PROTEINS: Structure, Function, and Genetics 11:159–175 (1991)

An Autoantibody to Single-Stranded DNA: Comparison of the Three-Dimensional Structures of the Unliganded Fab and a Deoxynucleotide–Fab Complex

J.N. Herron,[1] X.M. He,[2] D.W. Ballard,[3] P.R. Blier,[3] P.E. Pace,[3] A.L.M. Bothwell,[3] E.W. Voss, Jr.,[4] and A.B. Edmundson[5]

[1]Department of Pharmaceutics, University of Utah, Salt Lake City, Utah 84108; [2]George C. Marshall Flight Center, NASA, Huntsville, Alabama 35812; [3]Section of Immunobiology and Howard Hughes Medical Institute, Yale University Medical School, New Haven, Connecticut 06510; [4]Department of Microbiology, University of Illinois, Urbana-Champaign, Illinois 61801; and [5]Harrington Cancer Center, Amarillo, Texas 79106

ABSTRACT **Crystal structures of the Fabs from an autoantibody (BV04-01) with specificity for single-stranded DNA have been determined in the presence and absence of a trinucleotide of deoxythymidylic acid, $d(pT)_3$. Formation of the ligand–protein complex was accompanied by small adjustments in the orientations of the variable (V_L and V_H) domains. In addition, there were local conformational changes in the first hypervariable loop of the light chain and the third hypervariable loop of the heavy chain, which together with the domain shifts led to an improvement in the complementarity of nucleotide and Fab. The sugar–phosphate chain adopted an extended and "open" conformation, with the base, sugar, and phosphate components available for interactions with the protein. Nucleotide 1 (5′-end) was associated exclusively with the heavy chain, nucleotide 2 was shared by both heavy and light chains, and nucleotide 3 was bound by the light chain. The orientation of phosphate 1 was stabilized by hydrogen bonds with serine H52a and asparagine H53. Phosphate 2 formed an ion pair with arginine H52, but no other charge–charge interactions were observed. Insertion of the side chain of histidine L27d between nucleotides 2 and 3 resulted in a bend in the sugar–phosphate chain. The most dominant contacts with the protein involved the central thymine base, which was immobilized by cooperative stacking and hydrogen bonding interactions. This base was intercalated between a tryptophan ring (no. H100a) from the heavy chain and a tyrosine ring (no. L32) from the light chain. The resulting orientation of thymine was favorable for the simultaneous formation of two hydrogen bonds with the backbone carbonyl oxygen and the side chain hydroxyl group of serine L91 (the thymine atoms were the hydrogen on nitrogen 3 and keto oxygen 4).**

Key words: anti-ss-DNA autoantibody, deoxynucleotide–Fab complex, conformational changes in protein when ligand is bound

INTRODUCTION

Nucleic acids tend to be nonantigenic except in certain autoimmune diseases like systemic lupus erythematosus (SLE). Specific inbred strains of mice spontaneously develop a syndrome similar to human lupus, which affects multiple organ systems. A central feature of lupus is the production of anti-DNA autoantibodies. These antibodies interact with DNA self antigens to form circulating immune complexes which deposit in the tissues and lead to the activation of complement and a general inflammatory response.

A murine antigen-binding fragment (Fab) from a monoclonal lupus-related autoantibody (BV04-01 IgG2b) has been crystallized in ammonium sulfate both in the native (unliganded) form and in complexes with oligonucleotides of deoxythymidylic acid. In this article we compare the three-dimensional structures of the unliganded Fab and the complex of the Fab with a trinucleotide.

In solution, the BV04-01 antibody was shown to bind thermally denatured calf thymus and plasmid DNA (single-stranded, ss) with a base preference for thymine.[1–4] Binding of 5′-phosphorylated oligodeoxythymidylic acids was not detectable until the chain length exceeded five nucleotides. Ligand binding was relatively insensitive to changes in ionic strength.[2] This observation suggested that electrostatic interactions of protein constituents with phosphate groups of the DNA ligand were less important

Accepted for publication March 21, 1991.
Address reprint requests to A.B. Edmundson, Harrington Cancer Center, 1500 Wallace Blvd. Amarillo, TX 79106.

than expected for stabilization of the immune complexes. In contrast another anti-ss-DNA autoantibody (Hed 10) recognizes two phosphate groups.[5,6] Its active site covers four bases, two of which must be thymine. Moreover, two other DNA-binding antibodies (Jel 241 and Jel 318) differ in their reactivities from BV04-01 and Hed 10. Jel 241 interacts with four phosphate groups and six base pairs on duplex DNA,[7] while Jel 318 binds to triplex DNA.[8]

When cocrystallized with the Fab of BV04-01, the ligands can be much smaller than those detectable in solution assays. For example, complexes of Fab with tri-, tetra-, and pentanucleotides have been crystallized, and the presence of ligands has been verified in each case by X-ray diffraction methods. Crystals of the trinucleotide–Fab complex were suitable for structural analysis at 2.7 Å resolution, but the others are still too small.

Since these ligands bind with low affinity, it will be interesting to compare the structures of the complexes with those of high-affinity Fabs, such as the 4-4-20 antifluorescyl Fab.[9,10] The ligand–Fab complex of 4-4-20 crystallized from 2-methyl-2,4-pentanediol (MPD) in a form nearly isomorphous with the ammonium sulfate form of the unliganded BV04-01 Fab. With the partially refined structure of BV04-01 Fab as the starting model, the structure of the liganded 4-4-20 Fab was solved by molecular replacement.

MATERIALS AND METHODS

Determination of Nucleic Acid Sequences Corresponding to the Variable (V) Domains of the Light and Heavy Chains

Nucleic acid sequences were obtained by two procedures. The sequence of RNA isolated from the hybridoma was determined by the dideoxy primer extension method, using reverse transcriptase and oligonucleotide primers.[11] Primers consisted of an oligonucleotide specific for the constant domain (C_L or C_H1) and other nucleotides complementary to sequences around codon 70 for V_L and codons 40 and 70 for V_H. To confirm the sequences, rearranged V_L and V_H genes were cloned from genomic DNA. A recombinant phage library was constructed in the EMBL-3 vector with DNA partially digested with *Mbo*I. Phages hybridizing with probes for the joining (J) regions were isolated and characterized by standard procedures. The DNA sequences were determined by the method of Maxam and Gilbert.[12] Amino acid sequences deduced from these results were partially checked by repetitive end-group analyses of L and H chains.

Determinations of the Structures of the Unliganded Fab and the Trinucleotide–Fab Complex

The murine IgG_{2b} antibody was isolated from ascites fluid and converted into Fabs by methods previously described.[3] The unliganded Fab and the complex were crystallized under identical conditions at 15°C and pH 7.6 in 1.7 M ammonium sulfate. Before crystallization the complex was produced by mixing the Fab with 1.5 molar quantities of pTpTpT (P-L Biochemicals, Inc., Milwaukee, WI) for 1 hr. For the unliganded Fab the space group was triclinic $P1$, with a = 58.4, b = 43.7, and c = 41.5 Å; α = 83.4°, β = 89.4°, and y = 84.4°. The complex crystallized in a different space group (monoclinic $P21$), with a = 63.5, b = 99.2, c = 36.6 Å, and β = 95.3°.

A single crystal of each protein was used to collect X-ray data with a Nicolet P21 diffractometer, operated at 40 kV and 35 mA (Cu$K\alpha$ radiation). Though relatively small (0.15 mm in the shortest dimension), the crystals of the unliganded Fab diffracted to d spacings of 2.0 Å. The 2.0-Å data set included 27,398 unique reflections, of which 19,716 (72.0%) were observed with intensities $I > 2.0\ \sigma(I)$, where σ is the standard deviation based on counting statistics. Intensities of 13,413 reflections were measured to 2.66 Å resolution for the complex; 8,624 (64%)

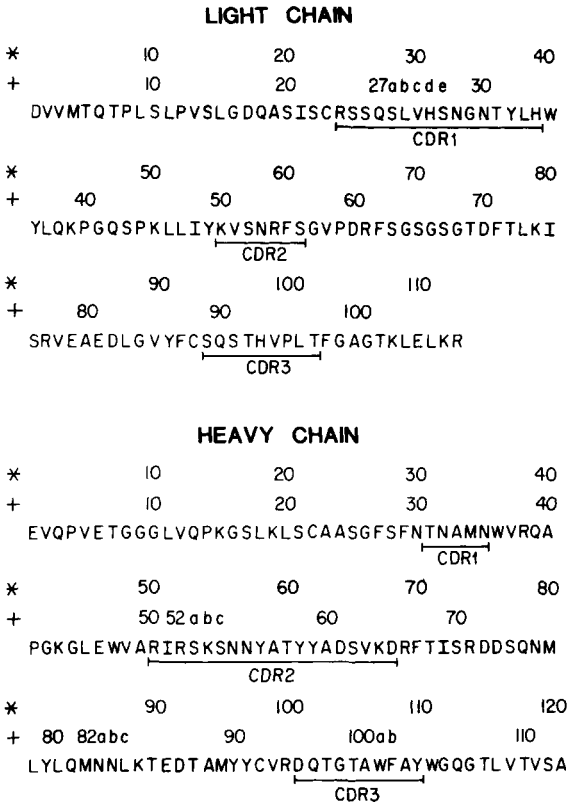

Fig. 1. Amino acid sequences of the "variable" domains of the light and heavy chains of the BV04-01 antibody, as deduced from nucleotide sequences. (*) The strict sequential numbering system. (+) A general numbering scheme introduced for immunoglobulins by Kabat et al.[26] The latter was used throughout the text. "Complementarity-determining regions" (CDR) in each chain are underlined.

TABLE I. Summary of Refinement Results*

Protein	Unliganded Fab	$d(pT)_3$–Fab complex	
Resolution limits (Å)	2.0–6.0†	2.66–6.0	
No. of reflections	18797	7759	
R-factor	0.246	0.191	
Average temperature factor ($Å^2$)	14.9	14.7	
	Actual rms deviation		Target value
rms‡ deviations from ideal distance (Å)			
Bond distance	0.019	0.018	0.025
Angle distance	0.032	0.033	0.030
Planar 1–4 distance	0.026	0.026	0.030
rms deviation from planarity (Å)	0.014	0.011	0.025
rms deviation from ideal chirality ($Å^3$)	0.179	0.155	0.150
rms deviation from permitted contact distance (Å)			
Single torsion contacts	0.199	0.190	0.300
Multiple torsion contacts	0.247	0.243	0.300
Possible hydrogen bond	0.213	0.227	0.300
rms deviation from ideal torsion angles (°)			
For planar group (0 or 180)	6.2	5.6	3.0
For staggered group (± 60 or 180)	25.5	23.0	15.0
For orthonormal group (± 90)	26.4	29.0	15.0

*Weighting factors for the refinements were obtained from the equation $\omega = (1/\sigma)^2$, where $\sigma = 40–268\ [\sin(\Theta)/\lambda - 1/6]$ for the unliganded Fab; and $\sigma = 45–642\ [\sin(\Theta)/\lambda - 1/6]$ for the $d(pT)_3$–Fab complex.
†Refinement statistics for the unliganded Fab at the same resolution limits (2.66–6.0 Å) as those of the complex: no. of reflections, 9676; *R*-factor, 0.206; average temperature factor, 15.0 $Å^2$; rms deviations from ideal values are nearly indistinguishable from those listed above.
‡rms, root mean square.

were observed at $I > 1.5\ \sigma(I)$. Data from 2.0 to 6 Å resolution (18,797 reflections) were used for refinement of the unliganded Fab, and data from 2.66 to 6 Å resolution (7,759 reflections) were used for the complex.

