cross peaks were assigned on the basis of the COSY

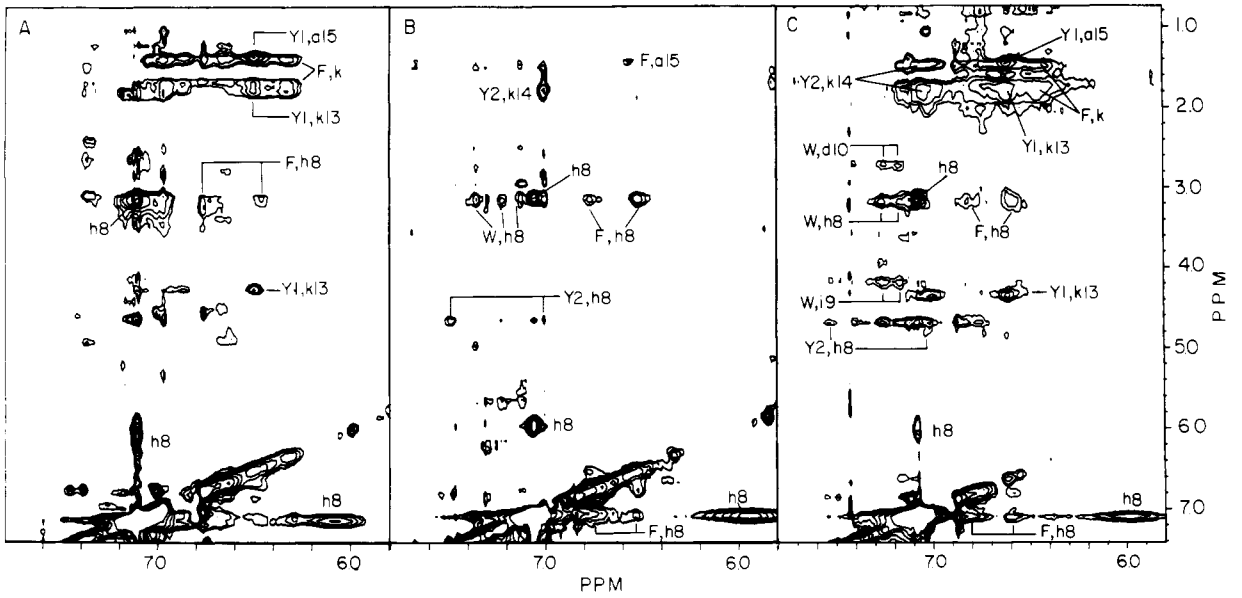

FIGURE 5: 2D TRNOE difference spectra obtained after specific chain labeling: (A) interactions of the light chain aromatic residues with the NCA peptide; (B) interactions of heavy chain aromatic residues with the NCA peptide; (C) interactions of aromatic residues from both chains with the NCA peptide. Assigned antibody residues are marked by capital letters and arbitrary numbers; peptide protons are marked by small letters and numbers denoting the location in the sequence of the residues to which the protons belong.

spectrum of the modified peptide with and without Fab. The cross peak at (4.15, 6.55) ppm is assigned to interaction of Y1 and alanine-15 $C_\alpha H$.

Interactions of C_δ Protons of Antibody Tyrosine Residues with the NCA Peptide. To observe the interactions of C_δ as well as C_ϵ protons of antibody tyrosine with the NCA peptide (acetyl-VEVPGSQHIDSQKKA-NH_2), which contains the complete epitope recognized by TE34, we prepared Fab in which all tryptophan and phenylalanine residues were perdeuterated while tyrosine was unlabeled. A section of the 2D TRNOE difference spectrum obtained for the labeled Fab in complex with NCA is presented in Figure 4C. The interactions assigned to lysine-14 appear in both panels B and C of Figure 4; however, interactions with arginine-13 in Figure 4B are replaced by interactions with lysine-13 in Figure 4C. The cross peaks in Figure 4C manifest much stronger intensities due to the much faster off-rate of NCA relative to that of the truncated peptide which contains a negatively charged C-terminus. Figure 4C also illustrates the chemical shift degeneracy of both lysine resonances in the NCA peptide.

