Sequencing and Modeling of Anti-DNA Immunoglobulin Fv Domains

COMPARISON WITH CRYSTAL STRUCTURES*

(Received for publication, September 30, 1993)

M. Michele Barry, Clifford D. Molt, Wayne F. Anderson**, and Jeremy S. Lee

From the Department of Biochemistry, the University of Saskatchewan, Saskatoon, Saskatchewan S7N 0W0, Canada and the **Department of Biochemistry, Vanderbilt University, Nashville, Tennessee 37232-0146

Models for the three-dimensional structures of the combining regions of six DNA-binding antibodies have been derived from the sequence data for their Fv domains presented here. Using the amino acid sequences and the canonical structure classes described by Chothia and Lesk (Chothia, C., and Lesk, A. M. (1987) J. Mol. Biol. 196, 901), model loops were selected from immunoglobulin domains of known structure for five of the six antibody hypervariable regions. Models for the third complementarity-determining region of the heavy chain were constructed from known immunoglobulin loops of similar length and sequence. Comparison of three of the models with the respective crystal structure indicates that this procedure can generate a working model of the antibody combining region that provides useful information on the nature of the interactions between antibodies and nucleic acids. As part of our continuing investigation into the structural basis of antibody-DNA recognition, the observed and predicted models for the combining regions of nucleic acid-binding antibodies have been examined. In general, single strand-specific antibodies have deep clefts where the antigen might bind, whereas duplex-specific antibodies present a relatively flat surface. In addition, on the basis of both sequence and structure, there is little to distinguish autoimmune antibodies from those produced by immunization. Testable hypotheses for how these antibodies might interact with single- and double-stranded nucleic acids are presented.

The wide range of antibody specificities is determined by the amino acid sequences and conformations of six hypervariable loops, or complementarity-determining regions (CDRs). Antibody combining sites are composed of three hypervariable loops from the variable light chain domain (V_{L}) and three loops from the variable heavy chain domain (V_{H}), referred to as L1, L2, and L3 and H1, H2, and H3, respectively. To understand how an antibody might interact with its antigen, the conformations of these six loops must be determined. We are interested in the structures of DNA-binding antibodies and the elucidation of the principles that govern antibody-DNA recognition.

Two different methods for using amino acid sequence information to predict the conformations of antibody hypervariable loops have been previously employed. The de novo approach generates loops ab initio based on energetic constraints, whereas the knowledge- or template-based approach uses loops from known crystal structures as models for the new loops. Several groups have reported some success predicting protein loop conformations de novo (2-4), whereas others have proven equally successful using the template-based approach (5-8), and still others have employed a combination of the two approaches (9). As the data base of high resolution Fab structures grows, it becomes increasingly reasonable to model the combining sites of other antibodies based on these known structures using a template-based approach.

The observed conformations of the antibody CDR loops may be influenced by interactions with framework residues and the other hypervariable regions of the molecule. One potential disadvantage of the template-based prediction of the conformation of antibody CDRs is that variation in the association of the V_{L} and V_{H} domains (10) can alter the observed conformation. However, given the estimated coordinate error in the observed structures, the very simple template-based approach using canonical structures (1) provides an adequate description of the antibody CDRs and the general properties of the combining regions of DNA-binding antibodies. In particular, the distribution of positively charged residues within the combining region may indicate areas of the antibody that are likely to interact with the negatively charged sugar phosphate backbone of the DNA. If the orientation of the DNA backbone with respect to the combining region can be fixed, then antibody residues that are in position to interact with the DNA can be identified. The ability to predict the conformation of the CDR loops from amino acid sequence data makes it possible to derive plausible models for specific antibody-DNA interactions that can be tested experimentally and improved upon even in the absence of crystallographic data for the Fab and Fab-DNA complexes.

To extend our studies of DNA-binding antibodies, we have determined the sequences of the Fv domains of six murine anti-nucleic acid antibodies and used the canonical structure model (1, 5) to predict the loop conformations. Three of these antibodies were derived from autoimmune mice, whereas the others were prepared by immunization of mice with the respective nucleic acid polymer. These results allow us to compare autoimmune and induced antibodies on the basis of both sequence and structure to determine whether there are any obvious differences between the antibodies obtained from these two sources.

The crystal structures of three of the Fab fragments have been determined and are compared with the modeled structures of their variable domains. The structure of Fab Hed 10 (11) has been determined to 2.4-Å resolution, whereas the structures of Fab Jel 72 and Fab Jel 318 have been determined to resolutions of 2.7 (12) and 2.8 (13) Å, respectively. The pre-

*This work was supported by Canadian Medical Research Council and Arthritis Society of Canada grants (to J. S. L.), an Alberta Heritage Foundation for Medical Research studentship (to C. D. M.), and Grant DK42502 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

To whom correspondence should be addressed. Tel.: 615-322-4336; Fax: 615-343-0704.

