Probing Antibody Diversity by 2D NMR: Comparison of Amino Acid Sequences, Predicted Structures, and Observed Antibody–Antigen Interactions in Complexes of Two Antipeptide Antibodies†

Rina Levy,‡ Olga Assulin,‡ 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 March 20, 1989; Revised Manuscript Received May 19, 1989

ABSTRACT: The interactions between the aromatic amino acids of two monoclonal antibodies (TE32 and TE33) with specific amino acid residues of a peptide of cholera toxin (CTP3) have been determined by two-dimensional (2D) transferred NOE difference spectroscopy. Aromatic amino acids are found to play an important role in peptide binding. In both antibodies two tryptophan and two tyrosine residues and one histidine residue interact with the peptide. In TE33 there is an additional phenylalanine residue that also interacts with the peptide. The residues of the CTP3 peptide that have been found to interact with the antibody are val 3, pro 4, gly 5, gln 7, his 8, and asp 10. We have determined the amino acid sequences of the two antibodies by direct mRNA sequencing. Computerized molecular modeling has been used to build detailed all-atom models of both antibodies from the known conformations of other antibodies. These models allow unambiguous assignment of most of the antibody residues that interact with the peptide. A comparison of the amino acid sequences of the two anti-CTP3 antibodies with other antibodies from the same gene family reveals that the majority of the aromatic residues involved in the binding of CTP3 are conserved although these antibodies have different specificities. This similarity suggests that these aromatic residues create a general hydrophobic pocket and that other residues in the complementarity-determining regions (CDRs) modulate the shape and the polarity of the combining site to fit the specific antigens.

Due to its flexibility, a short peptide antigen may present different antigenic determinants to the immune system. The strength of the humoral immune response to each of the peptide conformations will be modulated by the mechanism that generates antibody diversity: different antibodies may be coded by different germ line genes and may undergo different somatic mutations and recombinations during B-cell development (Tonegawa, 1983). Our understanding of the structural basis of antibody–antigen recognition would be significantly enhanced if we could compare the combining site structure of different antibodies against a particular antigenic determinant and learn how changes in amino acid sequences of the antibody modulate structure and binding affinity. Here we undertake such studies using NMR[1] spectroscopy to examine the structure of the antibody–antigen complex in solution.

Three-dimensional structure determination by NMR uses the transfer of magnetization that takes place between neighboring protons coupled by magnetic dipolar interactions (Noggle & Schirmer, 1971). The intensity of the transferred magnetization depends on the reciprocal of the 6th power of the distance between the interacting nuclei and is effective up to a distance of approximately 5 Å. NOESY is a two-dimensional NMR experiment, which measures magnetization transfer between pairs of nuclei (Wüthrich, 1986). The diagonal in the 2D presentation yields the attenuated spectrum, and whenever there is a transfer between two nuclei, an off-diagonal peak appears in the spectrum at the frequency values of the two interacting nuclei. Any pair of protons that are less than 5 Å apart is expected to yield an observable NOESY cross peak. Once the cross peaks are assigned to the interacting nuclei, the distance restraints obtained by the NOESY experiment can be used to calculate a model for the three-dimensional structure of the studied molecule (Wüthrich, 1986). Another type of magnetization transfer that can be observed in the NOESY spectrum is that due to chemical exchange. 2D NMR techniques, especially NOESY, have been used extensively to study the structure of small proteins of molecular weight up to 15 000. Study of larger proteins is hindered by the loss of spectral resolution caused by the increasing number of protons and the considerable broadening of the individual proton resonances caused by slower overall tumbling in solution.

Recently we developed a novel approach for studying peptide–antibody interactions using 2D transferred NOE difference spectroscopy (Anglister et al., 1989). The 2D difference spectrum between the NOESY spectrum of the Fab with a 4-fold excess of the peptide and that of the peptide-saturated Fab reveals those cross peaks that grow in intensity as excess peptide is added. These cross peaks are due to magnetization

† This work was supported by grants from the U.S.–Israel Binational Science Foundation (85-00299) and by the Fund for Basic Research administered by the Israel Academy of Science.

* Correspondence should be addressed to this author. He is the incumbent of the Graham and Rhona Beck Career Development Chair.

‡ The Weizmann Institute of Science.

§ Stanford University School of Medicine.

[1] Abbreviations: NMR, nuclear magnetic resonance; 2D, two dimensional; NOE nuclear Overhauser effect; NOESY, 2D NOE spectroscopy; TRNOE, transferred NOE; COSY, 2D *J*-correlated spectroscopy; CDR, complementarity-determining region of the antibody molecule; FDR, frame-determining region of the antibody molecule (the segments that build the conserved structure of the immunoglobulin fold); Fv, antibody fragment that is made of the variable region of the light chain and the variable region of the heavy chain and forms one combining site for an antigen; Fab, antibody fragment that is made of the Fv, the light-chain constant region, and the first constant region of the heavy chain; BSA, bovine serum albumin; ELISA, enzyme-linked immunosorbent assay. Name convention: (a) specific peptide residue, lower case name, e.g., gln 7; (b) specific antibody residue, upper case name, e.g., Tyr 32L.

0006-2960/89/0428-7168$01.50/0

transfer between protons of the Fab and neighboring protons of the bound peptide, followed by further transfer to the free peptide protons by exchange between bound and free peptide (intermolecular transferred NOE; Balaram et al., 1978). Additional cross peaks appearing in the difference spectrum are due to a combination of intramolecular interactions between bound peptide protons and exchange between bound and free peptide (intramolecular transferred NOE; Albrand et al., 1979; Clore & Gronenborn, 1982, 1983). The assignment of the cross peaks in the 2D TRNOE difference spectrum is achieved by combination of (a) specific deuteration of the aromatic amino acids of the antibody, (b) resonance assignments of the free peptide deduced from the COSY spectrum of the peptide, and (c) specific deuteration of the peptide residues.

In the present study we demonstrate the application of our approach for a comparison of antigen–antibody interactions in complexes of two monoclonal antibodies with a peptide. The antibodies are against a 15 amino acid peptide of the B subunit of cholera toxin (residues 50–64) known as CTP3, and both are cross reactive with the toxin (Anglister et al., 1988). The amino acid sequences of the two antibodies are determined and compared with highly homologous sequences of antibodies from the same gene family but with different specificities. Models for the combining site structure have been calculated on the basis of the known three-dimensional structure of other antibodies. Examination of the sequences and the models enables the assignment of the interactions to specific residues in the antibody sequences.