The assignment of cross peaks to tyrosine C_δ protons is based on their absence in the difference spectrum obtained for the Fab in which tryptophan and phenylalanine are perdeuterated while tyrosine residues are deuterated at the C_δ positions (Anglister & Zilber, 1990). The assignment of the peptide protons involved was previously reported. The cross peaks at (4.70, 7.52) and (4.70, 7.03) ppm are assigned to the interaction of the $C_\alpha H$ of the peptide histidine-8 with $C_\delta H$ and $C_\epsilon H$, respectively, of tyrosine (Y2). The assignment of these two cross peaks to a single tyrosine residue relies on the specific chain labeling and the calculated model for the antibody Fv (see below). Weak interactions of Y1 and Y2 with the $C_\epsilon H$ of lysine-13 and lysine-14 are observed, (2.99, 6.60) and (2.99, 7.02) ppm, respectively. The strong cross peak at (4.32, 6.60) ppm was previously assigned to the interaction of Y1 with a peptide lysine (Anglister & Zilber, 1990). On the basis of the observed interaction between Y1 and alanine-15 $C_\alpha H$ in Figure 4B, we conclude that the cross peak at (4.32, 6.60) ppm in Figure 4C is actually due to overlap of two cross peaks, one due to interaction with lysine-14 and one due to interaction with alanine-15. The C-terminal alanine proton has a chemical shift of 4.15 ppm in Ac-IDSQRKA and 4.33 ppm in Ac-VEVPGSQHIDSQKKA-NH_2 due to the conversion of the C-terminal into an amide. The weak cross peaks at (1.76, 6.73), (1.44, 6.73), (1.44, 6.96), (1.76, 6.96), and (1.44, 7.10) ppm are due to interactions of tyrosine C_δ protons of Y1 and Y2 and/or weak interactions of a third antibody tyrosine residue with the peptide.

Assignment of Interactions to a Specific Antibody Polypeptide Chain. In order to assign the peptide–antibody interactions to the specific polypeptide chain involved, two specifically labeled recombined Fab fragments were prepared. In Fab[L(1),H(2)], tryptophan, tyrosine, and phenylalanine residues of the heavy chain are perdeuterated while the aromatic residues of the light chain are unlabeled. In Fab[L-(2),H(1)], the aromatic residues of the light chain are perdeuterated while those of the heavy chain are unlabeled. The interactions of the light and heavy chain aromatic residues with the NCA peptide are shown in panels A and B of Figure 5, respectively. Inspection of Figure 5A,B reveals a single phenylalanine and at least one tyrosine residue of the light chain interacting with the peptide, with the heavy chain contributing one each of tyrosine, phenylalanine, and tryptophan. In Figure 5A the cross peaks at (1.44, 6.97), (1.76, 6.97), (1.44, 7.07), (1.76, 7.07) ppm were not assigned. These cross peaks appear also in Figure 4B and were attributed to interactions of C_δ protons of Y1 and Y2 and/or interactions of protons of a third antibody tyrosine with the peptide. According to Figure 5A these protons are of a light chain tyrosine. The sum of the two TRNOE difference spectra in Figure 5A,B is in excellent agreement with the spectrum previously obtained for the native unlabeled Fab (Figure 5C).

A Calculated Model for the Fv. The three-dimensional structure of the Fv of TE34 was modeled on the basis of the known conformations of the other antibodies. The different segments used in the modeling procedure are given in Figure 2. The conformations of the TE34 segments L1, L2, and H1 were found to be very similar to those of the corresponding segments in TE32 and TE33. The L3 segment in TE34 and in both TE32 and TE33 is similar in sequence and conformation between residues 91 and 93 but differs between residues 94 and 96. The conformation of H2 (residues 52a–55) in the

model of TE34 differs greatly from that of H2 in TE32 and TE33. This difference probably arises from the application of the HyHel-5 antibody coordinates for TE34 modeling (Padlan et al., 1989), as opposed to the use of J539 (Suh et al., 1986) in the modeling of both TE32 and TE33. The H2 segment has the sequences TYSG in TE32 and TE33, PDSG in J539, TYTG in TE34, and PGSG in HyHel-5. According to recent analysis of the conformations of antibody hypervariable loops (Chothia et al. 1989), Hyhel-5 is more suitable for modeling H2 in TE32, TE33, and TE34. In TE34 the H2 segment is located on the periphery of the combining site and therefore does not affect the structure of its central part. The second CDR of the TE34 heavy chain includes residues 50–52 as framework residues. In contrast to the differences noted for residues 52a–55, the main-chain conformation for residues 50–52 and the side-chain conformations for 50 and 51 are very similar in TE32, TE33, and TE34 and also to the conformation observed in other antibodies. Interestingly, residue Trp 50H occupies a central position in the combining site of all three antibodies.