With the murine McPC 603 Fab[13] as the starting model, the structure of the unliganded Fab was determined by molecular replacement methods.[14–16] To solve the rotation function with the program MERLOT,[17] the Fab molecule was first divided into V_L–V_H and C_L–C_H pairs of domains. After the orientation of the Fab in the unit cell was established, the amino acid sequences of the L and H chains of the McPC 603 Fab were replaced with those of the BV04-01 molecule. The structure of the latter was initially refined with the program PROLSQ.[18,19] At intervals following 30 cycles of refinement, F_o–F_c OMIT maps[20] were calculated and skeletal models were interactively fitted into modules of electron density with the program FRODO.[21,22] About 95% of the amino acid residues were placed in plausible locations by this procedure of alternating refinement and model building. The structure was next subjected to five cycles of simulated annealing with the program X-PLOR.[23,24] Finally, the stereochemistry of the model obtained with X-PLOR was idealized by 30 additional cycles of PROLSQ. Following the last refinement, Luzzati plots[25] were used to estimate the mean error in the atomic coordinates. The structure of the Fab–$d(pT)_3$ complex was determined by molecular replacement with the unliganded Fab as the starting model.

Comparison of Tertiary Structures

Since the unliganded Fab (triclinic) and the Fab–$d(pT)_3$ complex (monoclinic) crystallized in different orientations, the comparison of the tertiary structures was initiated by superimposing the atomic coordinates of individual domains of the complex on those of the unliganded protein. Using the least-squares rotation and translation program ROTMOL, kindly supplied by W. Steigemann and R. Huber, superpositions were performed with the following sets of C_α coordinates from each domain: V_L, L1–L105; C_L, L106–L214; V_H, H1–H113; and C_H1, H114–H211. After superposition, the positions of each pair of α-carbons (C_α) in the two structures were compared with the molecular graphics program INSIGHT (BIOSYM Technologies, Inc.). The distance between each pair of C_α atoms (displacement) was computed with a program (RMSTRAN) written by one of the authors (J.N.H.). C_α displacement values were plotted versus sequence number. Regions were considered as significantly different in the two structures when the displacement value was three times larger than the error in atomic coordinates as estimated from Luzzati plots.

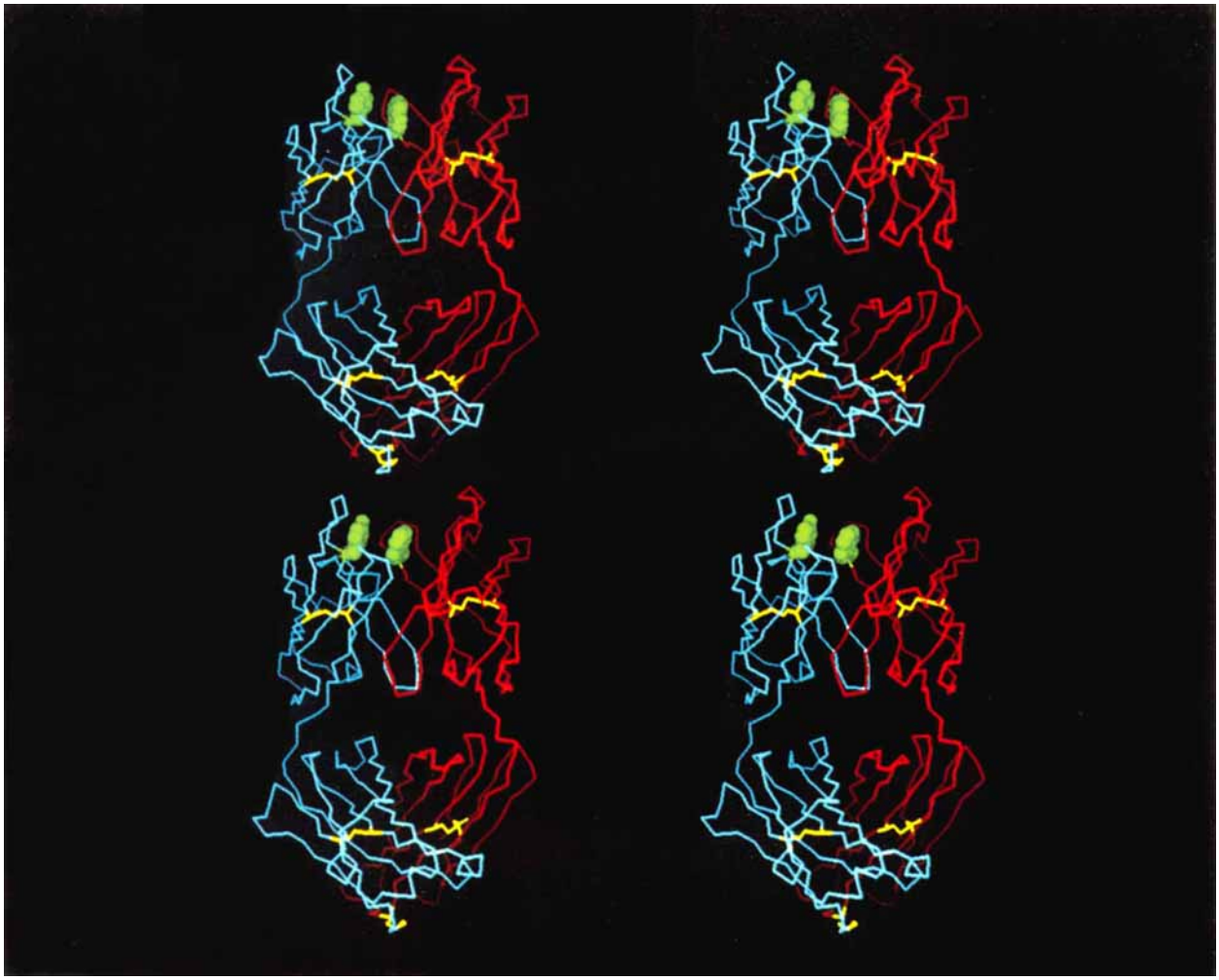

Fig. 2. **Top:** Stereo diagram of the C_α skeletal model of the unliganded Fab in triclinic crystals. The light chain is in blue and the heavy chain is in red. To identify the active site, space-filling models of two key residues (in green) are superimposed on the Fab: tyrosine L32 (CDR-1) of the light chain and tryptophan H100a (CDR-3) of the heavy chain. The inter- and intradisulfide bonds are represented by yellow connectors. **Bottom:** Stereo diagram of the C_α model of the protein moiety of the $d(pT)_3$–Fab complex (monoclinic crystals).

Analyses of Quaternary Structures

The overall quaternary structure of each molecule was initially evaluated by computing the elbow bend angle, which is defined as the angle subtended at the intersection of pseudodiads relating pairs of V and C domains. To estimate how closely these pseudodiads approached 180°, 39 C_α atoms from segments of β-pleated sheets of the V_L domain of each protein were rotated into the positions occupied by the corresponding atoms of the V_H domain. This procedure was repeated for the constant domain dimer. The pseudotwo-fold axes of rotation for the V and C pseudodiads were then used to compute the elbow bend angle.

Differences in quaternary structure were determined by first superimposing homologous structural elements (V_L–V_H dimers, C_L–C_H1 dimers, V_L domains, or C_L domains) and then computing whatever additional rotations and translations were required to superimpose any designated set of domains. C_α atoms from the following regions of each domain were used for comparisons: V_L, L2–L27b, L30–L105; C_L, L106–L151, L159–L167, L172–L212; V_H, H4–H7, H11–H24, H31–H95, H103–H113; and C_H1, H114–H124, H135–H155, H162–H185, H192–H203, H208–H210. If found in regions where there were changes in tertiary structure (see Table III), the appropriate C_α atoms were removed from calculations relating to quaternary structure.

Residues from the most ordered segments of β-pleated sheets of each domain were used to define a cylinder[26] with its long dimension corresponding to the X-axis. The Y-axis was drawn through a point in the middle of the first strand of one β-pleated sheet at residue L7 or H7 and the Z-axis was constructed orthonormal to X and Y. Rotations about X corre-

sponded to a "barrel roll" motion; rotations about *Y* denoted a change in the inclination (cant) of the cylinder; and rotations about *Z* defined a swivel motion in the plane of the central portion of the second β-pleated sheet (i.e., the one facing its counterpart across the pseudodiad).

With the V_L–V_H pairs overlaid as before, each cylinder in the complex was further adjusted until it coincided with the corresponding cylinder in the unliganded Fab. The least-squares rotation matrix for this additional adjustment was deconvoluted into the three rotations described above (roll, cant, and swivel).

RESULTS

Amino Acid Sequences of the V_L and V_H Domains

Amino acid sequences deduced for V_L and V_H from the nucleotide sequences are presented in Figure 1. The first 95 residues (V_K gene product) of the light chain are identical with those of TEPC 105 (subclass II[27,28]), although the J gene segments are dissimilar (J_K5 in BV04-01, J_K1 in TEPC 105). These genes, as well as the DQ 52 and JH3 minigenes encoding part of the BV04-01 heavy chain, are identical with known BALB/c germ line gene segments. Therefore, the BV04-01 antibody is probably encoded by sequences with few, if any somatic mutations.

The sequence of the V_L domain of BV04-01 differs from that of 4-4-20[29] in only six positions. Only two substitutions occurred in the three "complementarity-determining regions" (CDR), histidine L34 (CDR-1) for arginine in 4-4-20 and leucine L96 (CDR-3) for tryptophan (numbering system of Kabat et al.[28]; see Fig. 1). Arginine L34 neutralized the negative charge on the enolic group in the binding of the fluorescyl dianion in 4-4-20 and tryptophan L96 was one of the primary contact residues for the xanthonyl ring. Three other prominent members of the 4-4-20 active site, histidine L27d (CDR-1), tyrosine L32 (CDR-1), and serine L91 (CDR-3) were retained in the BV04-01 V_L domain. Histidine L93 (CDR-3) was not utilized for ligand binding in 4-4-20, but was present in the BV04-01 active site.

In the V_H sequences there were 42 differences, including numerous substitutions in all three CDRs (e.g., three of five residues in CDR-1 and seven of 19 in CDR-2). Alanine H33 (CDR-1) in BV04-01 replaced tryptophan, which was the major heavy chain contributor to the binding of the ligand in 4-4-20. In CDR-2, which was not directly involved in ligand interactions in 4-4-20, the most significant substitution was arginine H50 for glutamine in 4-4-20 (see later section). Arginine H52 and tyrosine H59, part of the lining of the binding groove in BV04-01, were found in both sequences.