Materials and Methods

Fab labeling and preparation, peptide synthesis and purification, and the procedure for NOESY measurements and difference spectra calculations have been described previously (Anglister et al., 1988, 1989). Measurements with TE32 Fab were carried out at 43 °C and with TE33 at 45 °C.

mRNA Sequencing. The mRNA coding for the TE32 and that coding for the TE33 antibodies were obtained by the method of Griffiths and Milstein (1985). The mRNA was sequenced by Sanger's dideoxy sequencing method (Sanger et al., 1977) as modified by Hamlyn et al. (1978) and Griffiths and Milstein (1985) for mRNA sequencing. The first primer for the heavy chain and the first one for the light chain were those used by Griffiths and Milstein. Two more primers were used for the TE33 light chain, 5′ CTC CAC TCT GCT GAT 3′ and 5′ AGA CTG GCC TGG TTT CTG 3′, and two more for the heavy chain, 5′ AGA GGT TTC CAA AGA 3′ and 5′ GAA ACC CTT TCC TGG AGT 3′, while for TE32 sequencing it was not necessary to use the third primer for either of the chains. Promega's Gemseq kit and AMV reverse transcriptase were used for the sequencing reactions (Promega, Madison, WI).

Specific Labeling of Chains. The specific labeling of the chains was accomplished as described by Anglister et al. (1985). The Bio-Gel DEAE ion-exchange column (Bio-Rad, Richmond, CA), used for chain separation, binds one of the chains of the reduced and alkylated Fab in 8 M urea and 50 mM glycine–Tris buffer, pH 8.3. The bound chain was eluted by 300 mM sodium chloride in the buffered urea solution. In order to identify the separated chains, a whole antibody reduced and alkylated was loaded on the column under identical denaturing conditions. Both chains of the antibody were bound to the column and were eluted with 300 mM sodium chloride in buffered urea solution. We therefore conclude that the heavy-chain fragment was that chain of the Fab that was not bound by the DEAE column.

Molecular Modeling. Detailed all-atom molecular models of the dimer formed from the light- and heavy-chain variable domains (Fv) of the antibodies TE32 and TE33 were constructed from the amino acid sequences of these molecules. The method, which has been used to model other antibodies (Chothia et al., 1986; Anglister et al., 1987), involves two steps. First, we piece together fragments taken from four variable chain dimers whose three-dimensional structure has been solved by X-ray crystallography (KOL, Marquart et al., 1980; MCP, Segal et al., 1974; NEW, Saul et al., 1978; 539, Won Suh et al., 1986). Second, we use a robust energy minimization method (Levitt, 1983) to refine the coordinates to give a low-energy conformation with acceptable stereochemistry. In this section, we refer to the particular antibody being modeled as the "target" structure and the antibodies used to provide the coordinates as the "known" structures.

Central to this modeling scheme is the need to determine which fragment of the known structures could serve as the basis for building the target structure. We use our modified alignment method that favors gaps in the loops between the segments of core secondary structure (Lesk et al., 1986) to align all the sequences (both of the target and known structures) to that of one of the known structures (we use the KOL sequence as this standard as it is longer than the other sequences). The alignment of each of the sequences to the same standard provides an alignment of all pairs of sequences to one another. Although this multisequence alignment is not optimal, it has proved adequate for all modeling of antibodies, which requires that chain segments rather than individual residues be correctly aligned.

Each variable domain is divided into 14 segments that delineate the nine β strands and the loops between them (see Figures 1 and 2). With these segment boundaries, the highly variable segments are flanked by well-conserved segments. Alignment of these flanking regions, which is facilitated by their high degree of homology, therefore, ensures correct alignment of the variable segments. Because the well-conserved segments have been defined so that their lengths are conserved, alignment of the segments is equivalent to alignment of every residue in the segment. For the hypervariable segments, which can be of different lengths, alignment of individual residues is done in the next step of the procedure.

After segment alignment is complete, each segment of the target protein is in correspondence with four segments, one from each of the four known structures. One of these segments is chosen as follows: (a) If any of the four known segments is the same length as the corresponding target segment, select the known segment (of the same length) that has the highest homology score to the sequence of the target segment [homology score is the weighted sum of the fraction of identical residues in the segment (weight of 0.25) and in the heavy or light chain as a whole (weight of 0.75)]. (b) If the target segment is shorter by one residue or more than the known segments, introduce a single deletion in the known segment at a position chosen both to maximize the homology and to minimize the disruption to the chain path (measured between the C^{α} coordinates adjacent to the deletion in the known segment). (c) If the target segment is longer than all known segments, introduce a single insertion in the target segment at a position chosen to maximize the homology. Figures 1 and 2 show the segments selected for the modeling of TE33; those used for TE32 are generally very similar.

A starting set of coordinates for the target structure is made by using the coordinates of the known structure selected for each particular segment. Such transfer of coordinates suffers

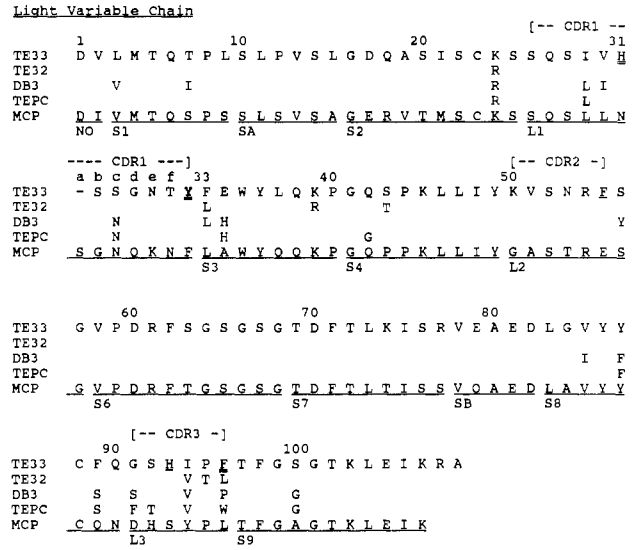

FIGURE 1: Comparison of the amino acid sequence of the variable regions of the light chains of two anti-CTP3 antibodies, TE32 and TE33, with the amino acid sequence of antibodies from the same gene family. DB3 is an anti-progesterone antibody (Deverson et al., 1987), TEPC-105 (abbreviated as TEPC) is a mouse myeloma antibody whose specificity is unknown (Moynet et al., 1985), and MCP is a myeloma antibody that binds phosphorylcholine and whose three-dimensional structure is known (Segal et al., 1974). The CDRs are marked above the TE33 sequence, using a broad definition that is based on the definitions of Chothia and Lesk (1987) and Bruccoleri et al. (1988). The full sequence is given for TE33, whereas for TE32, DB3, and TEPC, we only show the changes with respect to TE33. The aromatic residues in the CDRs of TE33 are underlined, while the subset of these residues found in this study to interact with the peptide antigen is doubly underlined. We also give the sequence of the light-chain variable region of MCP whose three-dimensional structure was used to calculate the structures of TE32 and TE33. The particular sequences used to provide initial coordinates for each segment of the TE33 model are underlined. The segments used for the modeling are marked by a two-letter segment name placed at the start of the particular segment.