The H3 segment, which is part of the third CDR of the heavy chain, contains only one residue in TE34 and four residues in both TE32 and TE33. This truncation totally alters the combining site structure of TE34. In TE32 and TE33 the sequence Ser 96, Trp 100a, Tyr 100b, not found in TE34, is at the tip of the loop created by CDR3 of the heavy chain, forming a protuberance in the center of the combining site. The empty space formed by shortened CDR3 in TE34 creates a pocket in the center of the combining site, the walls of which are composed of a continuum of aromatic side chains and one hydrophobic nonaromatic residue. All of these, Tyr 32L, Tyr 49L, Leu 50L, Phe 94L, Trp 96L, Tyr 32H, Trp 50H, and Phe 100cH, are exposed to the solvent. In contrast, the periphery of the combining site is composed mostly of nonaromatic residues.

On the basis of our calculated model and the assignment of the NOE interactions to their specific polypeptide chain, we are able to assign peptide–antibody interactions to their specific residues in the TE34 sequence: Tyr 32L interacts with alanine-15 and lysine-13; Phe 94L interacts with histidine-8, lysine-13, and lysine-14; Tyr 32H interacts with lysine-14 and histidine-8; Trp 50H interacts with histidine-8, isoleucine-9, and aspartate-10; Phe 100cH interacts with histidine-8 and alanine-15. Tyr 49L is in the periphery of the combining site and could in principle interact with alanine-15 and lysine-14. It is possible that the unassigned cross peaks at (1.44, 6.97), (1.76, 6.97), (1.44, 7.07), (1.76, 7.07), and (1.86, 7.07) ppm observed in Figures 4B and 5A are due to interactions of Tyr 49L.

Derivation of Restraints on Intermolecular Proton–Proton Distances. To translate cross-peak intensities into restraints on proton–proton distances, one needs a calibration usually obtained by measuring the intensity of a cross peak between a pair of proton separated by a known and fixed distance. To obtain such calibration, we measured the difference spectrum between the NOESY spectrum of Fab in the presence of 10-fold peptide excess and the NOESY spectrum of the uncomplexed Fab (without spectrum). All tryptophan and phenylalanine residues of the antibody were deuterated while tyrosine residues were unlabeled. The aromatic section of the difference spectrum is shown in Figure 6. In addition to the transferred NOE cross peaks, this difference spectrum shows the cross peaks due to intermolecular interactions of antibody protons that changed their chemical shift upon peptide binding. In the section of the NOESY difference spectrum showing

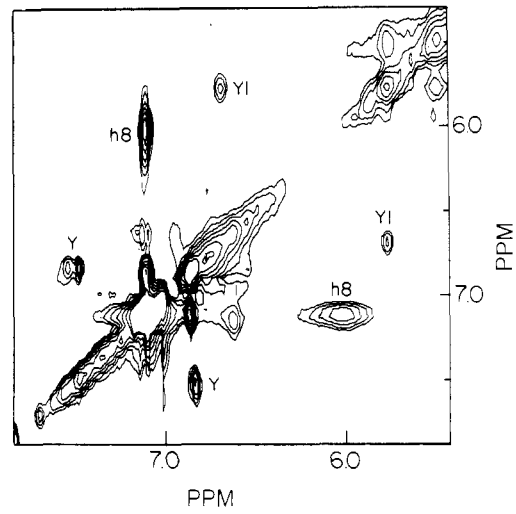

FIGURE 6: Aromatic section of the 2D difference spectrum between the NOESY spectra of the TE34 Fab with a 10-fold excess of NCA and the uncomplexed Fab. Antibody tryptophan and phenylalanine residues are perdeuterated. The cross peaks marked Y1 and Y are due to intraresidue interactions between the tyrosine C_ϵ and the C_δ protons; h8 are the exchange cross peaks of the peptide histidine $C_{\delta_2}H$.