1 The abbreviations used are: CDRs, complementarity-determining regions; V_{L} domain, variable light chain domain; V_{H} domain, variable heavy chain domain; r.m.s., root mean square.
Table I

| Antibody | Specificity | Source | Ref. |
|----------|-------------|--------|------|
| Hed 10   | Poly(dT)    | NZB/NZW | 14   |
| Hed 72   | Poly(dG)poly(dC) | C57Bl/6 | 12   |
| Jel 318  | Triplex DNA | C57Bl/6 | 15   |
| Jel 274* | dG-dC-rich duplex | NZB/NZW | 16   |
| Jel 242  | Duplex DNA  | NZB/NZW | 16   |
| Jel 201* | Poly(ADP-ribose) | C57Bl/6 | 45   |

* Jel 274 is a clone of antibody Jel 229.

** Jel 201 is a clone of antibody ADP-1.

** Predicted loop conformations for these Fab fragments are compared with those derived by crystallographic techniques. The predicted hypervariable loop conformations for the three antibodies that have not been structurally characterized are used to derive testable models of nucleic acid recognition.

** MATERIALS AND METHODS

** Hybridomas—Production and specificity testing of the hybridoma cell lines Hed 10, Jel 72, Jel 201, Jel 242, Jel 274, and Jel 318 have been previously described (Table I). Hybridomas were grown in RPMI 1640 medium (Life Technologies, Inc.) containing 8% fetal calf serum (Sigma), 50 μg 2-mercaptoethanol (Sigma), 0.1 μM sodium selenite (Sigma), and 50 μg/ml gentamycin (Roussel) prior to RNA isolation. Hybridoma supernatants were screened for antibody production and the retention of their original specificity with the aid of a solid-phase radioimmune assay (18).

** Antibody Isotyping—The immunoglobulin isotypes were determined by an enzyme-linked immunosorbent assay (19). The isotype-specific alkaline phosphatase-conjugated goat anti-mouse second antibodies (Life Technologies, Inc.) were used at the following dilutions: anti-mouse IgG1 at 1:10,000; anti-mouse IgG2a at 1:10,000; and anti-mouse IgG2b at 1:3750.

** Preparation of RNA—Total cellular RNA was prepared from hybridoma cell lines. Hed 10, Jel 72, Jel 201, Jel 242, Jel 274, and Jel 318 were grown in tissue culture by a modification of the procedure of Chomczynski and Sacchi (20). Briefly, cells were denatured by vigorous vortexing in 5 ml of 4 M guanidinium thiocyanate (Sigma), 25 μM Tris-HCl, pH 8.0, 0.5% N-lauroylsarcosine, and 0.1 M 2-mercaptoethanol. Following this, 0.5 ml of 3 M sodium acetate, pH 4.0, 5.0 ml of water-saturated phenol, and 1 ml of chloroform were added sequentially and mixed thoroughly after the addition of each reagent. The suspension was then cooled on ice for 15 min, followed by centrifugation at 10,000 xg for 20 min. The aqueous phase was transferred, and the RNA was precipitated using 1 volume of isopropanol, and the pellet was dried under a stream of nitrogen gas.

** Hybridoma supernatants were screened for antibody production and the retention of their original specificity with the aid of a solid-phase radioimmune assay (18).

** Model Building—An examination of the amino acid sequences of the anti-DNA antibodies revealed that many of the sequences possessed significant homology to antibodies with known crystal structures, particularly in their VH domains (see Fig. 1). In addition, most of the sequences examined, the hypervariable loops belonged to one of the canonical structures described by Chothia and Lesk (1). For this reason, we decided to attempt to model their combining sites using the loop designations and procedures described by them. The sequences of the hypervariable loops are presented in Table II. Also shown are those key residues that determine the loop conformation, the canonical structure to which each loop was assigned, and the name of the known structure from which the model was derived.

** Parent structures for the VH and VL frameworks were selected, and those amino acids that differed were replaced with the amino acids corresponding to the anti-DNA antibodies in such a way as to maximize side chain overlap. Model loops were grafted onto these framework structures by superimposing the backbone atoms (nitrogen, carbon, and α-carbon) of the 3 residues preceding and the 3 residues following the loop onto the corresponding atoms of the parent domains. For the special case of the H3 loop, models were selected based on their length and amino acid sequence homology. If an existing H3 loop of the same size was not available, then the nearest length loop was selected; the residue(s) that were deemed to be at the apex of the loop were added or excised; and the loop was closed by rotating the two newly exposed end residues about their Ψ and Φ torsion angles to bring the respective nitrogen and carbon atoms into conjugative bonding distance. A primary consideration in selecting a model H3 loop was whether both the modeled and existing loops contained arginine at position 94 and aspartic acid at position 101. In the McPC603 Fab structure, these 2 residues form a salt link at the base of the CDR. Thus, all of the modeled structures possessed the domain association as found in

** J. Tanha and J. S. Lee, unpublished results.
Sequences and Models of Anti-DNA Fv Domains

McPC603. The pseudo 2-fold rotation relating the \( \text{V}_{\text{L}} \) and \( \text{V}_{\text{H}} \) domains in McPC603 lies approximately midway between the extremes observed in known crystal structures of immunoglobulin Fab fragments. The resulting model Fv structures were then subjected to a limited energy minimization (30) to relieve unfavorable van der Waals contacts.