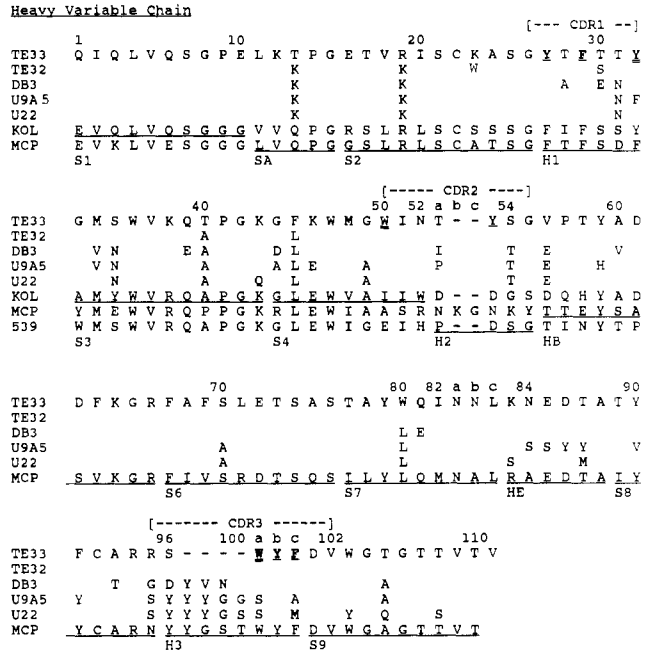

FIGURE 2: Comparison of the amino acid sequence of the variable regions of the heavy chains of two anti-CTP3 antibodies, TE32 and TE33, with the amino acid sequence of antibodies from the same gene family. DB3 is an anti-progesterone antibody (Deverson et al., 1987); U9A5 and U22B5 (abbreviated as U22) are anti-arsonate monoclonal antibodies (Meek et al., 1984). The CDRs are marked above the TE33 sequence. The full sequence is given for TE33, whereas for TE32, DB3, U9A5, and U22B5, we only show the changes with respect to TE33. The aromatic residues in the CDRs of TE33 are underlined, while the subset of these residues found in this study to interact with the peptide antigen is doubly underlined. We also give the sequences of the proteins (KOL, MCP, and 539) with known three-dimensional structures, which are used to calculate the structure of TE33. The particular sequences used to provide initial coordinates for each segment of the TE33 model are underlined. The segments used for the modeling are marked by a two-letter segment name placed at the start of the particular segment.

from two major defects: (a) Atoms are missing in those cases where the target residue is bigger than the known residue at a particular position. Our procedure uses randomized coordinates for these missing atoms, relying on the subsequent energy minimization to rectify defects in stereochemistry. (b) Breaks occur in the main chain between segments taken from different known structures. Both these defects in stereochemistry are eliminated by limited energy minimization (500 steps of conjugate gradients with a weak restraint holding the C^{α} atoms to their positions in the known structures). For TE33 and TE32, the number of missing atoms is 126 and 117, respectively (out of a total of 2025 and 2018 atoms, respectively), and the total energy values at the end of the refinement are −1534 and −1486 kcal/mol, respectively.

RESULTS

Amino Acid Sequences. The two anti-CTP3 antibodies TE32 and TE33 were sequenced by direct mRNA sequencing. The amino acid sequences of TE32 and TE33 are shown in Figures 1 and 2 together with sequences of other closely homologous antibodies. The light-chain sequences are typical to group II κ light chains (Kabat et al., 1987). There are 7 differences out of 114 between the TE33 and TE32 light-chain sequences (94% residue identity). One change occurs in the first FDR (frame determining region), three changes occur in the second FDR, and three changes occur in the third CDR. The heavy chains of TE32 and TE33 use a family IX variable region gene (Winter et al., 1985). There are 6 differences out of 114 residues between the TE33 and TE32 heavy-chain sequences (95% identity); all of them except one in the first CDR are in frame-determining regions. All the changes in the amino acid sequences are due to somatic point mutations in a parent heavy-chain gene and a parent light-chain gene. The most extensive changes occur in the third hypervariable region of the light chain in which a sequence of Ile-Pro-Phe in TE33 changes into Val-Thr-Ile in TE32. We have found that there is a high degree of sequence homology between the anti-CTP3 antibodies and the DB3 anti-progesterone antibody (Deverson et al., 1987). A total of 98 out of 114 residues are identical in the TE33 and DB3 light-chain variable region (86% identity) and 93 out of 117 residues are identical in the heavy-chain variable region (79% identity). The light chain of TEPC 105 (Moynet et al., 1985) is even more similar to TE33 with 102 out of 114 residues identical (89% identity). The amino acid sequences of the heavy-chain variable region of anti-arsonate antibodies, represented here by the two antibodies U9A5 and U22B5 (Meek et al., 1984), are also coded by a gene of family IX and are also highly homologous with the sequences of the anti-CTP3 antibodies. The degree of identity in amino acid sequence between the anti-CTP3 antibodies and the other antibodies given Figures 1 and 2 is typical to that found for antibodies in the same gene family.

