Structural Modeling of Sequence Specificity by an Autoantibody against Single-Stranded DNA†

Melissa J. Bobeck,‡ David Rueda,‡,§ Nils G. Walter,*,‡ and Gary D. Glick*,‡,||

Departments of Chemistry and Biological Chemistry, University of Michigan, Ann Arbor, Michigan 48109-1055

Received January 31, 2007; Revised Manuscript Received March 31, 2007

ABSTRACT: 11F8 is a sequence-specific pathogenic anti-single-stranded (ss)DNA autoantibody isolated from a lupus prone mouse. Site-directed mutagenesis of 11F8 has shown that six binding site residues ($^{R31}V_H$, $^{W33}V_H$, $^{L97}V_H$, $^{R98}V_H$, $^{Y100}V_H$, and $^{Y32}V_L$) contribute 80% of the free energy for complex formation. Mutagenesis results along with intermolecular distances obtained from fluorescence resonance energy transfer were implemented here as restraints to model docking between 11F8 and the sequence-specific ssDNA. The model of the complex suggests that aromatic stacking and two sets of bidentate hydrogen bonds between binding site arginine residues ($^{R31}V_H$ and $^{R96}V_H$) and loop nucleotides provide the molecular basis for high affinity and specificity. In part, 11F8 utilizes the same ssDNA binding motif of $^{Y32}V_L$, $^{H91}V_L$, and an aromatic residue in the third complementarity-determining region to recognize thymine-rich sequences as do two anti-ssDNA autoantibodies crystallized in complex with thymine. $^{R31S}V_H$ is a dominant somatic mutation found in the J558 germline sequence that is implicated in 11F8 sequence specificity. A model of the mutant ^{R31S}11F8·ssDNA complex suggests that different interface contacts occur when serine replaces arginine 31 at the binding site. The modeled contacts between the ^{R31S}11F8 mutant and thymine are closely related to those observed in other anti-ssDNA binding antibodies, while we find additional contacts between 11F8 and ssDNA that involve amino acids not utilized by the other antibodies. These data-driven 11F8·ssDNA models provide testable hypotheses concerning interactions that mediate sequence specificity in 11F8 and the effects of somatic mutation on ssDNA recognition.

Protein−ssDNA interactions are central to many biological processes such as replication (*1, 2*), transcription (*3, 4*), translation (*5*), recombination (*6*), gene maintenance (*7, 8*), and gene repair (*1*). Antibodies provide a unique model system for studying protein−ssDNA interactions because their highly conserved sequence and three-dimensional structure have been systematically mutated to differentiate residues necessary for structural integrity from those involved in antigen binding (*9*). Hence, site-directed mutagenesis can be applied in studying the role of residues that mediate antigen binding without altering the overall structure of the protein.

While antibodies form the basis of the normal humeral immune response, defects in tolerance can give rise to autoantibodies, some of which are pathogenic. Anti-DNA autoantibodies are a prominent serological hallmark of the autoimmune disease systemic lupus erythematosus (SLE).[1] Through a process involving antigen recognition, a subset of anti-DNA autoantibodies localize to the glomerular basement membrane (GBM) of kidney tissue and initiate an inflammatory response that can result in renal damage and death (*10*). While anti-DNA autoantibodies that bind both double-stranded and single-stranded (ss)DNA have been extensively studied, there is currently no method for differentiating pathogenic anti-DNA autoantibodies from those that are benign or for preventing formation of a complex between these autoantibodies and their antigens (*11*−*15*).

We have previously generated a panel of anti-ssDNA monoclonal autoantibodies (mAbs) from an autoimmune mouse (*16*). Three mAbs, 11F8, 9F11, and 15B10, arose from the same B-cell precursor (i.e., they are clonally related) with >99.5% sequence conservation. However, only 11F8 binds to ssDNA adherent to the GBM and displays sequence

† Supported by NIH Grant GM 42168.

* To whom correspondence should be addressed. G.D.G.: phone, (734) 764-4548; fax, (734) 615-8902; e-mail, gglick@umich.edu. N.G.W.: e-mail, nwalter@umich.edu.

‡ Department of Chemistry.

§ Current address: Department of Chemistry, Wayne State University, Detroit, MI 48202.

|| Department of Biological Chemistry.

[1] Abbreviations: anti-ssDNA, autoantibodies that bind single-stranded DNA; $^{a\#}V_X$, amino acid variable chain residue nomenclature in which a# represents the single letter code and residue number and H or L represents the heavy or light chain, respectively; $^{R31S}V_X$, amino acid nomenclature in which R31S represents a mutation to either the heavy (H) or light (L) chain; ^{R31S}11F8, structural mutation investigated by docking simulations; CDR, complementarity-determining region; SLE, systemic lupus erythematosus; GBM, glomerular basement membrane; mAb, monoclonal antibody; **1**, 21-mer high-affinity consensus sequence for 11F8; **1**(55), full-length high-affinity ligand from selection experiments; TCEP, tris(2-carboxyethyl)phosphine; K_d, equilibrium dissociation constant; 5-TAMRA SE, 5-carboxytetramethylrhodamine succinimidyl ester; ^{F}11F8, 11F8 species labeled with fluorescein (superscript F); ^{R}**1**, **1** variant labeled with tetramethylrhodamine (superscript R); trFRET, time-resolved fluorescence resonance energy transfer; R_0, Förster distance; Fab, antigen binding fragment; HADDOCK, High Ambiguity Driven protein−protein Docking; E_{vdw}, energy resulting from van der Waals interactions; E_{elec}, energy resulting from electrostatic interactions; E_{dist}, energy resulting from unmet distance restraints; BSA, buried surface area; E_{tot}, total energy of a complex; rmsd, root-mean-square deviation; SASA, solvent accessible surface area; r^2, linear correlation coefficient.