interactions between aromatic protons, one observes cross peaks due to intraresidue interactions between the tyrosine ϵ_1- and ϵ_2-protons and the δ_1- and δ_2-protons. These cross peaks are not observed when the C_δ protons of the antibody tyrosine residues are deuterated. A cross peak at (5.76, 6.68) ppm with its symmetrical counter cross peak is very well resolved, and another pair of cross peaks, (6.85, 7.48) and (6.85, 7.54) ppm, is the result of overlap between two tyrosine residues. The other cross peaks due to intra-tyrosine interactions are too close to the diagonal or overlap the resonances of the $C_{\delta_2}H$ of the bound and free peptide histidine. The intensity of the resolved cross peaks is then compared to the intensity of the exchange cross peak of $C_{\delta_2}H$ of the peptide histidine. Since this cross peak appears in all the TRNOE difference spectra and all of them are measured at the same mixing time and with the same molar ratio of peptide excess, it can serve as an internal standard for calibrating the intensities in all TRNOE difference spectra. Through this internal standard one gets the ratio between cross-peak intensities in the TRNOE difference spectrum and the cross peak due to intra-tyrosine interactions and subsequently the ratio between the distances of the corresponding pairs of protons. This analysis indicates that all the cross peaks due to intermolecular interactions observed in the TRNOE difference spectra are due to strong interactions between protons that are less than 3.5 Å apart. Measurements of the TRNOE difference spectra for the TE33 antibody at mixing times of 60 and 100 ms exclude the possibility that the cross peaks observed in the TRNOE difference spectra at 100 ms are due to spin diffusion.

Docking of the Peptide into the Antibody Combining Site. The list of the distance restraints used in the modeling of the TE34/CTP3 complex is given in Table I. C_ζ was used as a pseudoatom for tyrosine $C_\epsilon H$ and $C_\delta H$ protons. Since as shown in Figure 5B three tryptophan protons are involved in interactions with the peptide, we expect C_{ζ_3} and C_{η_2} to be two of these protons. For phenylalanine interactions C_ζ was chosen as a replacement for the pseudoatom. Each residue of the epitope was independently docked into the combining site to conform approximately to the NMR restraints while reasonable lengths for peptide bonds were maintained. The model for the complex was then refined under 35 NMR restraints, which are given in Table I. Initially, only the peptide was

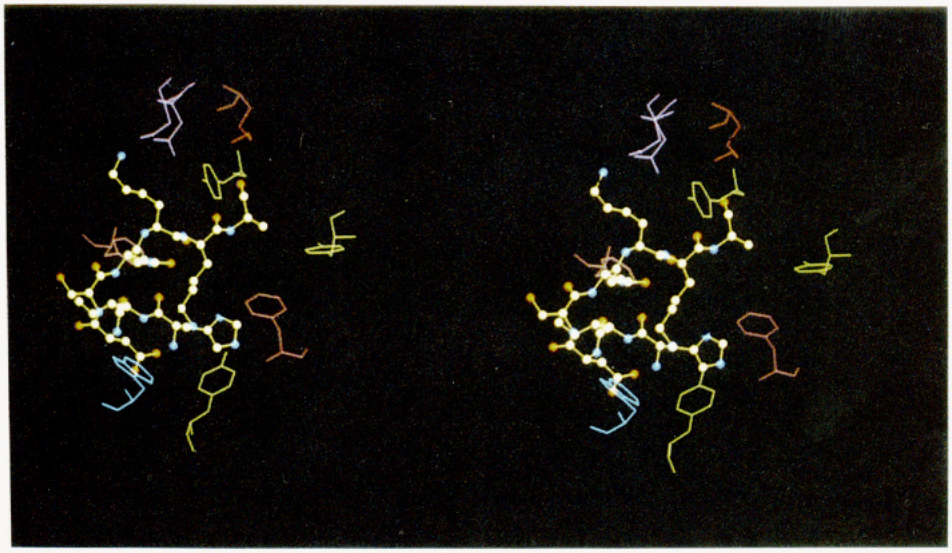

FIGURE 7: Structure of the TE34 complex with CTP3 (residues 8–15) obtained after energy minimization and on the basis of a calculated model for antibody Fv with NMR restraints: light blue, Trp 50H; light green, Tyr 32L, Tyr 49L, and Tyr 32H (top to bottom); pink, Phe 94L and Phe 100cH (left to right); red, Lys 31 eL; violet, Asp 31L and Asp 31cL. Atoms of residues 8–15 of CTP3 are represented by balls: orange, oxygen; light blue, nitrogen; yellow, carbon.