Antibody–Antigen Interactions Detected by 2D TRNOE Difference Spectroscopy. In a previous study (Anglister et al., 1989) we have shown how 2D TRNOE difference spectroscopy can be used together with specific deuteration of the

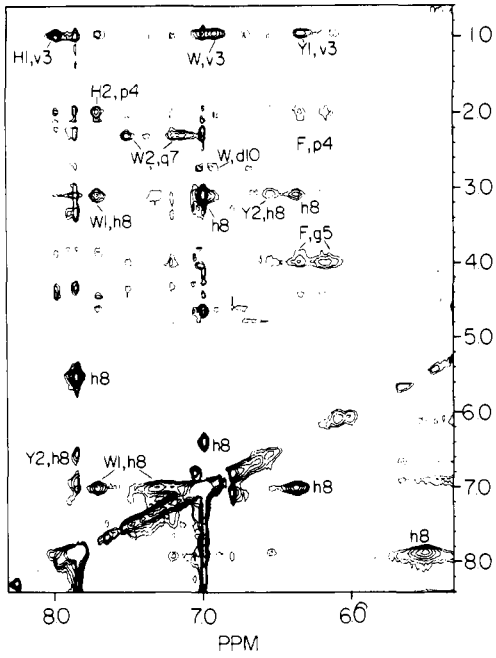

FIGURE 3: A portion of the 2D difference spectrum between the NOESY spectrum of TE33 Fab saturated with the peptide and the NOESY spectrum of the Fab in the presence of a 4-fold excess of the peptide. The Fab and the peptide are not labeled by deuteration. The portion presented shows the interactions between aromatic protons of the antibody and protons of the peptide. Assigned antibody residues are marked by capital letters and arbitrary numbers; peptide protons are marked by lower case letters and numbers denoting the location in the sequence of the residues to which the protons belong. Additional cross peaks marked by h8 are of the peptide his 8 and are due to exchange between bound and free peptide or intraresidue interactions in the bound peptide.

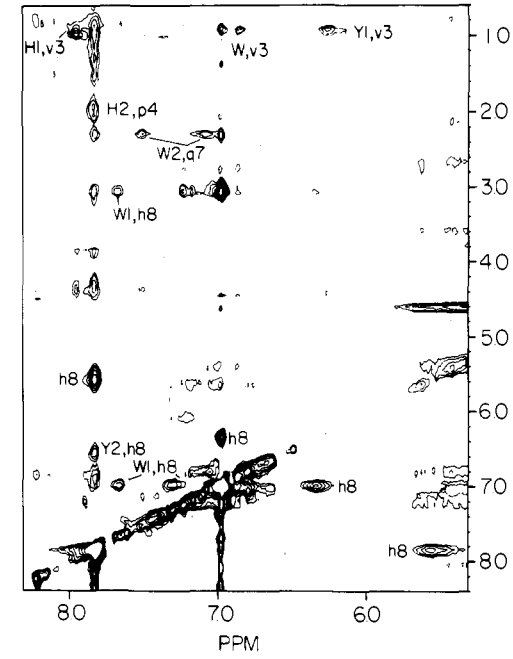

FIGURE 4: A portion of the 2D difference spectra between the NOESY spectrum of TE32 Fab saturated with the peptide and the NOESY spectrum of the Fab in the presence of a 4-fold excess of the peptide. The Fab and the peptide are not labeled. The section presented shows the interactions between aromatic protons of the antibody and protons of the peptide. The assignment of the observed interactions is based on comparison with the difference spectrum obtained for TE33 and the difference spectra obtained for specifically deuterated preparations of TE32 Fab shown in Figure 5. Assigned antibody residues are marked by captial letters and arbitrary numbers; peptide protons are marked by lower case letters and numbers denoting the location in the sequence of the residues to which the protons belong. Additional cross peaks marked by h8 are of the peptide his 8 and are due to exchange between bound and free peptide or intraresidue interactions in the bound peptide.

antibodies to identify the interactions between the aromatic amino acids of the antibody and the residues of a peptide antigen. Isotopic labeling of antibodies is both costly and time consuming. We therefore decided to measure NOESY spectra and calculate the 2D TRNOE difference spectrum using nonlabeled antibody. Figure 3 shows the difference between the NOESY spectrum of the Fab of TE33 in the presence of a 4-fold excess of CTP3 peptide and the NOESY spectrum of the peptide-saturated Fab. The spectrum is a remarkably accurate sum of the interactions that we previously observed for each type of aromatic amino acid by appropriate labeling of the antibody molecule (Anglister et al., 1989). The assignment of cross peaks to the individual type of antibody amino acid and specific peptide residue is based on our previous measurements with specifically labeled antibody and peptide. A new cross peak at 7.87, 6.89 ppm that did not appear in any of the previously measured difference spectra is attributed to interaction between $C_{\delta 1}H$ and $C_{\delta 2}H$ of antibody tyrosine with $C_{\delta 2}H$ of his 8 of the peptide. These protons are most probably of the antibody tyrosine whose $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ were found to interact with the same peptide proton. It should be noted, as well, that previously observed interactions between an antibody tyrosine and both pro 4 and asp 10 of the peptide are too weak to be observed in Figure 3. With the labeled antibody the corresponding cross peaks were enhanced by the partial deuteration of tyrosine, which resulted in a considerable narrowing of the tyrosine $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ resonances.

The quality of the difference spectrum for the nonlabeled Fab in terms of both resolution and signal-to-noise ratio indicates that a detailed comparison of antibody–antigen interactions in complexes of different monoclonal antibodies with the same specificity is feasible without extensive deuteration. Figure 4 shows a 2D TRNOE difference spectrum defining the interactions between aromatic amino acids of the second antibody TE32 and residues of the same CTP3 peptide. The 2D TRNOE difference spectra of the two unlabeled antibodies seem to be very similar, facilitating the assignment of the cross peaks in Figure 4. In the TE32–peptide complex a histidine imidazole proton of the antibody is found to interact with a methyl group of the peptide valine 3; this is also seen in TE33. Another histidine proton interacts with both a methyl group of valine 3 and the γ protons of the peptide proline 4. There are two antibody tryptophan residues that interact with the peptide: one of them interacts with $C_{\delta 2}H$ and with the β protons of histidine 8 of the peptide. A second antibody tryptophan residue interacts with the γ protons of the peptide glutamine 7. There are two tyrosine residues of the antibody interacting with the peptide: one of them interacts with the two imidazole protons of the peptide histidine 8 and the other interacts with the methyl of valine 3 of the peptide; this is also seen in TE33. In TE32 we do not find an antibody phenylalanine residue that interacts with the peptide whereas in TE33 we did observe an interaction of antibody phenylalanine with pro 4 and gly 5 of the peptide.