While CDRs 1 and 2 were the same length in the two proteins, CDR-3 was three residues longer in BV04-01. The twin tyrosine residues (H100a,b) in the 4-4-20 active site were missing from the BV04-01 sequence. Conversely, the tryptophanyl–phenylalanine sequence (H100a,b) important in BV04-01 was absent in 4-4-20.

Collectively, five of eight residues in the high-affinity binding slot of 4-4-20 (two of five in V_L and all three in V_H) were replaced in BV04-01. These substitutions are consistent with the differences in the geometries of the binding sites and the biological specificities of the antibodies.

Crystallographic Refinements

The results of the crystallographic refinements of the structures of the unliganded Fab and the $d(pT)_3$–Fab complex are summarized in Table I. For idealization of the stereochemistry, the refinements were continued until the root mean square (rms) deviations from ideal bond distances decreased to 0.020 Å or lower. Mean error values in atomic coordinates, as estimated from Luzzati plots,[25] ranged from 0.25 to 0.35 Å in the unliganded Fab and from 0.20 and 0.30 Å in the complex.

Stereo diagrams of the C_α skeletal models of the two forms of the protein are presented in Figure 2. Light chains are shown in blue, heavy chains in red, and disulfide bonds in yellow. At the top of each model, the side chains of tyrosine L32 and tryptophan H100a are displayed as prominent markers of the nucleotide binding site.

Both the native and liganded forms of the Fab were almost fully extended, with "elbow bend" angles of 172° and 175°, respectively (see Table II). The pseudodiad angle relating the V_L and V_H domains increased from 174° in the unliganded Fab to 179° in the complex. Lower rotation angles (169°) were obtained for the C_L–C_H1 pairs in both structures.

Binding Site for Oligodeoxynucleotides

The active site in the BV04-01 antibody was found to be a large irregular groove between the tips of the V_L and V_H domains. At the maximum width of the groove the walls were 11.5 Å apart, a distance suitable for the accommodation of single-stranded, but not duplex DNA. Stereo diagrams of the groove before and after binding of $d(pT)_3$ are presented as solvent-accessible surfaces[30] in Figure 3. Constituents of the unliganded Fab are shown in the upper panel, with the light chain in blue and the heavy chain in red. The active site of the $d(pT)_3$–Fab complex is depicted in the middle panel with the model of $d(pT)_3$ removed for clarity. In the lower panel the binding site is displayed with the ligand in place. Note the increase in size and other differences appearing in the overall structure and the groove when the complex was formed.

All three CDR loops of the light chain and heavy chain (see Fig. 1) contributed to the binding site. The groove was lined by a cluster of aromatic side chains, aliphatic hydrophobic residues, four posi-

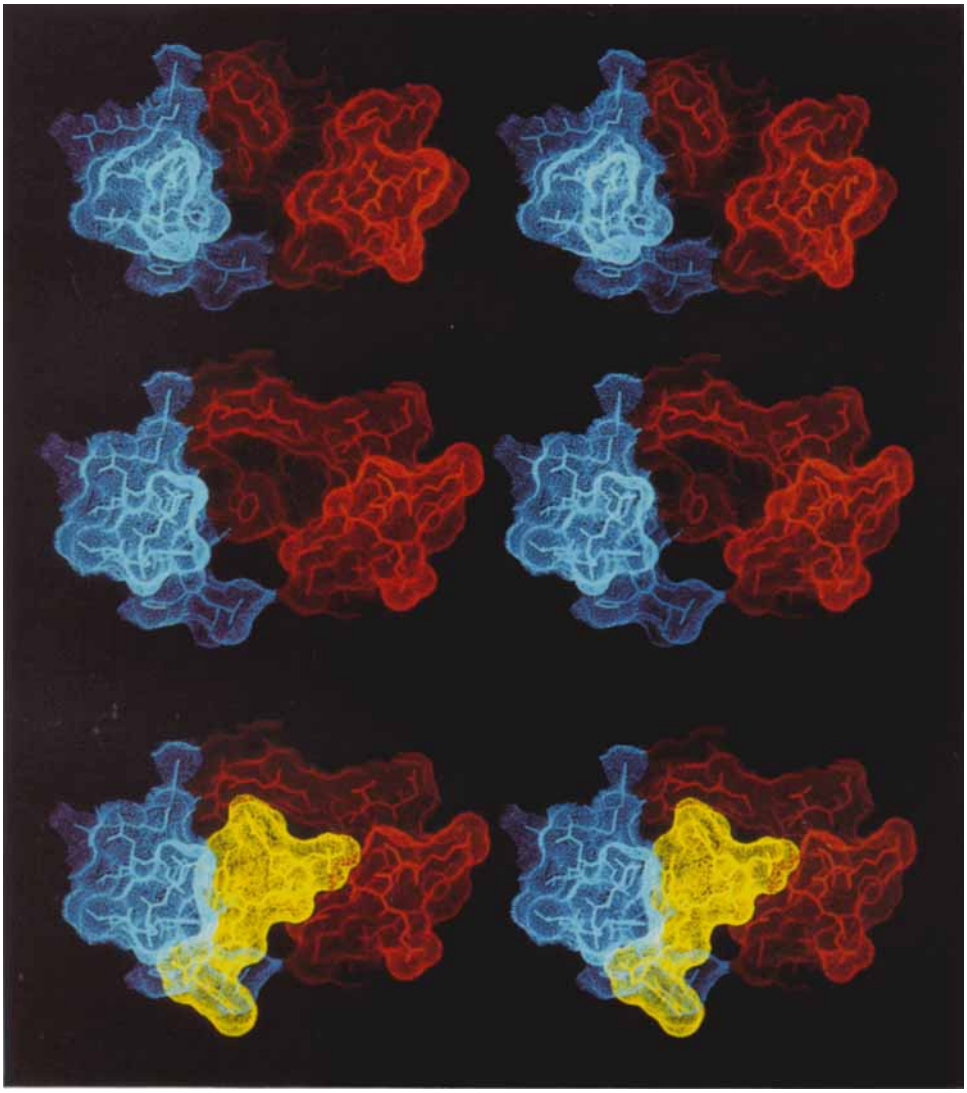

Fig. 3. **Top:** Stereo diagram of solvent-accessible surfaces[30] of residues lining the active site of the unliganded Fab. The light chain is in blue and the heavy chain is in red. **Middle:** Surface dot representation of the active site of the Fab in the complex with the $d(pT)_3$ removed. Note the expansion of the site and the changes in the heavy chain. **Bottom:** Binding site with $d(pT)_3$ in place (yellow).

tively charged side chains, and residues with polar groups suitable for hydrogen bonding. Light chain constituents included histidine L27d and tyrosine L32 from CDR-1, tyrosine L49 and lysine L50 from CDR-2, and serine L91, threonine L92, histidine L93, valine L94, and leucine L96 from CDR-3. The heavy chain contributed threonine H31 from CDR-1, arginine H50 and H52, serine H52a, asparagine H53, and tyrosine H58 from CDR-2, and aspartic acid H95, tryptophan H100a, phenylalanine H100b and alanine H101 from CDR-3.

As seen in the following section, not all of these residues participated in the binding of $d(pT)_3$. For example, the polypeptide backbone around alanine H101 and the four side chains of serine L91, leucine L96, arginine H50, and aspartic acid H95 provided a "false floor" over a potential cavity-type site in the V_L–V_H interface. The arginine and aspartic acid formed an ion pair directly over the cavity. These structural features thus favored the formation of a groove as the binding site for ssDNA.

Binding of the Ligand $d(pT)_3$

A stereo diagram of a skeletal model of $d(pT)_3$ fitted to "cage" electron density is shown in Figure 4 (upper panel). The cage density was contoured at $2\sigma(\rho)$, where σ is the standard deviation and ρ is the electron density, in an F_o-F_c OMIT map[20] in which the calculated structure factors did not contain contributions from the trinucleotide. The terminal 5′-

TABLE II. Elbow Bend and Pseudodiad Angles for the Unliganded Fab and $d(pT)_3$–Fab Complex*

Protein	V_L–H_H pseudodiad	C_L–C_H1 pseudodiad	Elbow bend angle
Unliganded Fab	174°	170°	172°
$d(pT)_3$–Fab complex	179°	170°	175°

*The elbow bend angle is defined as the angle subtended at the intersection of pseudodiads relating pairs of variable (V_L and V_H) and constant (C_L and C_H1) domains. Procedures for determining these angles are described in the text.

phosphate group, located on the lower left, was represented by the highest peak in the electron density map (peak height > 5σ(ρ)). Electron density modules assigned to the thymine base and deoxyribose ring of the central nucleotide, as well as parts of the remaining two phosphate moieties, were also prominent at the 4–5 σ(ρ) level in the map. Thymine was readily identified as a planar ring with three protuberances corresponding to carbonyl oxygen atoms in positions 2 and 4 and a methyl group in position 5 (these features are particularly clear in the central thymine in Fig. 4).

Skeletal models of $d(pT)_3$ and active site residues are shown in stereo in the bottom panel of Figure 4. Note that the flexible trinucleotide adopted an irregular conformation closely approximating the topography of the expanded active site (compare Fig. 4 with Fig. 3). This conformation was significantly different from the crystal structure of the dinucleotide sodium thymidylyl-5′,3′-thymidylate-5′-hydride,[31] which if extended hypothetically would have formed a helix with seven nucleotides per turn. In $d(pT)_3$ as bound, the thymine bases were in the *anti* configuration relative to the deoxyribose rings. The latter were assumed to have a C-2′-*endo* type puckering, although such structural details could not be conclusively demonstrated at the current resolution. Phosphate groups were arrayed on one side of the groove, with at least partial exposure to bulk solvent.

Nucleotide 1 (5′-end) was associated with the heavy chain, nucleotide 2 with both chains, and nucleotide 3 with the light chain (see bottom panels of Figs. 3 and 4). In Figure 4 hydrogen bonds between ligand and protein are indicated by green dots between atoms. Such interactions were assigned on the basis of distance criteria. Both hydrogen bonds and ion pairs were flagged if the distance between appropriate atoms was in the range of 2.6–3.2 Å. Other atoms were assumed to be in contact if they were separated by <4 Å (sum of van der Waals radii plus a tolerance of 0.4 Å).

There were 53 potential interactions of the protein with nucleotide 1, including 30 with thymine, 6 with deoxyribose, and 17 with the 5′-phosphate. For nucleotides 2 and 3 the interatomic contacts totalled 74 (thymine 66, deoxyribose 4, and phosphate 4) and 36 (thymine 17, deoxyribose 15, and phosphate 4).