The electrostatic potentials of crystal structures and of the modeled Fv fragments were calculated using the program DELPHI implemented within version 2.1.0 of INSIGHT II (Biosym Technologies, Inc.) using only the protein formal charges. For these calculations, the protein dielectric was set to 2.0, and the solvent dielectric to 80.0. The radius of the solvent probe was 1.8 Å, and the ionic strength was set to 0.145 M.

RESULTS AND DISCUSSION

Sequences of Anti-DNA Antibodies—The sequences of all six antibodies are shown in Fig. 1. It is of considerable interest to compare the sequences, gene usage, and homologies within this group since they all bind to a similar antigen. In addition, three of the antibodies (Hed 10, Jel 274, and Jel 242) were produced from autoimmune mice, and the origin of this type of antibody is obscure. A summary of gene usage among these antibodies is shown in Table III.

The antibodies were found to be from three different \( \text{V}_{\text{H}} \) gene families. This result is consistent with earlier studies that have found anti-nucleic acid antibodies mainly in the 5558 family, but also in the 7183, S107 and VH10 families (36–41). Thus, there appears to be little restriction on the usage of these families. Brigido and Stollar (41) reported anti-DNA antibodies mainly in the 5558 family, but also in the 7183, S107 and VH10 families (36–41). Thus, there appears to be little restriction on the usage of these families. Brigido and Stollar (41) re-
Sequences and Models of Anti-DNA Fv Domains

A.3

sequences and models of anti-DNA Fv domains

as judged from the percent homology to known V\textsubscript{H} genes, there are similar degrees of somatic mutation in this group of antibodies, suggesting that the autoimmune antibodies have also arisen by antigen stimulation. Overall, it is difficult to distinguish the autoimmune from the induced antibodies on the basis of sequence or gene usage. Hed 10, Jel 274, and Jel 242 appear to be typical DNA-binding antibodies. The only common theme is that neither of the single strand-specific antibodies use the V\textsubscript{H} 4558 gene family.

Although not shown in Table III, the light chains of Hed 10 and Jel 201 are 91.1% homologous. By comparison, the light chains of Jel 72 and Jel 242, both of which bind to duplex DNA, are only 32.7% related. Since Hed 10 and Jel 201 share an idiotypic (44) it is probable that this structure resides on the light chain.

Crystal Structures—To assess the accuracy of the modeling methods that we have employed, modeled structures of Hed 10, Jel 318, and Jel 72 have been compared with their structures as determined by x-ray crystallography. The crystal structures of all three anti-DNA Fab fragments were determined by molecular replacement (28) and have been refined using the PROLSQ (29) and X-PLOR (30) systems of programs. For Hed 10 (space group P\textsubscript{2}\textsubscript{1}\textsubscript{2}\textsubscript{1}2\textsubscript{1}, with unit cell dimensions of \(a = 81.9, b = 159.2, \) and \(c = 41.0 \) \(\text{Å} \), and has been refined to a \(R \) factor of 0.204 for 9342 reflections (88% complete) in the 2.4- and 6.0-Å resolution range is 0.200. Jel 318 crystallizes in the space group P\textsubscript{2}\textsubscript{1}\textsubscript{2}\textsubscript{1}2\textsubscript{1}, with unit cell dimensions of \(a = 94.6, b = 102.6, \) and \(c = 92.4 \) \(\text{Å} \), but with two Fab fragments in the asymmetric unit. The current \(R \) factor for 19,136 reflections (90% complete) in the resolution range of 2.7 to 6.0 \(\text{Å} \) is 0.189 (13).

Comparison of Crystal and Model Structures—The root mean square (r.m.s.) deviations from aligned positions for the backbone atoms of both the conserved \(\beta\)-sheet framework residues and for the hypervariable loops as well as the r.m.s. deviations from alignment for all the atoms in the loops are listed in Table IV. Since the modeled structures of the H3 hypervariable loops are the most likely to be in error, their conformations are compared with the crystallographically determined conformations in Fig. 2. The positions of the backbone atoms of all of the modeled loops, with the exception of H3, are within 1.0-Å deviation of the crystallographically observed loop conformations. The backbone atoms of the H3 loops deviate by only slightly more than 1.0 Å. When all of the side chain atoms are included, the r.m.s. deviation from alignment for most of the loops increases to between 1.0 and 2.0 Å. The large deviations of the side chain atoms for the H3 loops and the H2 loop of Jel 318 can be attributed to radically different orientations of a few residues with large side chains (Lys, Arg, and Tyr). In a free Fab fragment, these problematic residues would, for the most part, be exposed to solvent, and the side chains of these residues may exist in multiple conformations. Additionally, crystal packing interactions influence the conformational changes of some of these side chains in the observed structures. The variable domains of both Hed 10 and Jel 318 pack closely against constant domains of symmetry-related molecules.