10.1021/bi700212s CCC: $37.00
Published on Web 05/16/2007

FIGURE 1: **1**, a high-affinity ssDNA ligand of 11F8, is the truncated analogue of the consensus sequence selected from a pool of DNA oligonucleotides with a seven-nucleotide random region and lacking secondary structure (*18*). Mutations introduced during PCR amplification afforded the stable hairpin secondary structure that, together with the initial randomized region shown in bold, was selected by 11F8. Truncation of PCR primer sites on the full-length selected ligand [**1**(55)] as well as replacement of the original G_4 bulge with T_4 to form **1** does not alter the binding properties of 11F8 (*42*, *43*).

specificity for ssDNA (*17*, *18*). In vitro selection of the high-affinity ssDNA consensus sequence (**1**), in combination with DNA footprinting, affinity, and binding rate constant measurements, showed that sequence-specific recognition is achieved by 11F8 when the functional groups of **1** are presented within a well-defined secondary structure (Figure 1) (*18*). Site-directed mutagenesis revealed that six residues within the 11F8 complementarity-determining regions (CDRs) account for ~80% of the binding free energy of the complex (*19*). Two amino acids, $^{R31}V_H$ and $^{Y100}V_H$, mediate sequence specificity; however, the nature of intermolecular contacts could not be determined (*19*). $^{31}V_H$ and $^{100}V_H$ are the only two residues of 11F8 that differ from the equivalent residues in both 9F11 and 15B10, implying that contacts between these two residues and **1** should provide insight into the molecular basis of sequence specificity.

Efforts to obtain a crystal structure of the 11F8·**1** complex so far have been unsuccessful. To investigate the structural basis for ssDNA binding specificity, we here utilized a docking approach driven by experimental distance constraints from time-resolved fluorescence resonance energy transfer (trFRET) to build and optimize a model of the 11F8·**1** complex. The resulting lowest-energy ensemble of structures suggests that stacking and π-cation interactions promote the high affinity of the complex, while selectivity arises from two sets of bidentate hydrogen bonds and an aromatic residue in the center of the antigen binding site. Modeling of the mutant ^{R31S}11F8·**1** complex (reflecting a non-sequence-specific mAb sequence prevalent in the J558 germline) suggests possible differences in the recognition motif in the absence of the critical binding site arginine 31. Our models suggest that a single arginine acquired during antibody maturation affords an extended interface with ssDNA beyond the contacts accessible by other anti-ssDNA mAbs and that conformational differences and contacts facilitated by this arginine could mediate sequence specificity.

MATERIALS AND METHODS

11F8 Cysteine Variant Preparation and Validation. The His-tagged single-chain 11F8 construct cloned into pET-28b-(+) (Novagen, Madison, WI) was used as the template for 11F8 mutants (*19*). Plasmids encoding individual 11F8 cysteine mutants $^{S17C}V_H$, $^{T30C}V_H$, $^{S53C}V_H$, $^{Q105C}V_H$, $^{S10C}V_L$, $^{S43C}V_L$, $^{S60C}V_L$, $^{S67C}V_L$, and $^{T94C}V_L$ were constructed using site-directed mutagenesis following the Quickchange protocol (Stratagene, La Jolla, CA). Successful incorporation of each mutation was confirmed by DNA sequencing. Expression and purification were performed as described for the wild-type protein with the following modifications (*19*). Denatured 11F8 variants were isolated from the insoluble bacteria pellet and purified on Ni^{2+} agarose (Qiagen Inc., Santa Clarita, CA). Resultant fractions containing protein were refolded by rapid dilution (1:200) into DNA binding buffer [50 mM Na_iPO_4 (pH 8.0), 150 mM NaCl, and 0.1 mM tris(2-carboxyethyl)phosphine (TCEP)], supplemented with oxidized glutathione (0.4 mM), reduced glutathione (4 mM), and 1% (v/v) ssDNA agarose matrix, which was stirred for at least 24 h at 4 °C (*20*). The suspension was poured through a fritted glass column to concentrate the protein-bound ssDNA agarose, and the resin was washed with 5 column volumes of DNA binding buffer containing NaCl (400 mM). Protein was eluted from the ssDNA agarose with DNA binding buffer containing NaCl (2 M) and urea (2 M) and immediately exchanged into conjugation buffer [50 mM Na_iPO_4, 150 mM NaCl, and 0.1 mM TCEP (pH 8.0)] by dialysis. Protein (>98% pure by SDS−PAGE analysis) was concentrated with an Amicon Ultra YM10 membrane (Millipore, Bedford, MA) to micromolar concentrations. Purified cysteine mutants were stored at −80 °C or conjugated to fluorescein as described below.