In nucleotide 1 the 5′-phosphate group was partially immersed in a polar pocket consisting of threonine H31, alanine H33, serine H52a, and asparagine H53. Dotted green lines at the bottom of Figure 4 indicate hydrogen bonds (2.8 and 2.9 Å) from the hydroxyl group of serine H52a and the amide group of asparagine H53 to two of the oxygen atoms of the 5′-phosphate. A main chain carbonyl oxygen (threonine H31) was also 2.9 Å from a phosphate oxygen. The orientation of the phosphate group was also stabilized by 14 van der Waals contacts with protein atoms. Four of the six contacts of the deoxyribose ring (C2′, C3′, and C5′ atoms) were with the side chain of tryptophan H100a, and two (O3′) were with arginine H52. All atoms of thymine-1 were in contact with the protein except C2 and O2. One side of the ring (N1, C6, C5, and the C5 methyl group) interacted mainly with the side chain and backbone atoms of tryptophan H100a (total of 10 contacts). The N3, C4, and O4 atoms faced the backbone atoms of residues H96–H99. For hydrogen bonding, O4 was within 2.8, 3.2, and 3.1 Å of the amide nitrogen atoms of glycine H98, threonine H99, and alanine H100. Only the potential hydrogen bond with glycine H98 is marked with a dotted line in Figure 4. Collectively, the thymine-1 atoms were in contact with seven residues, all from CDR-3 (H95–H100a).

In nucleotide 2 the phosphate group interacted with arginine H52 to form an ion pair (2.8 Å; see Fig. 4). C1′ and C2′ of deoxyribose interacted weakly with the side chains of histidine L27d and tryptophan H100a but O3′ was within hydrogen bonding distance (3.1 Å) of one of the imidazolium NH groups of histidine L27d (this potential bond is not indicated in Fig. 4). Thymine-2 was intercalated between tyrosine L32 and tryptophan H100a, as illustrated in Figures 4 and 5. All rings in the triple stack were parallel, with a ring-to-ring distance of 3.5 Å. The stacking arrangement was reflected in the number and distribution of intermolecular contacts: 27 with tyrosine L32 and 26 with tryptophan H100a.

Additional interactions anchored the thymine-2. N3 was within hydrogen bonding distance (2.7 Å) of the backbone carbonyl oxygen of serine L91 and O4 was sufficiently close (2.9 Å) to the side chain hydroxyl group of L91 to form a second hydrogen bond. This double hydrogen bonding pattern is shown in Figure 4 and more clearly in Figure 5.

The insertion of histidine L27d between nucleotides 2 and 3 caused the sugar–phosphate chain to elongate and to bend. In nucleotide 3 the phosphate group was located in the general vicinity of histidine L27d (three contacts), but it was not strongly bound. The deoxyribose was situated between histidines

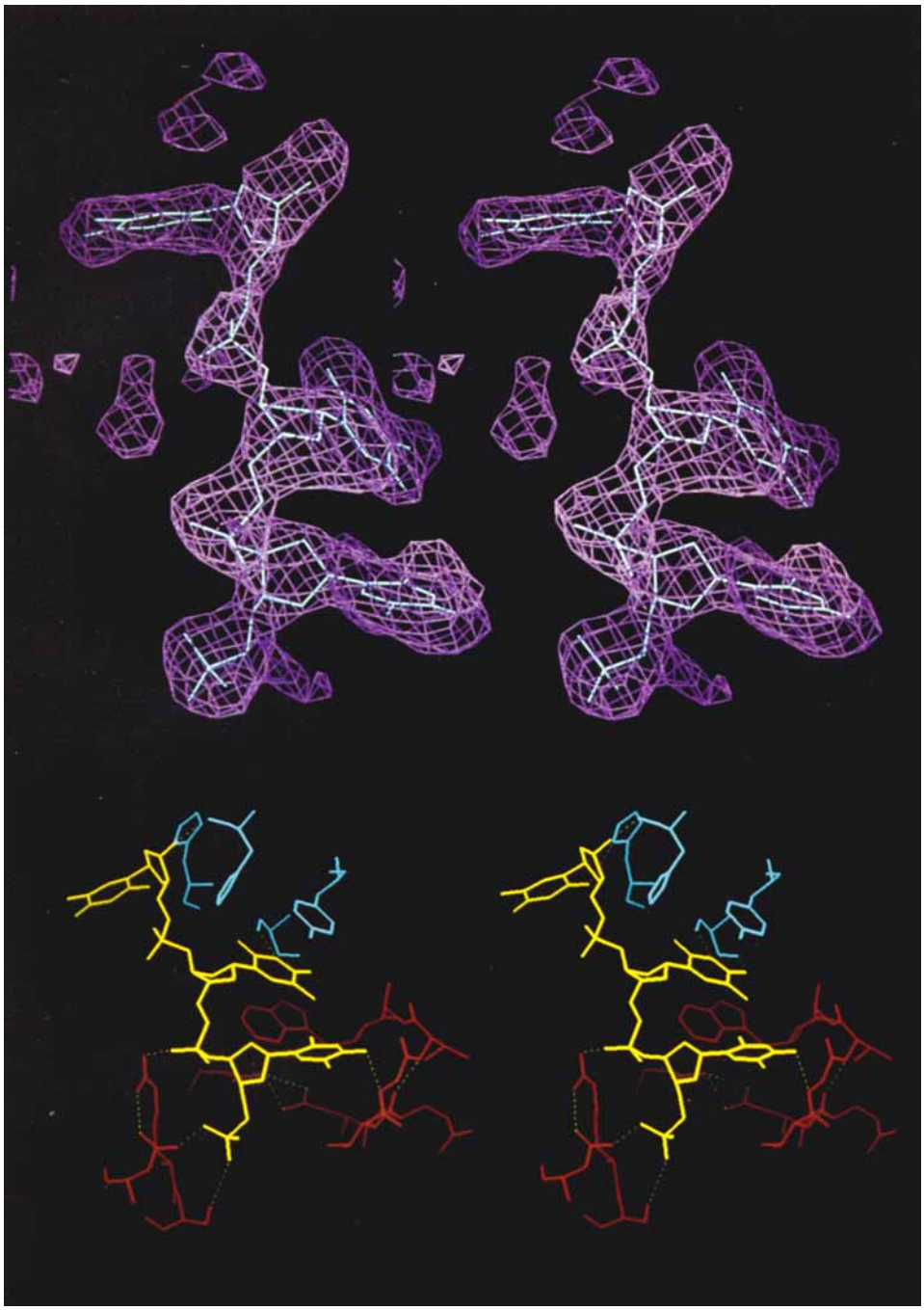

Fig. 4. **Top:** Stereo diagram of the three-dimensional "cage" electron density for $d(pT)_3$, with the model of the ligand superimposed. The F_o-F_c OMIT map,[20] from which this cage density was taken, was calculated without any contributions from the trinucleotide. The phosphate group at the 5′-end of the ligand is located at the bottom and the thymine at the 3′-end is displayed at the upper left. As bound, $d(pT)_3$ assumed an extended conformation different from the crystal structure of sodium thymidylyl-5′,3′-thymidylate-5′-hydrate.[31] **Bottom:** Stereo diagram of the skeletal model (yellow) of $d(pT)_3$ in the combining site of the Fab. The light chain components are in blue: histidine L93, histidine L27d, serine L91, and tyrosine L32 (left to right). Heavy chain constituents are in red. Charge–charge interactions or hydrogen bonds are designated by green dotted lines. To the lower left of the ligand arginine H52 participates in bidentate interactions with the oxygen of the central phosphate group and with the amide group of asparagine H53. The latter is also hydrogen bonded to one of the oxygen atoms of the 5′-phosphate. Serine H52a is hydrogen bonded to a second oxygen atom of the same phosphate. Behind the deoxyribose ring arginine H50 can be seen forming an ion pair (two dotted green lines) with aspartic acid H95 on lower right of the ligand. Other heavy chain components are tryptophan H100a, stacked with the central thymine base, alanine H100, threonine H99, glycine H98, hydrogen bonded to oxygen 4 of the lowest thymine base, threonine H97 and glutamine H96.

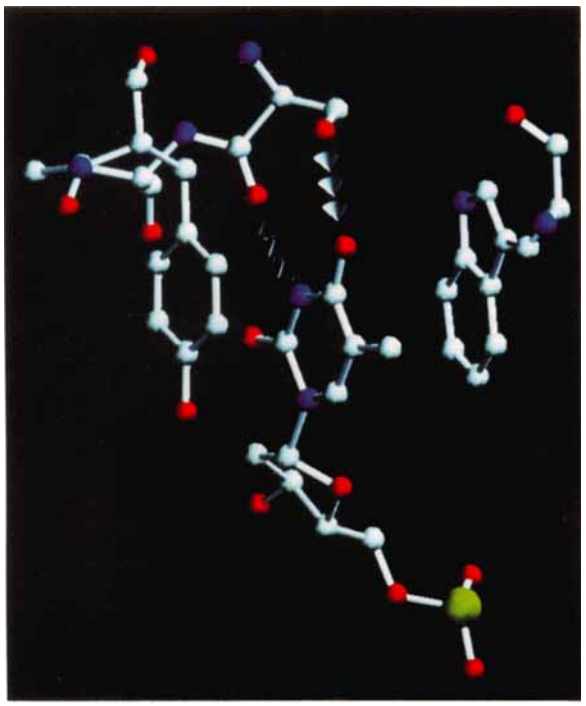

Fig. 5. A ball-and-stick model of the thymine base of the central nucleotide as bound to the Fab. The base is intercalated between tyrosine L32 on the left and tryptophan H100a on the right. Oxygen-4 of thymine is hydrogen bonded (white arrows) to the hydroxyl group on the side chain of serine L91, and the hydrogen atom of nitrogen 3 of the thymine is hydrogen bonded (gray arrows) to the carbonyl oxygen on the main polypeptide chain at serine L91. The 5′-phosphate of the nucleotide is shown on the lower right.

L27d and L93. C4′ and C5′ interacted with side chain atoms of histidine L27d, whereas C2′, O3′, and O4′ formed contacts with side chain and backbone atoms of L93 (see upper left part of Fig. 4). Thymine-3 was loosely associated with threonine L92, histidine L93, and valine L94 from CDR-3. Together C2 and O2 were close to 13 protein atoms and O2 was within hydrogen bonding distance (2.9 Å) of an imidazolium NH of histidine L93 (see Fig. 4).

Tertiary Structures of the Proteins in Two Crystal Forms

Differences in the tertiary structures are listed in Table III for regions in which the C_α displacement values were greater than 1.0 Å. This value was chosen because it was approximately three times the error in the atomic coordinates.[25]

Table III suggests that most differences in tertiary structure were attributable to packing interactions, except for those involved in ligand binding. The largest displacements occurred in the heavy chain segments involved in L–H interchain disulfide bonds (see H125–H134 in Table III). Half-cystine H128 was located at the distal end of the three-dimensional structure of the C_H1 domain. However, it was encoded by the first part of the C_H1 gene in a position characteristic of murine IgG_{2a} and IgG_{2b} subclasses, but unusual among IgG proteins in general (the "hinge" region gene normally includes this codon). The sequence of Ala-Pro-Gly-Cys-Gly-Asp-Thr-Thr-Gly-Ser for H125–H134 was compatible with the flexibility required to form the interchain disulfide bond. In monoclinic crystals the Thr-Thr-Gly-Ser part of the segment participated in close packing interactions with CDR-2 of an adjacent heavy chain. Interestingly, these packing interactions did not affect the tertiary structure of CDR-2. Note (in Table III) that glutamic acid L213 and half-cystine L214 also participated in sizable shifts, which were compatible with the changes in the heavy chain conformation around the disulfide bridge.