allowed to move, whereas in a further refinement the antibody was allowed to move as well under a weak force field acting to maintain its original conformation. The root-mean-square (rms) deviation between distances in the final model (Figure 7) and the NMR restraints is 0.83 Å. The deviations between the model and NMR distances are greatest for interactions involving Trp 50H (1.25 Å) and for intrapeptide interactions (1.01 Å). When these distances are excluded, the rms deviations for the remaining restraints drops to 0.58 Å. When the force constants attributed to the NOE restraints are increased by a factor of 5, the rms deviation between the restraints and the model drops to 0.60 Å. Decreasing or increasing this force constant effects the orientation of Tyr 32L, Phe 94L, and Trp 50H; however, these changes have a very minor effect on the calculated conformation of the peptide. The large violations could be explained by possible inaccuracies in the orientation of the side chains of the residues in the CDRs and the relative position of the hypervariable loops (Chothia et al., 1989) and especially of CDR2 of the heavy chain. These inaccuracies are only partially rectified by the relatively small number of restraints obtained by the NOESY experiments. It should be noted that although the interactions of Tyr 49L were not included in our restraints, in the calculated model there are interproton distances between Tyr 49L and alanine-15 methyl protons and lysine-14 β- and δ-protons which are less than 5 Å.

Discussion

The model for the combining site structure of the TE34/CTP3 complex, obtained by applying NMR-derived interproton distance restraints to the calculated Fv model, is shown in Figure 7. Only peptide residues 8–15 are shown since they represent the full peptide epitope. This segment of the peptide is highly polar, containing five charged residues: histidine-8, aspartate-10, lysine-13, lysine-14, and alanine-15 at the C-terminus. Despite the polarity of the peptide antigen, most of the antibody surface that it contacts is hydrophobic. The binding of the two molecules is made possible by the special folding of the bound peptide: (a) Formation of a β-turn manifested by a hydrogen bond between the isoleucine-9 oxygen and glutamine-12 NH. The formation of a β-turn by the peptide segment IDSQ increases the number of residues interacting with the central part of the antibody combining site, thus potentially increasing its binding energy. (b) Interactions with the aromatic residues of the antibody primarily via the peptide's backbone and β-protons. (c) Intramolecular electrostatic interactions in the bound peptide between the side chains of aspartate-10 and lysine-14, which orient their charges away from the hydrophobic surface of the antibody. There are also additional electrostatic interactions between the peptide and charged residues at the periphery of the combining site: (a) The negatively charged C-terminal carboxylate interacts with the ϵ-NH_3^+ of Lys 31eL. This interaction contributes significantly to the binding energy of CTP3 as manifested by the 2 orders of magnitude decrease in its binding constant when this carboxylate is converted into an amide (Anglister & Zilber, 1990). (b) According to the calculated model, the ϵ-NH_3^+ of lysine-13 interacts with the negatively charged γ-carboxyl groups of Asp 31L and Asp 31cL. Lysine-13 is exposed to the solvent and projects toward the periphery of the combining site. Its β- and δ-protons interact with Tyr 32L and Phe 94L. There are also polar interactions between the ϵ-NH_3^+ of lysine-14 and the Tyr 32H O_η. Histidine-8 interacts with Phe 100cH, Tyr 32H, and Trp 50H; however, the nitrogen atoms of its imidazole ring are exposed to the solvent. The pK values of the bound and free peptide histidine (pK = 6.5) are practically indistinguishable, providing support for the partial exposure to the solvent of histidine-8. The side chains of isoleucine-9, serine-11, and glutamine-12 project toward the solvent; however, the α- and β-protons of isoleucine-9 interact with the antibody. The total area of contact between the TE34 antibody and residues 8–15 of CTP3 was calculated to be 388 Å^2.

In the models of the three anti-CTP3 antibodies investigated, the surface of the antibody in contact with the peptide consists mainly of aromatic side chains, with five in both TE32 and TE34 and six in TE33. Interestingly, Tyr32H, Tyr32L, and Trp50H are found to interact with CTP3 in all three antibodies; however, they interact with different peptide residues in the TE34/CTP3 complex. Moreover, these aromatic residues are conserved in antibodies coded by the same V_H gene family which differ in their specificities. The major difference in the models of TE32, TE33, and TE34 is the shape of the combining sites of the three antibodies: a shallow groove partly