The crystal packing of Jel 318 involves contacts made by the H2 loop with the COOH terminus of a symmetry-related constant light chain domain (12). The conformation of lysine H53 is influenced by these interactions, with a possible hydrogen bond existing between the terminal nitrogen on the lysine side chain and the carbonyl oxygen of serine L208. The conforma-
tion of the side chain of this lysine differs markedly in the modeled loop, and it is this large difference in overall side chain orientation that contributes to the unusually large r.m.s. deviation for the atoms in this loop.

Despite some similarity in the crystal packing of Hed 10 and Jel 318, different residues make contacts. Among the crystal contacts in Hed 10 are made by residues of the L1 hyper-variable loop of the V_{L} domain interacting with a symmetry-related constant light domain. It is particularly interesting that the conformation of the L1 loop in Hed 10 was modeled satisfactorily, but the disposition of this loop with respect to the conserved framework residues of the V_{L} domain was different. When the modeled V_{H} and V_{L} domains are superimposed onto the crystal structures by their conserved framework residues, the differences in the positions of the \alpha-carbon atoms of the loops are indicative of whether the entire loop is shifted in its position (data not shown). The L1 loop of Hed 10 is moved inward, toward the combining region, by nearly 3.5 Å in the crystal structure relative to the modeled structure. This positioning of the L1 loop in the crystal allows it to make several crystal contacts and potential hydrogen bonds.

Immunoglobulin CDR loops are on the surface of the protein, as they must be to carry out their biological function. Because of this location, the CDR loops are also very likely to take part in crystal packing interactions. This places a significant constraint on the apparent accuracy one can hope to attain in predicting the structure of an antibody Fv domain. Both the CDR loop templates and the target structures are derived from crystal structures; therefore, crystal packing interactions can affect both the quality of the predicted model and our evaluation of its quality.

Another means to assess the quality of a model is to determine the degree to which two independent crystallographic structures are more similar to each other than to the model. The crystals of Fab Jel 72 contained two independent Fab molecules in the asymmetric unit, and these two molecules were refined independently. Comparison of the r.m.s. deviations between the backbone atoms of the CDR loops in these two copies of Jel 72 with those between the model and one of the Fab fragments (Table V) provides another assessment of the quality of the model. If one excludes H3, the r.m.s. deviations range from 0.1 to 0.6 Å between the two copies of Jel 72 in the crystal and from 0.2 to 0.6 Å between the model and one of the Fab fragments. The conformations of five of the six CDR loops are clearly predicted very well. If all of these factors are taken into account, satisfactory models have been derived for all of the

### Table III

| Antibody | Chain | Isotype | J gene | D gene | V\_\text{H} gene (\% homology) | V\_\text{subgroup} |
|----------|-------|---------|--------|--------|-----------------------------|--------------------|
| Hed 10   | H     | IgG2a   | JH2    | DSP2.3 | V\_\text{H} J606 (93.9%)    | Unknown            |
| Jel 72   | L     | \kappa | JK2    | Unknown| V\_\text{H} J558 (90.0%)    | V\_\text{\kappa}2  |
| Jel 318  | H     | IgG2a   | JH3    | DFL16.2| V\_\text{H} J558 (93.7%)    | V\_\text{\kappa}4  |
| Jel 318  | L     | \kappa | JK2    | Unknown| V\_\text{H} J558 (90.1%)    | V\_\text{\kappa}2  |
| Jel 274  | H     | IgG2a   | JH3    | DSP2.8 | V\_\text{H} J558 (87.0%)    | V\_\text{\kappa}2  |
| Jel 242  | L     | \kappa | JK2    |        |                             | V\_\text{\kappa}2  |
| Jel 242  | H     | IgG2    | JH2    |        |                             | Unknown            |
| Jel 242  | L     | \kappa | JK1    | DFL 16.1| V\_\text{H} 7183 (93.9%)    | Unknown            |
| Jel 201  | H     | IgG3    | JH3    |        |                             |                    |
| Jel 201  | L     | \kappa | JK2    |        |                             |                    |

* H, heavy; L, light.