Ligand binding produced significant changes in the tertiary structures of CDR-1 and CDR-3 in the light chain and CDR-3 in the heavy chain. Of the light chain differences only those in CDR-1 (L27c–L29) were sufficiently pronounced to be listed in Table III. In the interaction of histidine L27d with the phosphoester group between nucleotides 2 and 3, the imidazole ring moved about 2 Å closer to the body of the V_L domain to accommodate the ligand. This displacement altered the conformation of the L27c–L29 segment and contributed to the expansion observed in the active site (see next section).

It seems unlikely that differences observed in the segment H25–H30 (Table III) in the heavy chains are correlated with ligand binding. In CDR-3 (H96–H102 in Table III), however, the conformational changes accompanied two interactions with the trinucleotide. (1) Thymine of nucleotide 1 was inserted into the space under residues H97–H99 (Thr-Gly-Thr). In response the Thr-Gly-Thr segment shifted 2.5 Å toward the exterior of the active site. (2) The indole ring of tryptophan H100a moved almost 4 Å to stack with thymine of nucleotide 2.

Differences in Quaternary Structure

The elbow bend and pseudodiad angles presented in Table II suggest that the quaternary structures of the Fabs and the V_L–V_H dimers in particular may be slightly different in the two crystal forms. To investigate this possibility further, the V_L–V_H dimer of the complex was superimposed (with the program ROTMOL) on the corresponding domain pair in the unliganded Fab (see Table IV and Figure 6). After superposition the rms displacement values (distance between pairs of C_α atoms in the two structures) were 0.74 Å in the V_L domains and 0.89 Å in the V_H domains. For comparison the rms deviations in C_α positions after superposition of individual domains were found to be 0.45 Å for V_L and 0.54 Å for V_H. When the V_L and V_H domains were overlaid as a dimeric unit, the displacement values were thus 1.6

TABLE III. Major Differences in Tertiary Structure Between the Unliganded Fab (Triclinic) and the $d(pT)_3$–Fab Complex (Monoclinic)

Region	C_α Displacement (Å) rms	Maximum	Comments
Light Chain			
L1 (D)	2.05	2.05	Packing interactions in triclinic
L27c–L29 (VHSNG)	1.35	1.71	CDR-1—moved to accommodate ligand binding in monoclinic
L152–L158 (GSERQNG)	1.79	2.81	Packing interactions in triclinic
L168–L171 (SKDS)	1.30	1.59	Packing interactions in monoclinic
L213–L214 (EC)	1.88	2.28	Interchain disulfide, moved to accommodate changes in the heavy chain
Heavy Chain			
H1–H3 (EVQ)	2.27	3.70	Packing interactions in triclinic
H8–H10 (GGG)	1.20	1.54	Packing interactions in triclinic
H25–H30 (SGFSFN)	1.62	2.34	H27-Phe packs with L9-Leu in triclinic, but is tucked into a hydrophobic pocket in monoclinic
H96–H102 (QTGTAWFAY)	1.86	2.65	CDR-3—moved to accommodate ligand binding in monoclinic, temperature factors lower in monoclinic
H125–H134 (APGCGDTTGS)	2.94	5.60	Flexible interchain disulfide region— packing interactions in monoclinic
H156–H161 (SGSLSS)	1.69	2.84	Flexible region, packing interactions in monoclinic
H186–H191 (STWPSQ)	1.19	1.52	Packing interactions and lower temperature factors in monoclinic
H204–H207 (TTVD)	0.94	1.04	Packing interactions in triclinic
H211 (E)	1.67	1.67	COOH-terminus—moved in concert with H186–H192

times greater than the rms deviations when the domains were superimposed individually.

Under the assumption that changes in quaternary structure involved rotations of domains, the displacement of atoms should progressively increase with the distance removed from the rotation axis. Therefore, the expected proportions of C_α displacements beyond a given value (e.g., 1.0 Å) should be correlated with the magnitudes of the domain rotations. Such displacements were calculated from the above data for each domain and listed in Table IV. For V_L, 13.5% of the selected C_α atoms were displaced more than 1.0 Å. This percentage increased to 20.0% for V_H.

Active site residues (CDRs) were then examined to see if quaternary changes could be partially distinguished from tertiary changes. The positions of key contact residues in the complex were compared after the variable domains were placed in 2 different orientations: (1) each V domain was overlaid independently on its counterpart in the unliganded form; (2) the V_L–V_H dimers were superimposed as a unit. This comparison was used to eliminate differences due to tertiary structure. All atoms in the selected active site residues were included in the displacement calculations, which are summarized in Table V. In general, the displacements due to alterations in quaternary structures were smaller than those attributable to changes in tertiary structure. Small quaternary effects (1 Å displacements) were noted in L-CDR-1, L-CDR-3, and H-CDR-2, but only minimal displacements were observed for H-CDR-3.

Next we considered movements of entire domains. The "roll, cant, and swivel" rotation angles are presented in Table IV. These rotation angles probably reflect small rearrangements in the domains to accommodate the ligand.

Similar procedures were applied to the C_L–C_H1 dimer and the results are summarized in Table IV. With the exception of the roll angle of C_H1 (1.8°), the additional rotations required for superposition of individual domains were very small. When considered with the observation that the pseudodiad angles between C_L and C_H1 were identical in the two structures (Table II), the data given in Table IV suggest that the quaternary structure of the C_L–C_H1 pair did not change significantly when the trinucleotide was bound.

For comparison, the analyses were repeated with just the V_L or C_L domains superimposed (see Table

TABLE IV. Differences in Quaternary Structure Between the Unligand Fab and the $d(pT)_3$–Fab Complex

Domain*	Additional adjustments necessary to superimpose domains†				Displacement§ (Å)
	Roll‡ (°)	Cant (°)	Swivel (°)	Transl (Å)	
	Variable domains (V_L and V_H) superimposed**				
V_L	−0.9	−0.2	−2.8	0.1	0.74 (13.5%)
V_H	−3.6	1.3	−1.2	0.1	0.89 (20.0%)
	Constant domains (C_L and C_H) superimposed				
C_L	−0.2	−0.2	0.4	0.0	0.51 (2.1%)
C_H1	1.8	0.3	0.2	0.0	0.60 (5.6%)
	VL domains superimposed††				
V_H	−5.7	2.0	−3.2	0.5	1.45 (62.0%)
	CL domains superimposed				
C_H1	2.1	0.6	0.3	0.1	0.65 (9.9%)

*C_α atoms from the following regions of each domain were used for comparisons: (V_L) L2–L27b, L30–L105; (C_L) L106–L151, L159–L167, L172–L212; (V_H) H4–H7, H11–H24, H31–H95, H103–H113; (C_H1) H114–H124, H135–H155, H162–H185, H192–H203, H208–H210. These regions were selected because their tertiary structures were similar in the two crystal forms (see Table III).

†Differences in quaternary structure were evaluated by first superimposing homologous structural elements [variable domain dimer (V_L and V_H), constant domain dimer (C_L and C_H), V_L domain, or C_L domain] and then computing the additional rotation (roll, cant, swivel) and translation (transl) required to superimpose a given domain.

‡The roll, cant, and swivel rotations are defined in the text.

§rms displacement between a domain in the $d(pT)_3$–Fab complex and the homologous domain in the unliganded Fab before the additional adjustment (roll, cant, swivel, transl) was applied. Values in parentheses are the percentages of α-carbons which exhibited a displacement of more than 1.0 Å. When each domain in the liganded Fab was superimposed on its homologous domain in the unliganded Fab, the following rms deviations in the superimposed α-carbons were obtained: 0.45 Å (V_L), 0.50 Å (C_L), 0.54 Å (V_H), and 0.56 Å (C_H1).

**After superposition of the V_L–V_H pairs of domains application of the method of Colman et al.[37–39] gave a value of −3.7° for the additional rotation of V_H relative to V_L.

††After superposition of the V_L domains alone, the procedure of Colman et al. produced a value of −6.1 Å. Both values are similar to the roll angles obtained in our calculations.

IV). As expected, the apparent differences between the V_H domains were more accentuated when the two structures were aligned in this way. The roll, cant, and swivel rotation angles were all increased, the rms displacement value rose to 1.45 Å, and 62% of the C_α positions differed by more than 1.0 Å from their counterparts in the unliganded structure. When the C_L domains were overlaid, the rotation angles and displacement values of C_H1 in the complex exhibited increases, but these were of lower magnitude than the ones registered for V_H.

Comparison of Crystal Packing of the Unliganded Fab and the $d(pT)_3$–Fab Complex

Crystal packing diagrams of the triclinic (unliganded Fab) and monoclinic (complex) forms are shown in stereo in Figure 7. To help identify the active sites of the unliganded Fab (upper panel), light green models of tyrosine L32 and tryptophan H100a are added to skeletal models of the light (blue) and heavy (red) chains. In the lower panel of Figure 7, the binding site can be located by the presence of a yellow model of $d(pT)_3$.

Inspection of Figure 7 helped resolve two important issues. (1) Triclinic crystals of unliganded Fab were found to contain solvent channels suitable for the transport of oligonucleotides through the crystal lattice. Moreover, the active site was relatively unobstructed for receipt of these ligands. We are currently attempting to solve the structures of complexes prepared by diffusion of tri- and hexanucleotides into such crystals. (2) In monoclinic crystals it was desirable to see if packing interactions were responsible for the observed conformational changes in BV04-01 on ligand binding. The $d(pT)_3$ ligand did not participate in contacts with adjacent molecules in the crystal lattice. Its presence therefore did not directly dictate the crystal packing. Moreover, CDR-3 in V_H did not participate in intermolecular contacts and only two residues of CDR-1 of V_L were in the interface of adjacent molecules (see next paragraph). Since the observed conformational changes were most accentuated among these constituents, it seems likely that the altered structure of the protein in the complex was responsible for the differences in crystal packing and not vice versa.