Table I: NMR Distance Constraints Used To Model the TE34 Complex with CTP3 (Residues 8–15)

atom i^a	atom j^a	distance constraint (Å)[b]	distance in the model (Å)[c]	deviation (Å)[d]
his 8 Hα	Tyr 32H Cζ	5.0	5.3	0.3
his 8 Hα	Tyr 32H Cγ	5.0	5.8	0.8
his 8 Cβ	Trp 50H $H\zeta_2$	3.5	5.0	1.5
his 8 Cβ	Trp 50H Hη	3.5	3.9	0.4
his 8 Cβ	Phe 94L Cζ	5.5	6.3	0.8
his 8 Cβ	Phe 96H $H\epsilon_2$	3.5	4.0	0.5
his 8 $H\delta_2$	Phe 96H $H\epsilon_2$	4.0	4.9	0.9
ile 9 Hα	ser 11 Hα	3.0	4.4	1.4
ile 9 Hα	Trp 50H Hη	2.5	4.4	1.9
ile 9 Hα	Trp 50H $H\zeta_2$	2.5	3.9	1.4
asp 10 Hα	ser 11 Hα	3.0	4.4	1.4
asp 10 $H\beta_1$	ser 11 Hα	4.5	4.8	0.3
asp 10 $H\beta_2$	ser 11 Hα	4.5	4.3	0
asp 10 Hα	Trp 50H $H\zeta_2$	2.5	3.8	1.3
asp 10 Hα	Trp 50H Hη	2.5	3.5	1.0
asp 10 Cβ	Trp 50H Hη	3.5	3.8	0.3
asp 10 Cβ	Trp 50H $H\zeta_2$	3.5	4.8	1.3
asp 10 $H\beta_1$	lys 14 Cε	4.5	5.1	0.6
asp 10 $H\beta_2$	lys 14 Cε	4.5	3.7	0
lys 13 Hα	Tyr 32L Cζ	4.5	5.3	0.8
lys 13 $H\beta_1$	Tyr 32L Cζ	4.0	2.9	0
lys 13 $H\beta_2$	Tyr 32L Cζ	4.0	4.0	0
lys 13 Cγ	Tyr 32L Cζ	5.5	4.6	0
lys 13 Cδ	Tyr 32L Cζ	6.0	5.9	0
lys 13 Cε	Tyr 32L Cζ	6.5	6.7	0.2
lys 13 $H\beta_1$	Phe 94L Cζ	4.5	4.8	0.3
lys 13 $H\beta_2$	Phe 94L Cζ	4.5	5.8	1.3
lys 13 Cδ	Phe 94L Cζ	5.5	5.7	0.2
lys 14 Cγ	Phe 94L Cζ	5.5	6.5	1.0
lys 14 Cγ	Tyr 32H Cζ	6.0	6.3	0.3
lys 14 Cδ	Tyr 32H Cζ	5.5	5.6	0.1
lys 14 Cε	Tyr 32H Cζ	5.5	4.8	0
ala 15 Hα	Tyr 32L Cζ	4.5	5.4	0.9
ala 15 Cβ	Tyr 32L Cζ	5.5	5.2	0
ala 15 Cβ	Phe 96H Hζ	4.5	5.0	0.5

[a] Peptide residues are written with lower-case letters while antibody residues begin with capital letters. H, heavy chain residues; L, light chain residues. [b] The given value reflects the maximum distance from the NOE restraints and additional distance due to the use of pseudo-atoms (Wüthrich et al., 1983). [c] The distances in the final model obtained by energy minimization calculations. [d] The deviation between the restraint and the actual distance in the model. When the latter is smaller than the restraint, the deviation is set equal to zero.

surrounding Trp 100aH in TE32 and TE33 and a deep cavity in TE34 which lacks Trp 100aH. This cavity is the result of the deletion in TE34 of the sequence Ser 96H, Trp 100aH, Tyr 100bH which is found in the CDR3 of the heavy chain of both TE33 and TE32. These observations support our previous conclusion that aromatic residues inside the antibody combining site create a general "active" hydrophobic surface, while other residues, and especially variations in the length of the CDRs, modulate its polarity and shape to fit a specific antigen (Levy et al., 1989).

On the basis of our models for the three anti-CTP3 antibodies, a distinction can be made between two regions on the surface created by the six CDRs: (a) the center, which is mostly composed of the exposed side chains of aromatic residues; (b) the periphery, where a few aromatic residues are interspersed throughout predominantly nonaromatic residues. The interactions of our peptide occur mainly with the center, which forms the combining site. In contrast, in antibody–protein complexes, a considerable fraction of the interactions occurs with peripheral residues, a considerable fraction of the interactions occurs with peripheral residues, although calculations have shown that residues in the center of the combining site contribute the most to binding energy (Novotny et al., 1989). It remains to be determined whether the aromatic core described is a feature common to a large population of antibodies and whether this core persists in antibodies against other antigenic determinants dominated by polar and charged residues.