For further verification of our assignments, we have carried out the measurements and the 2D TRNOE difference spectra calculations with three different preparations of TE32-labeled Fab. Figure 5A shows interactions between TE32 tyrosine ($C_{\epsilon 1}H$ and $C_{\epsilon 2}H$) and histidine imidazole protons with the peptide protons. Interactions between an antibody tyrosine and both pro 4 and asp 10, which were too weak for detection with nonlabeled Fab, are now clearly observed. The same interactions are observed in the TE33–peptide complex. Figure

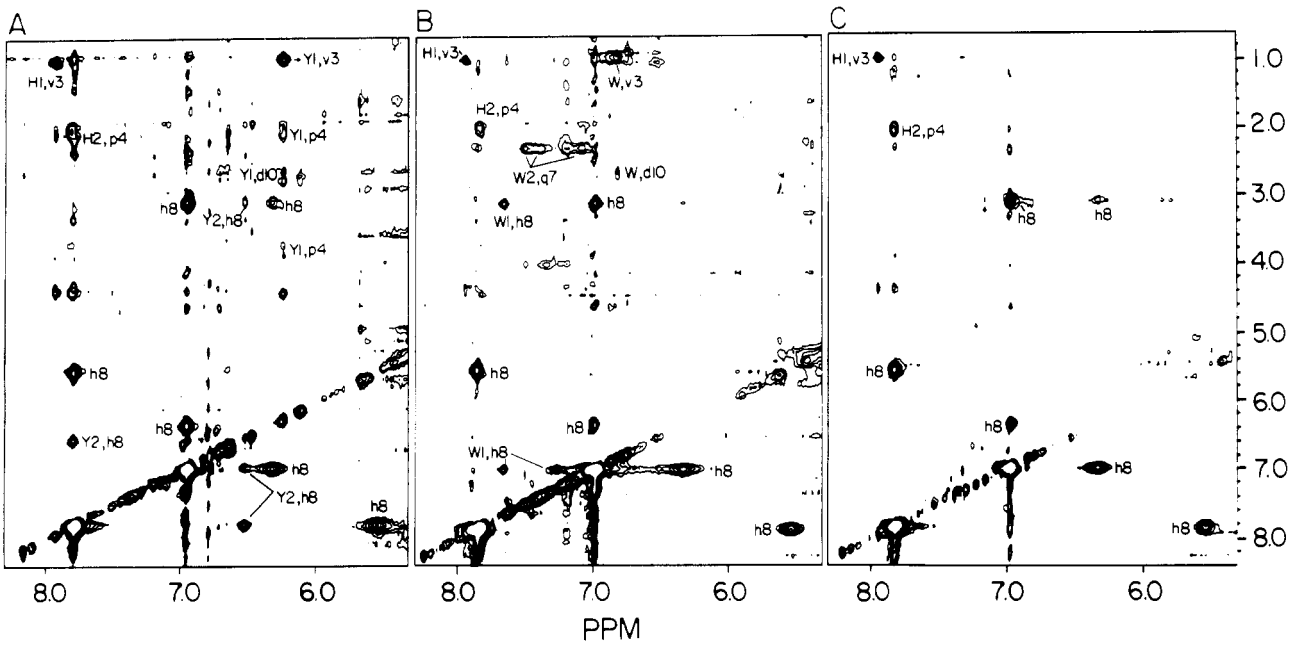

FIGURE 5: 2D difference spectra between the NOESY spectrum of TE32 Fab with a 4-fold excess of the peptide and the NOESY spectrum of the peptide-saturated Fab, showing interactions of specific types of antibody aromatic residues with peptide residues. Assignment to antibody residues is marked by capital letters and arbitrary numbers while assignment to peptide residues is marked by lower case letters and their location in the sequence. (A) Interactions of antibody tyrosine and histidine residues with peptide residues. Antibody phenylalanine and tryptophan residues are perdeuterated while tyrosine residues are deuterated at 2,6 phenyl positions. (B) Interactions of antibody tryptophan and histidine residues with peptide residues. Antibody phenylalanine and tyrosine residues are perdeuterated. (C) Interactions of antibody phenylalanine and histidine residues with peptide residues. Antibody tyrosine and tryptophan residues are perdeuterated.

5B shows interactions between TE32 tryptophan and histidine aromatic protons with protons of the peptide. Interaction between an antibody tryptophan and asp 10 of the peptide is observed in the difference spectrum of the labeled Fab. Only a weak cross peak corresponding to this interaction is observed in the nonlabeled Fab. It should be noted that the same tryptophan proton interacts with a methyl group of valine 3 of the peptide. It is clear from Figure 5C that there are no interactions between phenylalanine residues of TE32 and the CTP3 peptide.

Assignment of Antibody Tyrosine Residue to a Specific Chain. Rigorous assignment of the interactions observed in the NOESY difference spectra to specific residues of the antibody by sequential assignment methods, applied for small proteins using COSY and NOESY spectra (Wüthrich, 1986), is not feasible for the much larger Fab molecule. To facilitate the assignment procedure, we measured 1D difference spectra between the spectrum of the Fab and the complexed Fab of two preparations in which the chains were specifically labeled; the results are shown in Figure 6. Figure 6A shows the resonances of $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ of tyrosine residues of the antibody heavy chain that undergo changes in chemical shift upon peptide binding. Figure 6B shows the resonances of light-chain tyrosine $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ that change their chemical shift upon antigen binding. Figure 6C shows the contribution of tyrosine proton resonances of both chains that undergo changes in chemical shift upon peptide binding. Probably the major contributions to the difference spectrum are of amino acids in the combining site region. According to the chemical shifts of the tyrosine protons in the difference spectra, we conclude that the antibody tyrosine residue interacting with valine 3, proline 4, and aspartate 10 of the peptide is contributed by the light chain and that interacting with histidine 8 of the peptide is contributed by the heavy chain.

Titration of Histidine Residues. Two histidine protons of the antibody have been found to interact with the peptide. Since there are two histidine residues in the CDRs of the light

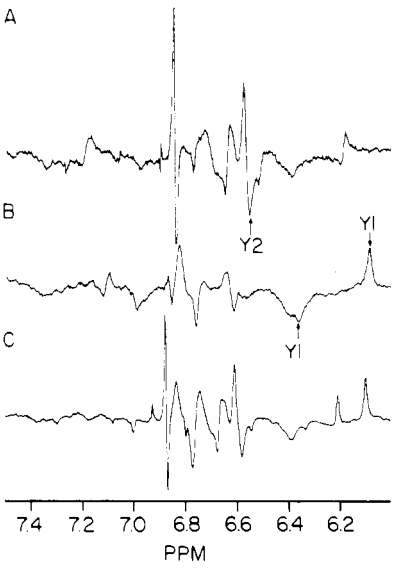

FIGURE 6: 1D difference spectrum between the spectrum of TE33 Fab and that of Fab saturated with the peptide, showing the contributions to the difference spectrum of tyrosine protons $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ of each of the chains. (A) All aromatic protons of antibody tryptophan, phenylalanine, and tyrosine are deuterated except heavy-chain tyrosine $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$. (B) All aromatic protons of antibody tryptophan, phenylalanine, and tyrosine are deuterated except light-chain tyrosine $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$. (C) All aromatic protons of antibody tryptophan, phenylalanine, and tyrosine are deuterated except tyrosine $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ of both chains. Y1 is the resonance of the $C_{\epsilon 1}H$ and the $C_{\epsilon 2}H$ of the antibody tyrosine interacting with val 3, pro 4, and asp 10 of the peptide. Y2 is the resonance of the $C_{\epsilon 1}H$ and the $C_{\epsilon 2}H$ of the antibody tyrosine interacting with his 8 of the peptide.