In both monoclinic and triclinic forms, the V_L domain of one molecule participated in head-to-tail interactions with the C_L domain in the next Fab (compare top and bottom panels in Fig. 7). Glycine L29 and asparagine L30 of CDR-1 faced the side chain of arginine L193 in the second molecule. Tyrosine L49,

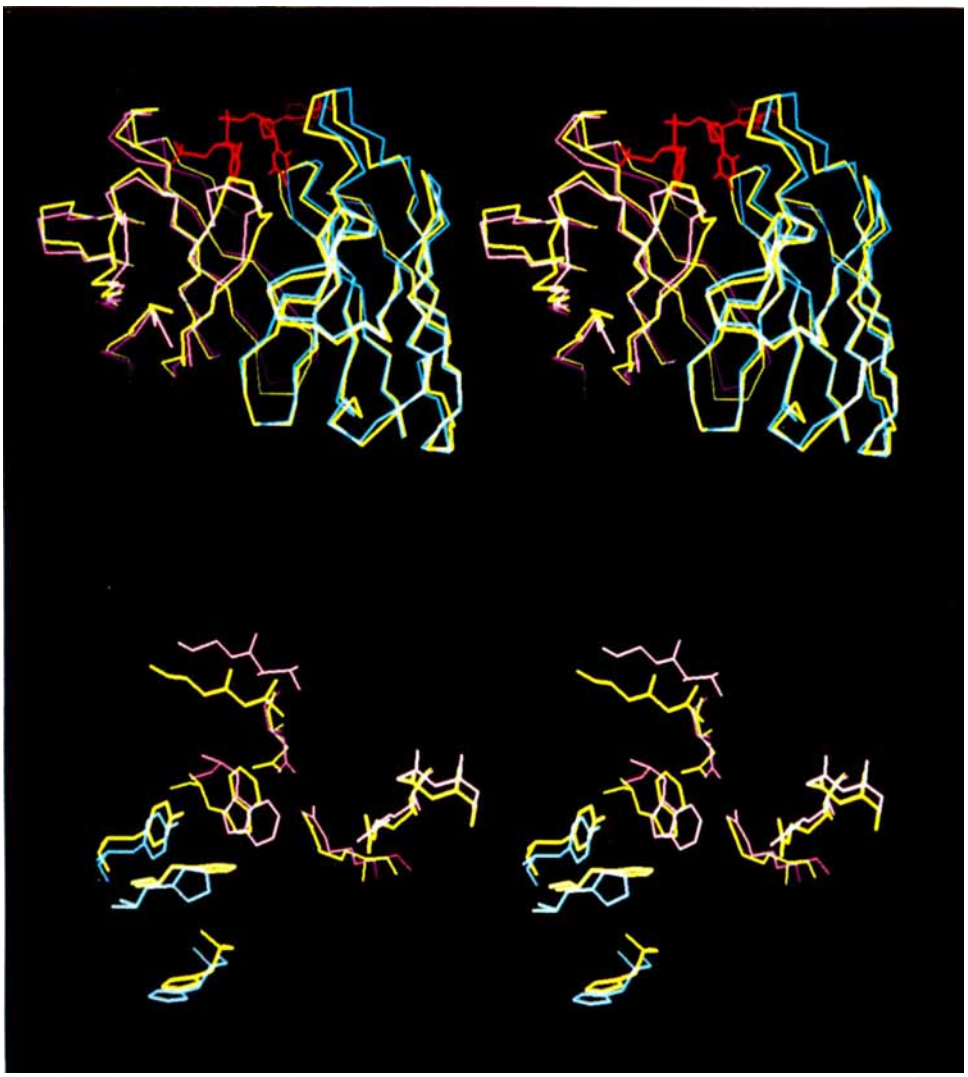

Fig. 6. **Top:** Stereo diagram of C_α models of the V_L–V_H dimers of the unliganded Fab and the $d(pT)_3$–Fab complex. Both domains of the unliganded protein are shown in yellow. In the complex V_L is blue and V_H is purple. A skeletal model of $d(pT)_3$ in red is included in the binding site. This photograph was obtained after superposition of the V_L–V_H pairs of domains with the program INSIGHT. The C_α atoms used in the alignment are listed in Table IV. This view emphasizes the expansion of the active site on ligand binding and the changes in tertiary structure of CDRs which line the active site. **Bottom:** Stereo diagram of key residues in the active site after superposition of the V_L–V_H pairs of domains of the unliganded Fab (yellow) and the $d(pT)_3$–Fab complex (light chain blue, heavy chain purple as above). From top to bottom the light chain residues are tyrosine L32, histidine L27d, and histidine L93. The heavy chain residues (yellow and purple) are glycine H98 (top left), threonine H97, aspartic acid H95 making an ion pair with arginine H50, tryptophan 100a (opposite tyrosine L32), arginine H52, and asparagine H53 (top right).

serine L52, asparagine L53, arginine L54, and aspartic acid L60 in and around CDR-2 maintained a highly polar interface with arginine L211 and glutamic acid L213, but the distances between them were too great (closest approach was 3.2 Å) for strong interactions. However, the side chains of lysine L50 and glutamic acid L187 were sufficiently close (2.9 Å) to form an ion pair. Triclinic crystals also exhibited a side-by-side packing motif in which the C_L domain was aligned with portions of the V_H and C_H1 domains of a second molecule (see Fig. 7). Arginine L155 of C_L was packed against the backbone atoms of threonine H7, glycine H8 and glycine H9 of V_H (closest approach of 3.1 Å of the guanidyl group to the amide nitrogen of glycine H8). Aspartic acid L151 of C_L was located opposite alanine H201 of C_H1.

In monoclinic crystals the CDR-2 loop was packed opposite threonine L202 of C_L and segment H131 to H134 (Thr-Thr-Gly-Ser) of C_H1 in a neighboring molecule. The amide groups of asparagines H53 and H54 formed a polar interface with the hydroxyl group of threonine L202. Constituents of CDR-2 and adjoining segments helped restrain the potential mobility of segment H131–H134 through a network of hydrogen bonds (as mentioned before, H131–

TABLE V. Displacement of Active Site Residues Due to Changes in Tertiary and Quaternary Structure

Sequence number	Amino acid	RMS displacement (Å) Tertiary structure	RMS displacement (Å) Quaternary structure
CDR-1 L27d	His	1.96	1.10
CDR-1 L27e	Ser	3.33	1.30
CDR-1 L28	Asn	3.18	1.31
CDR-1 L29	Gly	2.14	1.03
CDR-1 L30	Asn	2.14	0.97
CDR-1 L31	Thr	1.72	0.74
CDR-1 L32	Tyr	0.80	0.82
CDR-3 L91	Ser	1.19	0.72
CDR-3 L92	Thr	1.02	0.88
CDR-3 L93	His	1.10	0.93
CDR-3 L94	Val	1.07	0.90
CDR-1 H31	Thr	1.69	0.64
CDR-2 H50	Arg	1.06	0.51
CDR-2 H52	Arg	0.84	0.81
CDR-2 H52a	Ser	1.50	0.82
CDR-2 H53	Asn	1.36	0.99
CDR-3 H95	Asp	1.44	0.18
CDR-3 H96	Gln	2.29	0.23
CDR-3 H97	Thr	2.99	0.25
CDR-3 H98	Gly	2.59	0.11
CDR-3 H99	Thr	2.62	0.28
CDR-3 H100	Ala	2.71	0.33
CDR-3 H100a	Trp	2.53	0.53

H134 is part of the flexible region near the interchain -S-S- bond). For example, the hydroxyl group of tyrosine H55 was 3.0 Å from the backbone carbonyl oxygen of threonine H131 and 2.6 Å from the nitrogen of glycine H133. Side chains of lysine H52b and serine H74 were within hydrogen bonding distances (2.8 and 3.0 Å) of the side chain hydroxyl group and backbone carbonyl oxygen of threonine H132. The peptide nitrogen and carbonyl oxygen of arginine H71 participated in hydrogen bonds with the side chain hydroxyl group (3.0 Å) and backbone amide (2.7 Å) of serine H134. These packing interactions did not occur in triclinic crystals and segment H131–H134 exhibited a different tertiary structure (see Table III).

Similar observations were made with two other segments: L168–L171 (Ser-Lys-Asp-Ser) and H186–H191 (Ser-Thr-Trp-Pro-Ser-Gln). In both cases, the orientations of the segments in monoclinic crystals were fixed by packing restraints. It should be emphasized that having two versions of the Fab in different packing arrangements was very advantageous during the interpretation of electron density maps.

In most examples cited in this section, the predominant constituents in protein interfaces were polar side chains engaged in electrostatic interactions. These interactions included direct hydrogen bonding and ion pairing, as well as indirect interactions through intervening water molecules. In triclinic crystals the light chains were involved in substantially more packing interactions than heavy chains. The characteristic head-to-tail contacts were also observed in three crystal forms (two triclinic, one monoclinic) of the fluorescein–Fab complex of 4-4-20.[10] These interactions were maintained in three crystallizing media, ammonium sulfate, 2-methyl-2,4-pentanediol, and polyethylene glycol. It is now clear why the unliganded BV04-01 Fab and the ligand–Fab complex of 4-4-20 produce nearly isomorphous crystals.

DISCUSSION

Implications of Binding Patterns of $d(pT)_3$

The binding interactions of the nucleotides in $d(pT)_3$ help explain several observations on the structure and function of the BV04-01 autoantibody: e.g., the specificity for single-stranded rather than duplex DNA; the low affinity for ligands in solution, especially in fragments smaller than hexanucleotides; the base preference for thymine over purines and other pyrimidines like uracil; and the relative insensitivity of ligand binding to changes in ionic strength.

Figures 3 and 4 illustrate that the combining site in the unliganded BV04-01 molecule has to be expanded to accommodate even a single-stranded DNA fragment. This expansion can be considered as a conformational change "on demand." Although the present study does not define the limits of such demands, we suspect that combining sites recognizing duplex DNA have to be more voluminous than that of BV04-01 even in their resting states.

As a group, anti-DNA V domains tend to be relatively rich in arginine, particularly in the third hypervariable loop of the heavy chain.[32,33] Radic et al.[32] have suggested the appearance of arginine at multiple sites within the CDRs may be additive in leading to the binding of double-stranded DNA. In BV04-01, the arginine distribution does not fit the general pattern but instead is restricted to CDR-2 of the heavy chain. Moreover, one of the two arginines (H50) in CDR-2 participates in an ion pair, which stabilizes the topography of the active site but does not contribute directly to ligand binding. Thus the low frequency and limited involvement of arginine in BV04-01 are compatible with its specificity for single-stranded DNA.