Acknowledgments

We thank Professor Ruth Arnon and Dr. Chaim Jacob for the hybridomas, Professor Fred Naider for very helpful discussions, Mrs. Rina Levy for excellent technical assistance, Mrs. Ruth Sharon for help with the computer programs, and Dr. Ora Goldberg for peptide synthesis.

Registry No. CTP3, 89157-28-8; *N*-acetyl-IDSQKKA, 124042-27-9.

References

Akolkar, P. N., Sikder, S. K., Bhattacharya, S. B., Liao, J., Gruezo, F., Morrison, S. L., & Kabat, E. A. (1987) *J. Immunol. 138*, 4472–4479.

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

Anglister, J., & Zilber, B. (1990) *Biochemistry 29*, 921–928.

Anglister, J., Bond, M. W., Frey, T., Leahy, D., Levitt, M., McConnell, H. M., Rule, G. S., Tomasello, J., & Whittaker, M. (1987) *Biochemistry 26*, 6058–6064.

Anglister, J., Jacob, C., Assulin, O., Ast, G., Pinker, R., & Arnon, R. (1988) *Biochemistry 27*, 717–724.

Anglister, J., Levy, R., & Scherf, T. (1989) *Biochemistry 28*, 3360–3365.

Bilofsky, H. S., & Burks, C. (1988) *Nucleic Acids Res. 16*, 1861–1864.

Chothia, C., Lesk, A. M., Levitt, M., Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., & Poljak, R. J. (1986) *Science 233*, 755–758.

Chothia, C., Lesk, A. M., Tramontano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulip, W. R., Colman, P. M., Spinelli, S., Alzari, P. M., & Poljak, R. J. (1989) *Nature 342*, 877–883.

Clore, G. M. & Gronenborn, A. M. (1983) *J. Magn. Reson. 53*, 423–442.

Clore, G. M., Gronenborn, A. M., Carlson, G., & Meyer, E. F. (1986) *J. Mol. Biol. 190*, 259–267.

Colman, P. M., Laver, W. G., Varghese, J. N., Baker, A. T., Tulloch, P. A., Air, G. M., & Webster, R. G. (1987) *Nature 326*, 358–363.

Griffiths, G. M., & Milstein, C. (1985) in *Hybridoma Technology in the Biosciences and Medicines* (Springer, T. A., Ed.) pp 103–115, Plenum Publishing Corp., New York.

Jacob, C. O., Sela, M., & Arnon, R. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 7611–7615.

Jones, T. A. (1978) *Acta Crystallogr., A34*, 827–828.

Juszczak, E., Near, R. I., Gefter, M. L., & Margolies, M. N. (1984) *J. Immunol. 133*, 2603.

Kabat, E. A., Wu, T. T., Reid-Miller, M., Perry, H. M., & Gottesman, K. S. (1987) *Sequences of Proteins of Immunological Interest*, 4th ed., National Institutes of Health, Bethesda, MD.

Kofler, R., Strobel, R., Balderas, R. S., Johnson, M. E., Noonan, D. J., Duchosal, M. A., Dixon, F. J., & Theofilopoulos, A. N. (1988) *J. Clin. Invest. 82*, 852–860.

Lesk, A., Levitt, M., & Chothia, C. (1986) *Protein Eng. 1*, 77–88.

Levitt, M. (1983) *J. Mol. Biol. 170*, 723–764.

Levitt, M., & Sharon, R. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 7557–7561.

Levy, R., Assulin, O., Scherf, T., Levitt, M., & Anglister, J. (1989) *Biochemistry 28*, 7168-7175.
Liu, A. Y., Robinson, R. R., Hellström, K. E., Murray, E. D., Jr., Chang, P., & Hellström, I. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 3439-3443.
Marquart, M., Deisenhofer, J., Huber, R., & Palm, W. (1980) *J. Mol. Biol. 141*, 369-391.
Meek, K., Jeske, D., Slaoui, M., Leo, O., Urbain, J., & Capra, J. D. (1984) *J. Exp. Med. 160*, 1070-1086.
Meyer, E. F., Clore, G. M., Gronenborn, A. M., & Hansen, H. A. S. (1988) *Biochemistry 27*, 725-730.
Novotny, J., Bruccoleri, R. E., & Saul, F. A. (1989) *Biochemistry 28*, 4735-4749.
Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J., & Davies, D. R. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 5938-5942.
Panka, D. J., Mudgett-Hunter, M., Parks, D. R., Peterson, L. L., Herzenberg, L. A., Haber, E., & Margolies, M. N. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 3080-3084.
Saul, R., Amzel, L., & Poljak, R. (1978) *J. Biol. Chem. 253*, 585-597.
Segal, D., Padlan, E., Cohen, G. H., Rudikoff, S., Potter, M., & Davies, D. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 4298-4302.
Sheriff, S., Silverton, E. W., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., & Davies, D. R. (1987) *Proc. Natl. Acad. Sci. U.S.A. 84*, 8075-8079.
Suh, S. W., Bhat, T. N., Navia, M. A., Cohen, G. H., Rao, D. N., Rudikoff, S., & Davies, D. (1986) *Proteins: Struct., Funct., Genet. 1*, 74-80.
Winter, E., Radbruch, A., & Krawinkel, U. (1985) *EMBO J. 4*, 2861-2867.
Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
Wüthrich, K., Billeter, M., & Braun, W. (1983) *J. Mol. Biol. 169*, 949-961.