chain of both TE32 and TE33, a question arises as to whether these two protons are of the same histidine residue or of two different residues. In order to answer this question, we titrated the pH of a solution of TE33 Fab in which all aromatic residues except histidine were perdeuterated. We have found that

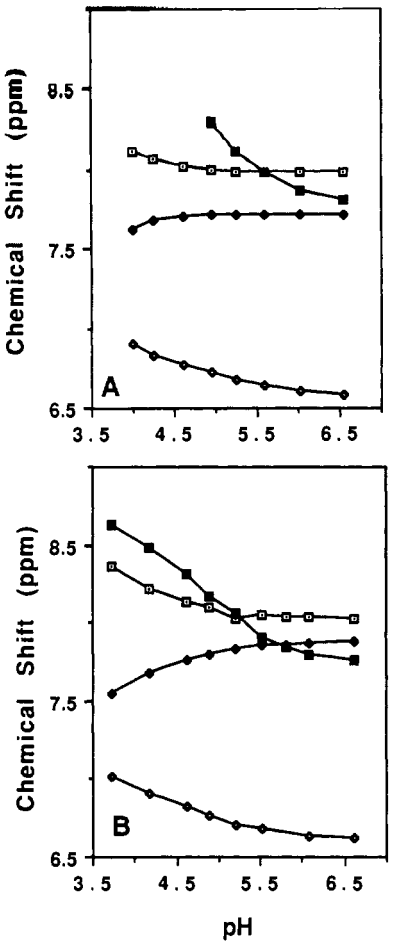

FIGURE 7: pH dependence of the chemical shift of antibody histidine protons that change their chemical shift values at low pH. (A) Histidine protons in TE33 Fab in the presence of a 4-fold excess of the peptide: (⊡) H1 (see Figure 3), (◆) H2 (see Figure 3), (■) Hc1, and (◇) Hc2. (B) Histidine protons in TE33 Fab without peptide: (⊡) H1, (◆) H2, (■) Hc1, and (◇) Hc2.

four histidine protons change their chemical shift at unusually low pH values. Figure 7A shows the changes in the chemical shift of the imidazole protons of the titrated histidine residues in Fab with a 4-fold peptide excess. Two of the protons change their chemical shift continuously from pH 6.6 and below. Two other protons change their chemical shift only below a pH of 5. These two protons have been found to interact with the peptide (Figure 3). On the basis of the similar behavior of these two protons, which is different from the behavior of all other histidine protons of the antibody and the peptide, we conclude that they are of the same histidine residue. Figure 7B shows the changes in the chemical shift of the titrated histidine residues in Fab without peptide. Two protons change their chemical shift continuously from pH 6.6 with very similar chemical shift values to those of the histidine protons of the complex that are not interacting with the peptide. The two other protons that were found to interact with the peptide exhibit considerably smaller changes in chemical shift between pH 6.6 and pH 5.2. One of the protons that was found to interact with the peptide exhibits another peculiar behavior both in the Fab and in the complexed Fab; its chemical shift changes to higher field (lower chemical shift values) at lower pH. The protons marked H1 and Hc1 in Figure 7 exchange with the deuterium of the solvent after several weeks; however, H1 is even more protected from exchange in the Fab complex with the peptide. We therefore assign these two protons as $C_{\epsilon 1}H$ of two different antibody histidine residues. The resonances marked H2 and Hc2 in Figure 7 are assigned as $C_{\delta 2}H$ of two antibody histidine residues. The protons Hc1 and Hc2 exhibit behavior similar to that of the imidazole protons of His 198L, that is, a conserved residue in the constant domain of both λ and κ light chains, as observed by Arata and co-workers (Arata & Shimizu, 1979; Akira et al., 1980).

Discussion

A view of the combining site in the calculated model of TE33 Fv is presented in Figure 8. As the models of TE33 and TE32 were calculated independently of the NMR results, the results of the NMR experiments can be used to verify and refine the models. Both the molecular models and the NMR results suggest that the aromatic amino acids play an important role in the combining site of the two antibodies. According to the NMR results for both TE32 and TE33, two tyrosine and two tryptophan residues and one histidine residue interact with the peptide. A phenylalanine residue is found to interact with the peptide in TE33 but not in TE32.

It is expected that the antibody residues interacting with the antigen will be of one of the six CDRs. According to Wu and Kabat (1970) the light-chain CDRs are 24–34, 50–56, and 89–97 and the heavy-chain CDRs are 31–35, 50–65, and 95–102. Two more recent studies by Chothia and Lesk (1987) and Brucolleri et al. (1988) on analysis and theoretical modeling of the three-dimensional structure of the antibody Fv suggest slightly different definitions of the CDRs. A new broader definition of CDRs, which encompasses the definition of Chothia and Lesk and of Brucolleri et al., is as follows: light-chain CDRs, 26–32, 50–55, and 91–96; heavy-chain CDRs, 26–32, 50–55, and 95–102.

Examination of the amino acid sequences of TE32 and TE33 reveals one tryptophan in the second CDR of the heavy chain (Trp 50H) and another tryptophan residue in the third CDR of the heavy chain (Trp 100aH). According to the calculated model, both tryptophan residues are exposed to the solvent and are in the combining site. Therefore, we conclude that one of these tryptophan residues interacts with his 8 of the peptide and the other interacts with gln 7 of the peptide.