### Table IV

| Antibody | Chain | Isotype | J gene | D gene | V\_\text{H} gene (\% homology) | V\_\text{subgroup} |
|----------|-------|---------|--------|--------|-----------------------------|--------------------|
| Hed 10   | H     | IgG2a   | JH2    | DSP2.3 | V\_\text{H} J606 (93.9%)    | Unknown            |
| Jel 72   | L     | \kappa | JK2    | Unknown| V\_\text{H} J558 (90.0%)    | V\_\text{\kappa}2  |
| Jel 318  | H     | IgG2a   | JH3    | DFL16.2| V\_\text{H} J558 (93.7%)    | V\_\text{\kappa}4  |
| Jel 318  | L     | \kappa | JK2    | Unknown| V\_\text{H} J558 (90.1%)    | V\_\text{\kappa}2  |
| Jel 274  | H     | IgG2a   | JH3    | DSP2.8 | V\_\text{H} J558 (87.0%)    | V\_\text{\kappa}2  |
| Jel 242  | L     | \kappa | JK2    |        |                             | V\_\text{\kappa}2  |
| Jel 242  | H     | IgG2    | JH2    |        |                             | Unknown            |
| Jel 242  | L     | \kappa | JK1    | DFL 16.1| V\_\text{H} 7183 (93.9%)    | Unknown            |
| Jel 201  | H     | IgG3    | JH3    |        |                             |                    |
| Jel 201  | L     | \kappa | JK2    |        |                             |                    |

* r.m.s. deviations for the backbone atoms of conserved framework residues (4-6, 9, 11-13, 19-25, 33-49, 53-55, 61-76, 84-90, and 97-106) using the residue numbering system of Kabat et al. (25).

* r.m.s. deviations for the backbone atoms of conserved framework residues (3-12, 17-25, 33-52, 55-60, 68-82, 88-95, and 102-112) using the residue numbering system of Kabat et al. (25).
FIG. 2. Conformations of modeled loops (thin lines) compared with crystallographically observed conformations (thick lines). The loops have been superimposed according to a least-squares fit of their backbone atoms. A, H3 loop from Hed 10; B, H3 loop from Jel 72; C, H3 loop from Jel 318. The r.m.s. differences in aligned positions for these loops are presented in Table II.

TABLE V
r.m.s. deviation in backbone atom positions

| CDR | No. of atoms | Model vs. Fab1* | Fab1 vs. Fab2 |
|-----|--------------|----------------|---------------|
| L1  | 36           | 0.62           | 0.20          |
| L2  | 9            | 0.16           | 0.07          |
| L3  | 18           | 0.47           | 0.27          |
| H1  | 21           | 0.49           | 0.40          |
| H2  | 15           | 0.49           | 0.85          |
| H3  | 39           | 2.64           | 0.53          |

* Fab1 and Fab2, Fab fragments 1 and 2, respectively.

loops, except H3. Considering the simplicity of the template-based modeling, even the H3 loop results are not discouraging.

In evaluating the Fv domain model structures or drawing conclusions from them, a couple of provisos must be kept in mind. One is that the CDR loops, especially H3, tend to have high temperature factors, indicating that motion or alternative conformations are possible. A second is that there may be conformational changes upon binding antigen. In the case of anti-DNA Fab BV04-01, the temperature factors for the H3 loop are reduced in Fab complexes with the trinucleotide d(pT3), and there is a significant alteration in the structure of the loop (31). The longer glycine-containing H3 loops, such as that of Jel 72, may be relatively flexible and assume a conformation that maximizes interaction with the DNA during binding. Nonetheless, a good approximation of the antibody combining region can be an advantage in designing further experiments.

DNA Binding Models—An examination of the relative disposition of the antibody CDRs and the electrostatic potential of the crystallographically determined structures has aided us in deriving testable hypotheses for how these antibodies may interact with their respective nucleic acid antigens. Descriptions of how these antibodies might interact with nucleic acids and suggested mutations to test their validity are presented in more detail elsewhere for Jel 318 (12) and Jel 72 (13). We present here models for the interaction with nucleic acids of the Fv domains whose structures we have derived from our sequence data.

It has been proposed that antibody combining sites are enriched in relatively solvent-exposed aromatic residues and that the burying of these large rigid side chains by antigen contributes to higher binding constants (33). DNA-binding antibodies
are also rich in aromatic amino acids, some of which no doubt interact with DNA. A comparison of the model and crystal structures shows that the solvent accessibility of the individual residues of each loop, including H3, is similar in both structures (data not shown). Aromatic amino acid residues pertinent to the binding of antigen could be discerned on the basis of their solvent accessibility.

An examination of the electrostatic surface potential of antibody combining sites can be useful in determining where charged groups on the antigen may bind. For DNA-binding antibodies, the primary concern is where the negatively charged phosphate groups of the DNA backbone will interact. Electrostatic surface potentials of the model Fv domains were calculated and compared with the surface potentials of the crystal structures, where applicable, in Fig. 3. Although there are some differences in the potential surfaces, the general features of the combining regions of the models do correspond with the calculated potentials of the observed structures.