Arginine can directly interact with three structural elements of DNA, the phosphodiester backbone, base-paired guanine, and unpaired or based-paired cytosine.[34,35] In BV04-01, such interactions were limited to the phosphodiester backbone, and only one arginine side chain (H52) was actually utilized for binding. Lysine residues were not present in suitable locations to interact with the phosphate groups and histidine side chains were not protonated at the pH (7.6) of the binding studies. The paucity of strong electrostatic interactions probably

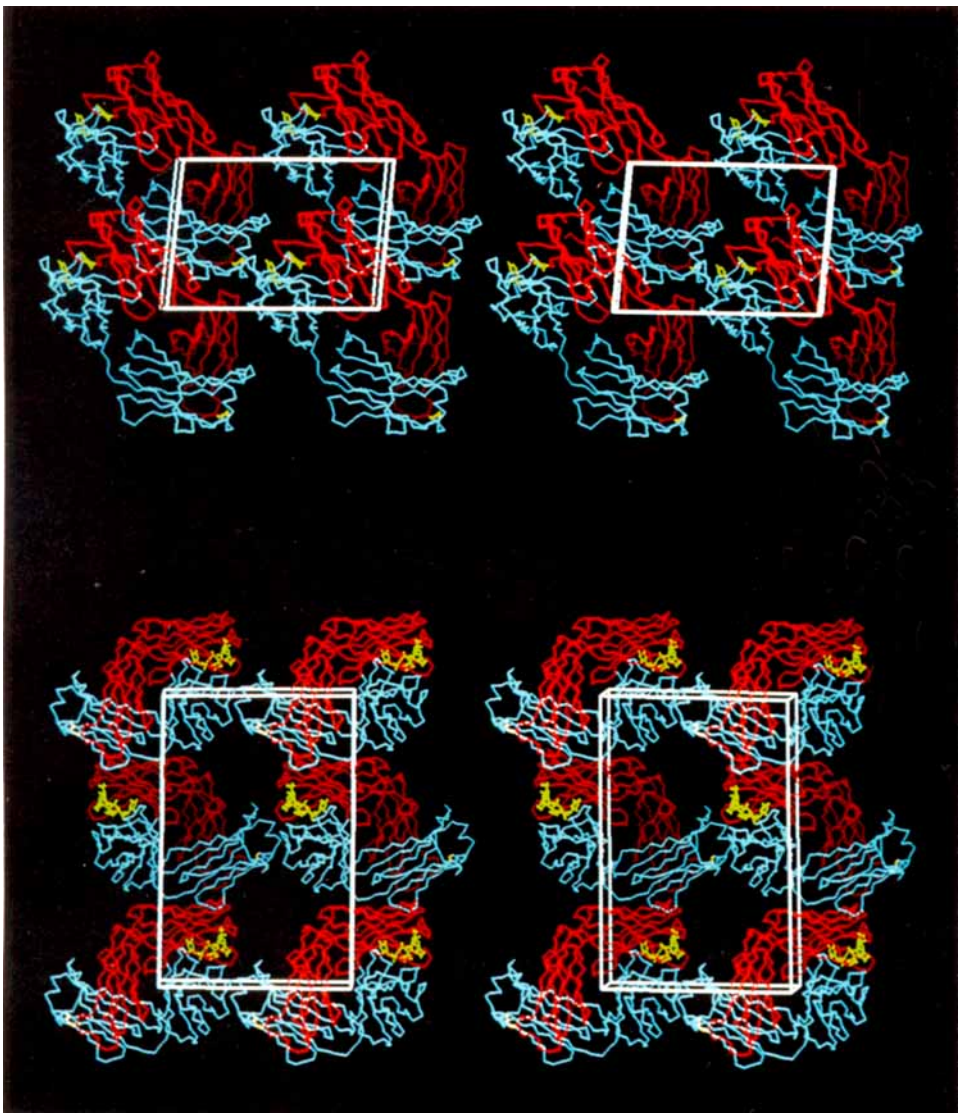

Fig. 7. **Top:** Packing diagram of unliganded Fab molecules in a triclinic space group. An outline (white lines) of the unit cell is superimposed (the front face is the *ab* plane). Light chains are in blue and heavy chains are in red. The binding site is labeled by light green models of tyrosine L32 on the light chain and tryptophan H100a on the heavy chain. At the other end of each molecule the interchain disulfide bond is shown as a yellow connector between the light and heavy chains. **Bottom:** Packing diagram of molecules of the $d(pT)_3$–Fab complex in a monoclinic space group. In the unit cell the *ab* plane is the front face. The ligand is in yellow.

accounted for at least part of the low affinity of BV04-01 for $d(pT)_3$ and other oligodeoxynucleotides and also explained the minimal effects of high salt concentrations on ligand binding in solution. These properties proved to be advantageous for the present work, because high concentrations of ammonium sulfate could be used to produce cocrystals of ligand and protein.

With the possible exception of the deoxyribose in nucleotide 3, the sugar moieties did not participate in sufficiently large numbers of interactions to have major effects on the binding pattern of $d(pT)_3$. However, deoxyribose rings 1 and 2 bracketed the indole ring of tryptophan H100a. The C2′ atom of deoxyribose 1 was in van der Waals contact (3.4 Å) with CZ3 of tryptophan. If these two sugars were replaced with ribose as in RNA, it seems likely that potential interactions with tryptophan would be compromised by steric restraints accompanying the addition of a hydroxyl group to C2′.

In comparison with the phosphate and sugar moieties, the thymine bases participated in relatively large numbers of contacts with the protein. The key docking interactions were provided by the central thymine, which could probably be considered the "immunodominant" part of the ligand. Thymines-1

and -3 were involved in fewer and less specific interactions.

The preference of BV04-01 for pyrimidines rather than purines may be attributable to a combination of factors. Adenine would not be favored in nucleotide 2 because the exocyclic amino group in position 6 would not be compatible with the double hydrogen bonding pattern between thymine and serine L91. Like thymine, guanine has a keto oxygen (position 6) adjacent to an endocyclic amino group (position 1), but the nucleotide would have to be translated slightly from its present location to make two hydrogen bonds with serine. A guanine in nucleotide 1 would also have to be moved to maintain the hydrogen bond between the keto oxygen and the protein. Guanines placed in adjacent nucleotides would not satisfy the stringent spacial requirements in the BV04-01 site [if fitted onto the sugar—phosphate backbone of $d(pT)_3$, the purine bases would overlap]. Moreover, the *syn* orientation is favored in deoxynucleotides of guanine.[36] Pyrimidines prefer the *anti* configuration where there is minimal steric hindrance between base and sugar. The *anti* configuration seemed to be necessary for the binding of the first two bases of $d(pT)_3$.

The specificity conferred by thymine was underscored by the observation that substitution of uracil in 5′-phosphorylated oligonucleotides led to a 20-fold decrease in the binding constant with BV04-01.[2] In contrast to uracil, oxygen 4 of thymine would be more electronegative than oxygen 2 because of the tendency of the 5-methyl group to push electrons into the pyrimidine ring. Oxygen 4 would therefore be more likely to form hydrogen bonds with the protein. This predicted hydrogen bonding pattern was observed in thymines-1 and -2 of the $d(pT)_3$–Fab complex. The 5-methyl group also contributed to the van der Waals interactions of bases-1 and -2 (eight and three contacts), but none to thymine-3.

In summary the binding of $d(pT)_3$ illustrated a variety of interactions that constituted a basic package for low-affinity binding of single-stranded DNA. For more selective recognition and higher affinity binding, it seems desirable for charge–charge and double hydrogen bonding interactions to play more dominant roles. In BV04-01 the charge–charge interactions were limited by the number and distribution of arginines, but there were apparently compensatory features which ensured the effectiveness of the antibody. For example, binding capacity for additional nucleotides was still available, since $d(pT)_3$ did not occupy all of the space in the groove. Moreover, solution assays indicated that relative binding constants were elevated 5- to 10-fold when the ligand was increased in size from 10 to 18 nucleotides.[4] This finding is especially pertinent to the binding of DNA in vivo and may reflect a combination of interactions inside and outside the designated active site.

Structural Changes Accompanying Ligand Binding

The present results support the proposal of Colman et al.[37–39] that antigen binding can be accompanied by movements of V domains to improve the complementarity of the interacting molecules. In BV04-01 the rearrangements of the V domains were similar in magnitude to those observed when the unliganded D1.3 Fv (V_L–V_H dimer) was compared to the complex of D1.3 Fv with lysozyme.[40] However, much larger changes in the local conformations of the CDRs were observed for BV04-01. Small but significant adjustments in the side chains and backbone of the Fab active site were observed on binding of the 19-residue peptide corresponding to the C-helix of myohemerythrin.[41] Interestingly, the bound peptide assumed the conformation of a type II β-turn rather than an α-helix.

Other examples of locally "induced fit" models include the complexes of HyHel-5[42] and HyHel-10 Fabs[43,44] with lysozyme. In the D1.3 Fab the binding of lysozyme was classified as a "lock-and-key" model and few, if any, conformational changes were observed.[45] Similar conclusions were reached in the X-ray analyses[46] of six complexes of a human light chain dimer with peptides preselected for binding activity by the peptide scanning procedures of Geysen et al.[47] When the peptide showing the highest affinity for the protein was deliberately modified by the addition of two β-alanine residues to the carboxyl end, however, the structure of the ligand was changed from a compact, internally stabilized conformation to an extended form which produced significant distortions of the protein (induced fit model). In the case of the dimer it was thus possible to incorporate design features which directed otherwise similar ligands to bind by different mechanisms. We are attempting to reach the same level of sophistication with antibodies like BV04-01.

Possible Correlations of the Present Work With Features of Systemic Lupus Erythematosus (SLE)

Autoantibodies binding DNA contribute to the formation of the lesions in lupus and act as important indicators of the course of the clinical disease.[48] The origins of these autoantibodies are still obscure.

Thymine and guanine tend to be immunodominant in many of the DNA autoantigens in SLE.[52] If BV04-01 is typical, the principal feature responsible for the prominence of thymine is probably the double hydrogen bonding potential in positions 3 and 4, especially when the ring is held in a fixed orientation by stacking interactions. The immunodominance of guanine could be explained in the same way, since it also has favorable geometry for the formation of two hydrogen bonds through a ring nitrogen and a keto oxygen in positions 1 and 6.

Single-stranded DNA and oligonucleotides are immunogenic, but the elicited anti-DNA antibodies are not always closely similar to the autoantibodies in SLE. Such observations have led to suggestions that the primary specificities of the anti-DNA autoantibodies might be directed toward other antigens, such as bacterial surface antigens or recurrent epitopes on macromolecules from diverse sources (reviewed in 48). Phosphocholine[49,50] and the K30 capsular polysaccharide of *Klebsiella pneumoniae*[51] are examples of bacterial antigens eliciting antibodies related to anti-DNA autoantibodies.

Two properties of the BV04-01 antibody are compatible with possible stimulation by antigens other than DNA: the low affinity for oligonucleotides and the extensive conformational changes when confronted with relatively small ligands. The loss of tolerance to self may be partly attributable to the presence of antibodies that appear normal until they assume a "new" (self-reactive) conformation on exposure to single-stranded DNA or its fragments.

ACKNOWLEDGMENTS

We thank Luke Guddat, Leif Hanson, and Lin Shon for many stimulating discussions, Lynne Miller for aid with the bibliography, Twyla Slay and Kristin Saunders for typing the manuscript, Silicon Graphics for gifts of graphics work stations, and Biosym Technologies for supplying their Insight and Discover programs. This work was supported by N.I.H. Grants AI 22898 (to J.N.H.), AI 20960 (to E.W.V.), and CA 19616 (to A.B.E.), by the Howard Hughes Medical Institute (to A.L.M.B.), and by the Harrington Cancer Center (to A.B.E.)