There is one tyrosine residue in the CDRs of the light chain (Tyr 32L), and there are four tyrosine residues in the CDRs of the heavy chain (Tyr 27H, 32H, 53H, and 100bH). On the basis of the specific labeling of the chains, we believe that Tyr 32L is the antibody tyrosine found to interact with val 3, pro 4, and asp 10 of the peptide. In the calculated model this tyrosine is partly exposed to the solvent and is in the combining site. The tryptophan proton interacting with asp 10 and val 3 of the peptide should therefore be in the vicinity of the tyrosine residue interacting with the same peptide residues. We therefore conclude that this tryptophan proton is $C_{\delta 1}H$ of Trp 100aH. According to the calculated model $C_{\delta 1}H$ of Trp 100aH is near $C_{\epsilon 1}H$ and $C_{\epsilon 2}H$ of Tyr 32L. Tyr residues 32H, 53H, and 100bH are in the combining site and exposed to the solvent (see Figure 8). At this stage we believe that Tyr 32H interacts with his 8 of the peptide, but we cannot rule out the possibility that the antibody tyrosine is actually Tyr 53H or Tyr 100bH.

There are two histidine residues in the CDRs, His 31L and His 93L, both in the light chain. The calculated model predicts that His 31L points in the direction of the combining site while His 93L points away from the combining site. In the model, His 31L is near tyrosine 32L, which is thought to interact with val 3, pro 4, and asp 10 of the peptide. On the basis of the proximity between His 31L and Tyr 32L in the model and the observation that there is a histidine proton of the antibody interacting with val 3 of the peptide and another histidine proton (of the same residue) interacting with pro 4, we believe

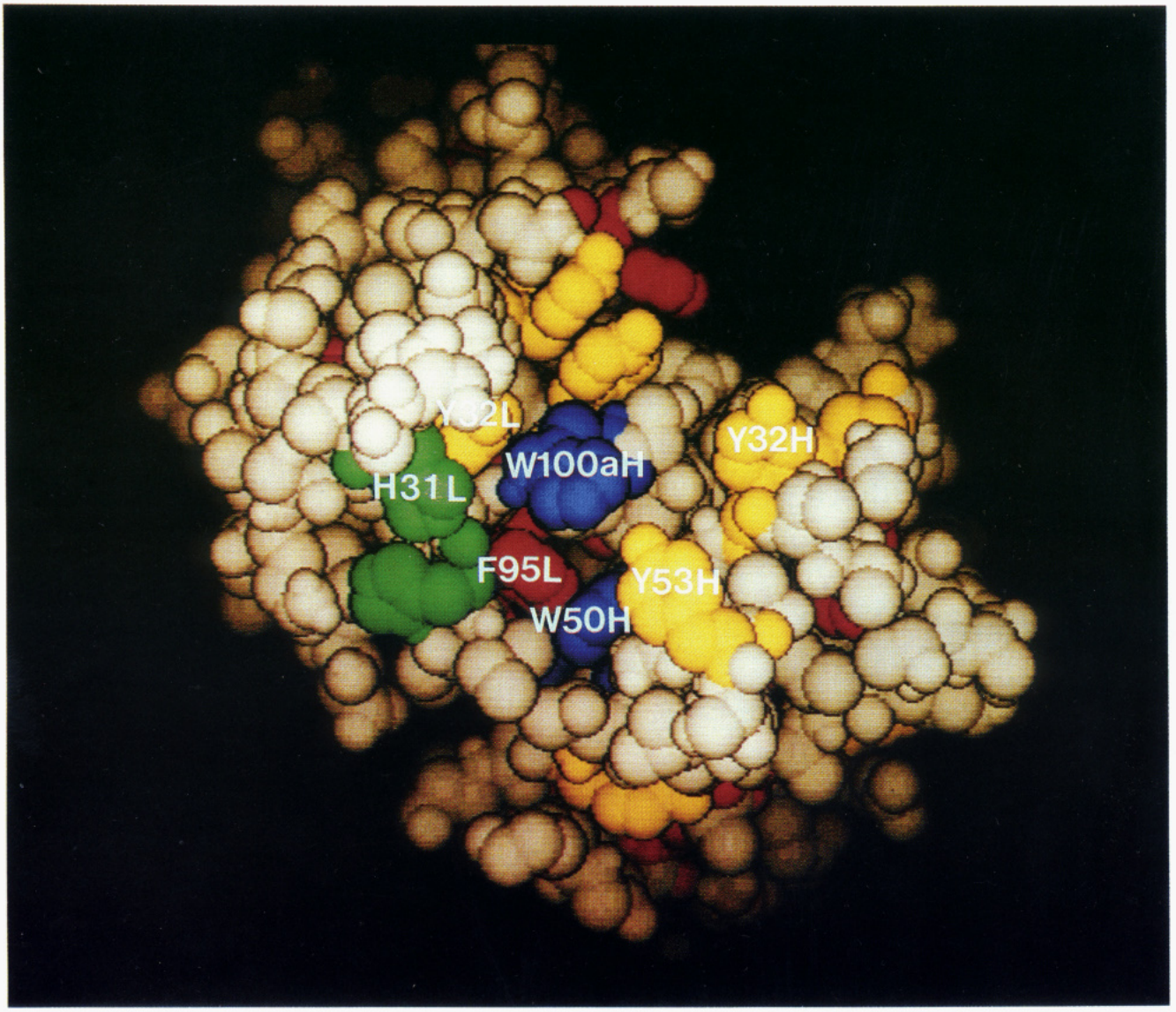

FIGURE 8: A model for the Fv antigen combining site of TE33 calculated from known Fv structures. Aromatic residues have been colored to aid identification as follows: His is green, Phe is red, Trp is blue, and Tyr is yellow. The shading of this space-filling model employs strong depth-cuing to emphasize the clefts and depressions in the antigen binding site. Some of the aromatic CDR residues found in TE33 also have close analogues in one of the known antibodies, MCP, used as a source of coordinates (Tyr 32L, Tyr 27H, Phe 29H, Tyr 32H, Trp 100aH, Tyr 100bH, and Phe 100cH). These side chains are modeled more reliably than those for which there is no close analogue in a known structure (His 31L, Phe 55L, His 93L, Phe 96L, Trp 50H, and Tyr 53H).

that the antibody histidine involved in the peptide binding is His 31L. The model also explains why the histidine protons involved in the binding change their chemical shifts at unusually low pH. According to the model His 31L is partly buried in a hydrophobic environment formed by Tyr 32L, Trp 100aH, Phe 96L, and His 93L.

The phenylalanine residues in the CDRs are Phe 55L, Phe 96L, Phe 29H, and Phe 100cH. All these residues except Phe 96L are buried inside the Fv structure. Phe 96L of TE33 is replaced by Leu in TE32. This replacement explains why an antibody phenylalanine is seen to interact with the peptide in a complex with TE33 but not with TE32. This difference may account for the stronger binding of the peptide to TE33 compared with TE32.