Single Strand-specific Nucleic Acid-binding Antibodies—Of the six antibodies that we have examined, two are specific for single-stranded nucleic acids: Hed 10, which binds poly(dT), and Jel 201, which is specific for poly(ADP-ribose). The combining site of Hed 10 forms a long cleft roughly parallel to the VH-VL domain interface. The interior of this cleft is particularly enriched in tyrosines, mostly from the H3 loop, and is bordered on the VH side by a large ridge of positive electrostatic potential created by lysine and arginine residues in CDR-L1 and CDR-H2 (Fig. 3A). The current model places the sugar phosphate backbone of the DNA along this positively charged ridge while the thymine bases penetrate into the cleft and stack with the side chains of the tyrosines in the H3 loop.

The VH domain of Jel 201 is significantly homologous to Hed 10 (>90% amino acid identity), but its VH domain is different. Specifically, it has a shorter H2 loop and a much longer H3 loop, resulting in the modeled combining region forming a shallower cleft than in Hed 10. The H3 loop of Jel 201 was modeled from the H3 loop of McPC603, despite the fact that it lacks the aspartic acid residue at position H101 necessary to form the salt link with arginine H94. Although the model for this loop may not be correct in detail, the general features of the combining region suggest that adenine bases of the nucleic acid antigen may stack between the tyrosine residues and form specific hydrogen bonds with the side chains of the asparagine residues in CDR-H3. The ribose moieties would then interact with the polar residues in CDR-L1 and CDR-L3. This would place the phosphate groups of the nucleic acid in the region of positive potential centered on the H3 and L2 hypervariable loops (Fig. 3G).

Double- and Triple-stranded DNA-binding Antibodies—The combining regions of the double strand-specific (Jel 72, Jel 274, and Jel 242) and triple strand-specific (Jel 318) anti-DNA antibodies share several features that distinguish them from the single-stranded DNA-specific antibodies. The combining regions of Jel 72, Jel 274, Jel 242, and Jel 318 possess relatively flat surfaces that are bordered by knob-like ridges composed of residues in the H2 loop on the VH side or residues in the L1 or H3 loop on the VH side of the combining region. These projecting knobs often possess strong positive electrostatic potential, whereas there is commonly an area of negative potential near the L3 side of the combining region (Fig. 3, C–F). If the DNA were to bind with its helix axis parallel to the VH-VL domain interface, then it would be bracketed by the positively charged knobs, but the sugar phosphate backbone of the DNA would have to pass through the area of negative electrostatic potential. The possibility exists that this region serves as a cation-binding site that could act as a bridge between the negatively charged side chains and the DNA. In a heavy atom derivative of Jel 72, the cation UO$_4^{2-}$ binds with high occupancy in this area of the combining site around glutamic acid H50 at the base of CDR-H2. If cation binding is required for DNA binding, then neutralizing the electronegative areas by substituting glutamines or asparagines for glutamic or aspartic acid residues should result in lowered affinities for DNA, whereas a glutamic acid to arginine mutation may increase affinity for DNA.

Another possibility is that the DNA binds with its helix axis inclined to the VH-VL domain interface. The protruding knobs would then be in position to penetrate the major and minor grooves of the DNA and make specific contacts with the DNA bases, an idea first suggested by Stollar (34). In this orientation, the DNA helix would be mostly localized near the H3 side of the combining region, away from the negatively charged areas, and the mutations suggested above should have little effect on DNA binding. A consequence of this model is that it places the residues of the H3 loop within the major groove of the DNA. For Jel 242, substitutions can be made at aspartic acid L28 and glutamic acid H50 to discriminate between the two orientations. To increase the potential for specific hydrophobic bonding with the DNA bases, valine H99 could be changed to threonine, and tyrosine H98 substituted with asparagine or glutamine to increase specificity for DNA containing adenes or to arginine for greater guanine specificity. In the crystal structure of a zinc finger protein bound to DNA, arginines were observed to interact quite specifically with guanine bases (35). A similar type of interaction may be occurring in anti-DNA antibodies. The H3 loop of Jel 72, which binds the right-handed duplex of poly(dG)-poly(dC), is particularly arginine-rich, whereas the H3 loop of Jel 242, which binds B-DNA with little sequence specificity, contains only 1 arginine residue.