NMR Identification of Protein Surfaces Using Paramagnetic Probes

Andrew M. Petros,[‡] Luciano Mueller,[§] and Kenneth D. Kopple*
SmithKline Beecham Pharmaceuticals, P.O. Box 1539, King of Prussia, Pennsylvania 19406
Received March 16, 1990; Revised Manuscript Received June 22, 1990

ABSTRACT: Paramagnetic agents produce line broadening and thus cancellation of anti phase cross-peak components in two-dimensional correlated nuclear magnetic resonance spectra. The specificity of this effect was examined to determine its utility for identifying surface residues of proteins. Ubiquitin and hen egg white lysozyme, for which X-ray crystal structures and proton NMR assignments are available, served as test cases. Two relaxation reagents were employed, 4-hydroxy-2,2,6,6-tetramethylpiperidinyl-1-oxy and the gadolinium(III) diethylenetriaminepentaacetate complex ion. Correlations were sought between reagent-produced decreases of side-chain cross-peak volumes in double-quantum-filtered proton correlation (DQF-COSY) spectra and the solvent-exposed side-chain surface area of the corresponding residues. The lanthanide complex produced strong effects ascribable to association with carboxylate groups but was not otherwise useful in delineating surface residues. The nitroxyl, on the other hand, produced clear distinctions among the Val, Leu, and Ile residues that generally paralleled side-chain exposure in the crystal, although consistent correlations were not observed with residues of other types. Although an instance of possible specific protein-nitroxyl association was noted, the nitroxyl appears to be a tool for identifying hydrophobic surface residues.

For most enzymes and for globular proteins in general, surface residues play a critical functional and immunological role. Identification of a protein surface may therefore be of value in several ways. For hormonally active proteins, knowledge of the surface may aid in the design of peptide agonists and antagonists, especially when the functional regions are composed of residues not in a continuous sequence. For enzymes, mapping of surface residues may aid in identifying substrate or inhibitor binding sites, which in turn will facilitate inhibitor design. Furthermore, identification of surface residues may supplement NOE[1] and coupling constant data in the determination of protein structures by NMR.

Several methods have been proposed for using NMR spectroscopy to explore the surface structure of proteins and peptides. The photo-CIDNP technique has been used successfully to look at solvent-exposed aromatic side chains of several proteins (Kaptein et al., 1978; Kaptein, 1982; Stob et al., 1988). Other methods that have been employed or proposed include temperature dependence of amide proton chemical shifts (Kopple et al., 1969), proton-deuteron exchange rates (Wagner & Wuthrich, 1982), pH-rate profiles

* To whom correspondence should be addressed at SmithKline Beecham Pharmaceuticals, L-940, P.O. Box 1539, King of Prussia, PA 19406-0939.
‡ Present address: Abbott Laboratories, Dept. 47G. Bldg. AP9, Abbott Park, IL 60064.
§ Present address: Squibb Institute for Medical Research, Room D4159, P.O. Box 4000, Princeton, NJ 08543.

[1] Abbreviations: NOE, nuclear Overhauser effect; NMR, nuclear magnetic resonance; 1D, one dimensional; 2D, two dimensional; CIDNP, chemically induced dynamic nuclear polarization; COSY, correlated spectroscopy; DQF, double quantum filtered; HyTEMPO, 4-hydroxy-2,2,6,6-tetramethylpiperidinyl-1-oxy; Gd[DTPA], gadolinium(III) diethylenetriaminepentaacetate complex ion; hew, hen egg white; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, totally correlated spectroscopy; tri-NAG, tri-*N*-acetylglucosamine.

0006-2960/90/0429-10041$02.50/0