REFERENCES

1. Ballard, D.W., Lynn, S.P., Gardner, J.E., Voss, E.W., Jr. Specificity and kinetics defining the interaction between a murine monoclonal autoantibody and DNA. J. Biol. Chem. 259:3492–3498, 1984.
2. Ballard, D.W., Voss, E.W., Jr. Base specificity and idiotype of anti-DNA autoantibodies with synthetic nucleic acids. J. Immunol. 135:3372–3386, 1985.
3. Gibson, A.L., Herron, J.N., Ballard, D.W., Voss, E.W., Jr., He, X.-M., Patrick, V.A., Edmundson, A.B. Crystallographic characterization of the Fab fragment of a monoclonal anti-ss-DNA antibody. Mol. Immunol. 22:499–502, 1985.
4. Smith, R.G., Ballard, D.W., Blier, P.R., Pace, P.E., Bothwell, A.L.M., Herron, J.N., Edmundson, A.B., Voss, E.W., Jr. Structural features of a murine monoclonal anti-ss-DNA autoantibody. J. Indian Inst. Sci. 69:25–46, 1989.
5. Cygler, M., Boodhoo, A., Lee, J.S., Anderson, W.F. Crystallization and structure determination of an autoimmune antipoly (dT) immunoglobulin Fab fragment at 3.0 Å resolution. J. Biol. Chem. 262:643–648, 1987.
6. Anderson, W.F., Cygler, M., Braun, R.P., Lee, J.S. Antibodies to DNA. BioEssays 8:69–74, 1988.
7. Braun, R.P., Lee, J.S. Variations in duplex DNA conformation detected by the binding of monoclonal autoimmune antibodies. Nucl. Acids. Res. 14:5049–5065, 1986.
8. Lee, J.S., Burckholder, G.D., Latimer, L.J.P., Haug, B.L., Braun, R.P. A monoclonal antibody to triplex DNA binds to eucaryotic chromosomes. Nucl. Acids. Res. 15: 1047–1061, 1987.
9. Gibson, A.L., Herron, J.N., He, X.-M., Patrick, V.A., Mason, M.L., Lin, J.N., Kranz, D.M., Voss, E.W., Jr., Edmundson, A.B. Differences in crystal properties and ligand affinities of an antifluorescyl Fab (4-4-20) in two solvent systems. Proteins: Struct. Funct. Genet. 3:155–160, 1988.
10. Herron, J.N., He, X.-M., Mason, M.L., Voss, E.W., Jr., Edmundson, A.B. Three-dimensional structure of a fluorescein-Fab complex crystallized in 2-methyl-2,4-pentanediol. Proteins: Struct. Funct. Genet. 5:271–280, 1989.
11. Maizels, N., Bothwell, A. The T-cell-independent immune response to the hapten NP uses a large repertoire of heavy chain genes. Cell 43:715–720, 1985.
12. Maxam, A., Gilbert, W. A new method for sequencing DNA. Proc. Natl. Acad. Sci. U.S.A. 74:560–564, 1977.
13. Satow, Y., Cohen, G.H., Padlan, E.A., Davies, D.R. The phosphorylcholine binding immunoglobulin Fab McPC 603: an X-ray diffraction study at 2.7 Å. J. Mol. Biol. 190: 593–604, 1986.
14. Rossmann, M.G., Blow, D.M. The detection of subunits within the crystallographic asymmetric unit. Acta Crystallogr. 15:24–31, 1962.
15. Crowther, R.A. Fast rotation function. In: "The Molecular Replacement Method: A Collection of Papers on the Use of Non-crystallographic Symmetry." Rossmann, M.G., ed. New York: Gordon and Breach, 1972: 173–178.
16. Lattman, E.E., Love, W.E. A rotational search procedure for detecting a known molecule in a crystal. Acta Crystallogr. B26:1854–1857, 1970.
17. Fitzgerald, P.M.D. Merlot, An integrated package of computer programs for the determination of crystal structures by molecular replacement. J. Appl. Crystallogr. 21:273–278, 1988.
18. Hendrickson, W.A., Konnert, J.H. Stereochemically restrained crystallographic least-squares refinement of macromolecule structures. In: "Biomolecular Structure, Conformation, Function and Evolution." Srinivasan, R., Subramanian, E., Yathindra, N., eds. Vol. 1, New York: Pergamon Press, 1981: 43–57.
19. Hendrickson, W.A. Stereochemically restrained refinement of macromolecular structures. Methods Enzymol. 115:252–270, 1985.
20. Bhat, T.N., Cohen, G.H. An electron density map suitable for the examination of errors in a macromolecular model. J. Appl. Crystallogr. 17:244–248, 1984.
21. Jones, T.A. A graphics model building and refinement system for macromolecules. J. Appl. Crystallogr. 11:268–272, 1978.
22. Pflugrath, J.W., Saper, M.A., Quiocho, F.A. New generation graphics system for molecular modeling. In: "Methods and Applications in Crystallographic Computing." Hall, S., Ashida, T., eds. Oxford: Clarendon Press, 1984: 404–407.
23. Brünger, A.T., Kuriyan, J., Karplus, M. Crystallographic R-factor refinement by molecular dynamics. Science 235: 458–460, 1987.
24. Brünger, A.T., Karplus, M., Petsko, G.A. Crystallographic refinement by simulated annealing: application to crambin. Acta Crystallogr. A45:50–61, 1989.
25. Luzzati, P.V. Traitement statistique des erreurs dans la determination des structures cristallines. Acta Crystallogr. 5:802–810, 1952.
26. Edmundson, A.B., Ely, K.R., Abola, E.E., Schiffer, M., Panagiotopoulos, N. Rotational allomerism and divergent evolution of domains in immunoglobulin light chains. Biochemistry 14:3953–3961, 1975.
27. Lazure, C., Hum, W.T., Gibson, D.M. Sequence diversity within a subgroup of mouse immunoglobulin kappa chains controlled by the IgK-Ef2 locus. J. Exp. Med. 154:146–155, 1981.
28. Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M., Gottesman, K.S. "Sequences of Proteins of Immunological Interest." U.S. Department of Health and Human Services, Public Health Service, National Institutes of Health, Bethesda, MD, 1987.
29. Bedzyk, W.D., Johnson, L.S., Riordan, G.S., Voss, E.W., Jr. Variable region primary structure of a high affinity monoclonal anti-fluorescein antibody. J. Biol. Chem. 264:1565–1569, 1989.
30. Connolly, M.L. Solvent-accessible surfaces of proteins and nucleic acids. Science 221:709–713, 1983.
31. Camerman, N., Fawcett, J.K., Camerman, A. Molecular structure of a deoxyribose-dinucleotide, sodium thymidy-

lyl-(5′ yields to 3′)-thymidylate-(5′) hydrate (pTpT), and a possible structural model for polythymidylate. J. Mol. Biol. 107:601–621, 1976.
32. Radic, M.Z., Mascelli, M.A., Erikson, J., Shan, H., Shlomchik, M., Weigert, M. Structural patterns in anti-DNA antibodies from MRL/1pr mice. Cold Spring Harbor Symp. Quant. Biol. 54:933–946, 1989.
33. Eilat, D., Webster, D.M., Rees, A.R. V region sequences of anti-DNA and anti-RNA autoantibodies from N2B/NZW F_1 mice. J. Immunol. 141:1745–1753, 1988.
34. Seeman, N.C., Rosenberg, J.M., Rich, A. Sequence-specific recognition of double helical nucleic acids by proteins. Proc. Natl. Acad. Sci. U.S.A. 73:804–808, 1976.
35. McClarin, J.A., Frederick, C.A., Wang, B.-C., Greene, P., Boyer, H.W., Grable, J., Rosenberg, J.M. Structure of the DNA-Eco RI endonuclease recognition complex at 3 Å resolution. Science 234:1526–1541, 1986.
36. Saenger, W. "Principles of Nucleic Acid Structure." New York: Springer-Verlag, 1984: 77.
37. Colman, P.M., Laver, W.G., Varghese, J.N., Baker, A.T., Tulloch, P.A., Air, G.M., Webster, R.G. Three-dimensional structure of a complex of antibody with influenza virus neuraminidase. Nature (London) 326:358–363, 1987.
38. Colman, P.M., Tulip, W.R., Varghese, J.N., Tulloch, P.A., Baker, A.T., Laver, W.G., Air, G.M., Webster, R.G. Three-dimensional structures of influenza virus neuraminidase-antibody complexes. Phil. Trans. R. Soc. London B323: 511–518, 1989.
39. Colman, P.M. Structure of antibody-antigen complexes: Implications for immune recognition. Adv. Immunol. 43: 99–132, 1988.
40. Bhat, T.N., Bentley, G.A., Fischmann, T.O., Boulot, G., Poljak, R.J. Small rearrangements in structures of Fv & Fab fragments of antibody D1.3 on antigen binding. Nature (London) 347:483–485 (1990).
41. Stanfield, R.L., Fieser, T.M., Lerner, R.A., Wilson, I.A. Crystal structures of an antibody to a peptide and its complex with peptide antigen at 2.8 Å. Science 248:712–719, 1990.
42. Sheriff, S., Silverton, E.W., Padlan, E.A., Cohen, G.H., Smith-Gill, S.J., Finzel, B.C., Davies, D.R. Three-dimensional structure of an antibody-antigen complex. Proc. Natl. Acad. Sci. U.S.A. 84:8075–8079, 1987.
43. Padlan, E.A., Silverton, E.W., Sheriff, S., Cohen, G.H., Smith-Gill, S.J., Davies, D.R. Structure of an antibody-antigen complex: crystal structure of the HyHEL-10 Fab-lysozyme complex. Proc. Natl. Acad. Sci. U.S.A. 86:5938–5942, 1989.
44. Davies, D.R., Padlan, E.A., Sheriff, S. Antibody-antigen complexes. Annu. Rev. Biochem. 59:439–473, 1990.
45. Amit, A.G., Mariuzza, R.A., Phillips, S.E.V., Poljak, R.J. Three-dimensional structure of an antigen-antibody complex at 2.8 Å resolution. Science 233:747–753, 1986.
46. Edmundson, A.B., Harris, D.L., Schley, B.T., Tribbick, G., Geysen, H.M. Principles and pitfalls in designing site directed peptide ligands. Submitted for publication.
47. Geysen, H.M., Rodda, S.J., Mason, T.J., Tribbick, G., Schoofs, P.G. Strategies for epitope analysis using peptide synthesis. J. Immunol. Methods 102:259–274, 1987.
48. Schwartz, R.S., Stollar, B.D. Origins of anti-DNA autoantibodies. J. Clin. Invest. 75:321–327, 1985.
49. Eilat, D., Hochberg, M., Pumphrey, J., Rudikoff, S. Monoclonal antibodies to DNA and RNA from NZB/NZW F_1 mice: antigenic specificities and NH2 terminal amino acid sequences. J. Immunol. 133:489–494, 1984.
50. Diamond, B., Scharff, M.D. Somatic mutation of the T15 heavy chain gives rise to an antibody with autoantibody specificity. Proc. Natl. Acad. U.S.A. 81:5841–5844, 1984.
51. Kabat, E.A., Liso, J., Brelting, H., Franklin, E.C., Geltner, D., Frangione, B., Koshland, M.E., Shyong, J., Osserman, E.F. Human monoclonal macroglobulins with specificity for *Klebsiella* K polysaccharides that contain 3,4-pyruvylated-D-galactose and 4,6-pyruvylated-D-galactose. J. Exp. Med. 152:979–995, 1980.
52. Munns, T.W., Liszewski, M.K., Hahn, B.H. Antibody-nucleic acid complexes. Antigenic domains within nucleosides as defined by solid-phase immunoassay. Biochemistry 23:2964–2970, 1984.