Jel 274 also binds B-DNA, but has a sequence preference for DNA that is GC-rich (16). The amino acid sequence of the VH domain of Jel 274 is significantly homologous to the VH domain of Jel 72 (>84% identity), whereas their VH domains are less similar. This observation, coupled with the fact that both antibodies bind very similar nucleic acids (Table 1), suggests that their VH domains are major determinants of the binding affinity for double-stranded DNA that contains guanine and cytosine bases. Both antibodies possess a long L1 and a short H2 hypervariable loop. The distribution of positive charge in the combining region (Fig. 3I) suggests that Jel 274 binds double-stranded DNA at angle, with the long axis of the DNA inclined to the VH-VL domain interface. In this orientation, the H3 loop is positioned within the major groove, whereas residues of the H2 loop are in position to interact with the minor groove of the DNA, similar to the model for Jel 72 binding A-type poly(dG)-poly(dC) (13). The arginine residue at position H99 could thus be a major determinant of sequence specificity by interacting specifically with the guanine bases of the DNA.

In summary, the sequences, gene usage, and model or crystal structures of six DNA-binding antibodies have been examined. The antibodies were obtained either from autoimmune sera or by immunization with nucleic acids. Overall, it is difficult to distinguish between these two sets of antibodies based on sequence or gene usage. Whereas no present modeling protocol can precisely predict Fv domain structure, the shape and overall charge distribution of the modeled combining regions generally correspond to those of the crystallographically determined structures. The nature of the antigens for the antibodies discussed here is that electrostatic forces are likely to play a significant role in the interaction with antibody. Therefore, identifying those regions of the antibody that are likely to interact with the negatively charged groups of the DNA backbone is an important first step in understanding antibody-DNA interactions. A simple knowledge-based approach to modeling the combining regions of anti-DNA antibodies serves to elucidate.
Sequences and Models of Anti-DNA Fv Domains
Fig. 3. Electrostatic surface potentials of modeled and observed structures superimposed on ribbon drawings of Fv domains. The view in all cases is looking directly down into the combining region, with the approximate positions of the hypervariable loops indicated. For all of the antibodies, the V\textsubscript{H} domain is shown in green, and the V\textsubscript{L} domain in yellow. Contours are displayed at +1 (blue) and −1 (red) kcal/mol. A, Hed 10 crystal structure; B, Hed 10 model structure; C, Jel 318 crystal structure; D, Jel 318 model structure; E, Jel 72 crystal structure; F, Jel 72 model structure; G, Jel 242 model structure; H, Jel 201 model structure; I, Jel 274 model structure.
potential mechanisms for recognition of nucleic acids. These models may aid in designing antibodies with enhanced or altered specificity for DNA by indicating key residues for substitution. If the affinity for specific DNA sequences can be increased, these antibodies may prove more efficacious for the crystallization of Fab-DNA complexes. The resulting detailed structures of Fab-DNA complexes will assist in the design of antibodies that not only bind to specific DNA sequences, but also catalyze cleavage of the DNA.

Acknowledgments—The sequencing expertise of Brenda Haug and Jamshed Tanha is gratefully acknowledged.