Comparison of the sequences of the light and heavy variable chains of TE32 or TE33 with other sequences in the same gene family (see Figures 1 and 2) shows that 10 of the 13 aromatic residues in the CDRs are always conserved (His 31L, Tyr 32L, Phe 55L, His 93L, Tyr 27H, Phe 29H, Tyr 32H, Trp 50H, Tyr 53H, and Tyr 100bH). The aromatic residues that change include (a) Phe 96L, which becomes Leu, Pro, and Trp in TE32, DB3, and TEPC, respectively, (b) Trp 100aH, which becomes Ser in both U9A5 and U22, and (c) Phe 100cH, which becomes Ala in U9A5 and Met in U22. The great similarity of DB3, an anti-progesterone antibody, to our anti-peptide antibodies, TE32 and TE33, is of some interest. The *only* aromatic CDR residue that is different is Phe 96L, which is only partly exposed and unlikely to dramatically affect the shape of the binding site (see Figure 8) or the affinity to CTP3 as it is also replaced by Leu in TE32. Other differences, which are all in the heavy-chain CDR3, include the change Arg 95H → Gly and a three residue insert of Tyr-Val-Asn after Ser 96H. The Arg to Gly change removes the large Arg side chain hidden below Trp 100aH in Figure 8; this will make the pocket deeper and less polar. The three-residue insertion adds Tyr and Val side chains above that of Trp 100aH, filling the depression that exists in TE33 and TE32.

Solid-phase binding experiments (ELISA) carried out with TE32 and TE33 antibodies and a representative monoclonal anti-progesterone antibody reveal that the binding of TE32

and TE33 to progesterone linked to BSA is only slightly stronger than the background binding to BSA and is much weaker than the binding of these antibodies to CTP3 linked to BSA. The monoclonal anti-progesterone antibody did not show any specific binding to CTP3 linked to BSA (Levi and Anglister, unpublished results). We therefore conclude that although the amino acid sequences of the anti-CTP3 antibodies are very similar to the amino acid sequences of the anti-progesterone antibodies (Deverson et al., 1987) and although all the antibody aromatic amino acids that have been found to interact with the antigen are conserved, the cross reactivity between the two groups of antibodies is very weak if not negligible.

The conservation of those aromatic amino acids involved in peptide binding suggests that the role of the aromatic amino acids in the CDRs is to create a general hydrophobic depression for antigen binding; other residues in other positions modulate the shape of the combining site to fit the specific antigen. A comparison of the interactions observed in the two anti-CTP3 antibodies with the interactions observed in two anti-lysozyme antibodies (recognizing different epitopes of lysozyme) also demonstrates the importance of aromatic amino acids in antigen binding. In the D1.3 antibody, whose structure was solved by Amit et al. (1986), 9 out of the 15 antibody residues directly interacting with the antigen were found to be aromatic. In the Hyhel-5 antibody, whose structure was solved by Sherif et al. (1987), 5 out of 17 residues interacting with the antigen are aromatic. In both anti-lysozyme antibodies residue Tyr 32L interacts with the antigen.

The 2D TRNOE difference spectroscopy combined with theoretical model building provides a very powerful method to study the interactions between antibodies and a peptide antigen. Although the experiments described in the present study have been carried out in a system for which the ligand off-rate is fast relative to the spin–lattice relaxation time, there are ways to apply the NOESY difference spectroscopy for systems in which the ligand off-rate is slower. This may be done by using a truncated or modified peptide with weaker affinity to the antibody, a spin-labeled peptide, or a fully deuterated peptide. Experiments to obtain further information about intramolecular interactions in the Fab and in the bound peptide are under way. This information is most important to further verify and refine the calculated model.

The elucidation by X-ray crystallography of the three-dimensional structure of the complex of DB3 antibody with progesterone is in progress (Stura et al., 1987). There is 83% sequence identity between the anti-CTP3 antibodies and DB3, whereas the degree of sequence identity between the anti-CTP3 antibodies and the other antibodies whose three-dimensional structure has been determined is only about 50%. It should be possible to calculate more accurate models for TE32 and TE33 from the coordinates of the DB3 antibody when they are available.

Acknowledgments

We are grateful to Prof. Ruth Arnon and Dr. Chaim Jacob for the gift of the TE33 and TE32 hybridomas and to Dr. Ora Goldberg for peptide synthesis.

References

Albrand, J. P., Birdsall, B., Feeny, J., Roberts, G. C. K., & Burgen, A. S. V. (1979) *Int. J. Biol. Macromol. 1*, 37.

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

Anglister, J., Frey, T., & McConnell, H. M. (1985) *Nature 315*, 65–67.

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

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

Arata, Y., & Shimizu, A. (1979) *Biochemistry 18*, 2513–2520.

Balaram, P., Bothner-By, A. A., & Dadok, J. (1972) *J. Am. Chem. Soc. 94*, 4015–4017.

Bruccoleri, R. E., Haber, E., & Novotny, J. (1988) *Nature 335*, 564–568.

Chothia, C., & Lesk, A. M. (1987) *J. Mol. Biol. 196*, 901–917.

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.

Clore, G. M., & Gronenborn, A. M. (1982) *J. Magn. Reson. 48*, 402–417.

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

Deverson, E., Berek, C., Taussig, M., & Feinstein, A. (1987) *Eur. J. Immunol. 17*, 9–13.

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

Hamlyn, P. H., Brownlee, G. G., Cheng, C. C., Gait, M. J., & Milstein, C. (1978) *Cell 15*, 1067–1075.

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.

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

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

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.

Moynet, D., McClean, S. J., Ng, K. H., Anctil, D., & Gibson, D. M. (1985) *J. Immunol. 134*, 3455–3460.

Sanger, F., Nicklen, S., & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463–5467.

Saul, F., 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.

Shimizu, A., Honzawa, M., Yamamura, Y., & Arata, Y. (1980) *Biochemistry 19*, 2784–2790.

Stura, E. A., Feinstein, A., & Wilson, I. A. (1987) *J. Mol. Biol. 193*, 229–231.

Suh, S. W., Bhat, T. N., Navia, M. A., Cohen, G. H., Rao, D. N., Rudikoff, S., & Davies, D. (1986) *Proteins 1*, 74–80.

Tonegawa, S. (1983) *Nature 302*, 575–581.

Winter, E., Radbuch, A., & Krawinkel, U. (1985) *EMBO J. 4*, 2861–2867.

Wu, T. T., & Kabat, E. A. (1970) *J. Exp. Med. 182*, 211–250.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.