REFERENCES
1. Chothia, C., and Lesk, A. M. (1987) J. Mol. Biol. 196, 901
2. Brucoleri, R. E., and Karplus, M. (1987) Biopolymers 26, 137
3. Fine, R. M., Wang, H., Shenkin, P. S., Yarmush, D. L., and Levintal, C. (1986) Protein Struct. Funct. Genet. 1, 342
4. Moult, J., and James, M. N. G. (1986) Protein Struct. Funct. Genet. 1, 146
5. Chothia, C., Lesk, A. M., Tramantano, A., Levitt, M., Smith-Gill, S. J., Air, G., Sheriff, S., Padlan, E. A., Davies, D., Tulp, W. R., Colman, P. M., Steiner, S., Aizari, P. M., and Poljak, R. J. (1989) Nature 342, 877
6. de la Paz, P., Sutton, B. J., Dabley, M. J., and Rees, A. R. (1988) EMBO J. 7, 415
7. Chothia, C., Lesk, A. M., Levitt, M., Amit, A. G., Mariuza, R. A., Phillips, S. E. V., and Poljak, R. J. (1986) Science 233, 755
8. Smith-Gill, S. J., Mainhart, C., Levitz, T. B., Feldmann, R. J., Drohan, W., and Brooks, B. B. (1987) J. Mol. Biol. 194, 713
9. Martin, A. C. R., Chothiam, J. C., and Rees, A. R. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 9268
10. Steipe, B., Plichtchun, A., and Huber, R. (1992) J. Mol. Biol. 235, 739
11. Cigler, M., Boodhoo, A., Lee, J. S., and Anderson, W. F. (1987) J. Biol. Chem. 262, 643
12. Mol, C. D., Muir, A. K. E., Cigler, M., Lee, J. S., and Anderson, W. F. (1994) J. Biol. Chem. 269, 3615-3622
13. Mol, C. D., Muir, A. K. E., Lee, J. S., and Anderson, W. F. (1994) J. Biol. Chem. 269, 3605-3614
14. Lee, J. S., Dembovski, D. F., and Mason, T. R. (1992) Biochemistry 31, 4940
15. Lee, J. S., Burkholder, G. D., Latimer, J. P., Haug, B. L., and Braun, R. P. (1987) Nucleic Acids Res. 15, 1047
16. Zucan, R. P., and Lee, J. S. (1986) Nucleic Acids Res. 14, 5049
17. Lee, J. S., Woodswood, M. L., and Latimer, J. P. (1984) Biochemistry 23, 3277
18. Lee, J. S., Lewis, R. J., Morgan, R. A., and Singh, B. (1981) Nucleic Acids Res. 9, 1707
19. Wein, D. M., Herzenberg, L. A., Blackwell, C., and Herzenberg, L. A. (1986) Handbook of Experimental Immunology, Vol. 1, 4th Ed., Blackwell Scientific Publications, Inc., Cambridge, MA
20. Chomczynski, P., and Sacchi, N. (1987) Anal. Biochem. 162, 156
21. Ariv, H., and Leder, P. (1972) Proc. Natl. Acad. Sci. U. S. A. 69, 1408
22. Kaarring, N., Griffiths, G. M., Markham, A. F., and Milstein, C. (1986) Nature 304, 320
23. Bernet, F. C., Koest, T. F., Williams, G. J. B., Mayer, E. F., Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T., and Tasumi, M. (1977) J. Mol. Biol. 112, 535
24. Satow, Y., Cohen, G. H., Padlan, E. A., and Davies, D. R. (1986) J. Mol. Biol. 199, 555
25. Kabat, E. A., Wu, T. T., Reed-Miller, M., Perry, H. M., and Gottesman, K. S. (1987) Sequences of Proteins of Immunological Interest, 4th Ed., United States Department of Health and Human Services, National Institutes of Health, Bethesda, MD
26. Sub, S. W., Bhat, S. N., Navia, M. A., Cohen, G. H., Rao, D. N., Rudikoff, S., and Davies, D. R. (1986) Proteins Struct. Funct. Genet. 1, 74
27. Sheriff, S., Silverton, E., Padlan, E. A., Cohen, G. H., Smith-Gill, S. J., Finzel, B. C., and Davies, D. R. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 8675
28. Rosemann, M. G. (1980) Acta Crystallogr. Sect. A Found. Crystallogr. 46, 73
29. Hendrickson, W. A., and Konnert, J. H. (1980) in Biomolecular Structure, Conformation, Function and Evolution (Srinivasan, R., ed) Vol. 1, pp. 43-57, Pergamon Press, Oxford
30. Brungard, A., Kuran, J., and Karplus, M. (1987) Science 235, 458
31. Herron, J. N., Ho, X.-M., Ballard, D. W., Miller, R. P., Pace, P. E., Bothwell, A. L. M., Voss, E. W., Jr., and Edmondson, A. B. (1991) Proteins Struct. Funct. Genet. 11, 159
32. Deleted in proof
33. Paiflan, E. A. (1990) Proteins Struct. Funct. Genet. 7, 112
34. Stollar, B. D. (1986) CRC Crit. Rev. Biochem. 20, 1
35. Pavletich, N. P., and Pabo, C. C. (1991) Science 253, 809
36. Shlomchik, M. J., Ausin, A. H., Pieskacy, D. S., and Weigert, M. G. (1987) Proc. Natl. Acad. Sci. U. S. A. 84, 9150
37. Behar, S. M., and Scharff, M. D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 3970
38. Shlomchik, M., Massalley, M., Hain, R. A., Richard, M. Z., Pieskacy, D., Marashk, Rothstein, A., and Weigert, M. (1990) J. Exp. Med. 117, 265
39. Eilat, D., Webster, D. M., and Rosen, A. R. (1986) J. Immunol. 141
40. Radie, M. Z., Massalley, M. A., Eriksen, J., Hain, R., and Weigert, M. (1991) J. Immunol. 146, 176
41. Biggio, M. M., and Stollar, B. D. (1991) J. Immunol. 146, 2005
42. Trepocz, H., Jr., and Barrett, K. J. (1987) J. Immunol. 138, 2323
43. Bellon, B., Manheimer-Lory, A., Menestier, M., Moran, T., Dimitriu-Bona, A., Alt, F., and Bona, C. (1987) J. Clin. Invest. 79, 1044
44. Sibley, J. T., Latimer, J. P., and Lee, J. S. (1988) J. Immunol. 140, 3502
45. Sibley, J. T., Braun, R. P., and Lee, J. S. (1990) Clin. Exp. Immunol. 84, 563
46. Padlan, E. A., Silverton, E. W., Sheriff, S., Cohen, G. H., Smith-Gill, S. J., and Davies, D. R. (1989) Proc. Natl. Acad. Sci. U. S. A. 86, 5938
47. Herron, J. N., Ho, X.-M., Mason, M. L., Voss, E. W., Jr., and Edmondson, A. B. (1989) Proteins Struct. Funct. Genet. 5, 271