Fluorescein Labeling. Fluorescein 5-maleimide (Molecular Probes, Eugene, OR) was dissolved (50 mg/mL) in DMSO, and a 50-fold molar excess was added to a stirring solution of concentrated folded protein (15−100 μM) in conjugation buffer. The reaction proceeded at room temperature with stirring for 2 h and then at 4 °C with stirring for an additional 18 h. Unreacted dye was removed by passing the reaction mixture over a 2.5 mL bed volume PD-10 column (Amersham, Piscataway, NJ) equilibrated with conjugation buffer. Eluted protein was dialyzed against 4 L of conjugation buffer for 24 h to remove trace amounts of unconjugated dye. The degree of labeling was calculated for each 11F8 variant by comparing the protein absorbance at 280 nm [extinction coefficient of 11F8 of 51 970 cm^{-1} M^{-1} (*21*)] with the absorbance of the fluorophore at 495 nm [extinction coefficient of fluorescein of 83 000 cm^{-1} M^{-1} (*22*)] and correcting for the UV absorbance contributed by fluorescein at 280 nm. Labeling efficiencies for 11F8 variants ranged from 60 to 90%. Site specificity of labeling was evaluated using wild-type 11F8 as a negative control. No dye conjugation to wild-type 11F8 was observed using the same protocol, confirming that the four cysteine side chains involved in disulfide bridge interactions do not react with maleimide dye. Equilibrium dissociation constants (K_d) were determined in 20 mM Tris and 150 mM NaCl (pH 8.0) at 25 °C for each cysteine mutant and corresponding fluorophore conjugate as previously described (*23*).

DNA Variant Preparation and Validation. Amine variants of **1** containing a single copy of Amino-Modifier C6 dT or C6 dC (Glen Research, Sterling VA) were synthesized using standard phosphoramidite chemistry. Oligonucleotides were synthesized by replacing cytosine or thymine at T4, C8, C9, T10, and C12 as well as the two positions flanking the 5′ and 3′ ends of the hairpin with a C5 alkylamine nucleobase, and amine-modified 21-mer hairpin variants were purified

to >95% by HPLC as previously described (*16*). Prior to conjugation, DNA was ethanol precipitated to remove Tris salt. Each modified oligonucleotide was alkylated with 5-carboxytetramethylrhodamine succinimidyl ester (5-TAMRA SE, Molecular Probes). Briefly, 100 μg of a 25 mg/mL stock of each oligonucleotide in H_2O was alkylated with 5-TAMRA (200 μg) dissolved in DMSO (14 μL) in sodium tetraborate labeling buffer (0.1 M, pH 8.5) with a total reaction volume of 100 μL as described previously (*24*). The reaction was allowed to proceed at room temperature with stirring for 8 h. Each reaction mixture was precipitated twice with EtOH, and labeled oligonucleotides were purified by reverse-phase HPLC using a Platinum EPS C_8 column (Alltech, Deerfield, IL) equilibrated at 95% (v/v) 0.1 M triethylammonium acetate (pH 6.6) and 5% (v/v) acetonitrile. DNA was eluted with a linear gradient from 5 to 65% acetonitrile over 30 min, similar to procedures described previously (*24*). Analytical reverse-phase HPLC analysis showed that each TAMRA-labeled oligonucleotide was >95% pure. Each oligonucleotide was equilibrated with NaCl (4 M) for 12 h to ensure that all counterions were replaced with Na^+, and sequences were desalted with a 2.5 mL bed volume PD-10 column (Amersham). The structural integrity of all modified hairpins was tested by their affinity for 11F8. Equilibrium dissociation constants (K_d) were determined in 20 mM Tris and 150 mM NaCl (pH 8.0) at 25 °C as previously described (*23*).

Equilibrium Fluorescence Anisotropy Measurements. The fluorescence anisotropy of each fluorescein-labeled 11F8 (^{F}11F8) and tetramethylrhodamine-labeled **1** (^{R}**1**) was determined first as a single species and then in complex with unlabeled **1** and 11F8, respectively, and finally in complex with each ^{R}**1** and ^{F}11F8 mutant, respectively. Excitation and emission bandpass filters were utilized to observe the anisotropies of fluorescein (excitation at 480 ± 10 nm and emission at 530 ± 10 nm; Andover Corp., Salem, NH) and tetramethylrhodamine (excitation at 560 ± 10 nm and emission at 590 ± 10 nm) in isolation. Titration of ^{F}11F8 and ^{R}**1** with a >10-fold excess of labeled DNA and protein, respectively, resulted in the same anisotropies that were observed with a 1:1 molar ratio so that experiments were completed with equimolar amounts of 11F8 and **1**. The anisotropy of each complex (20 nM 11F8 or mutant and 20 nM **1** or variant) was measured, after equilibrating at room temperature for 5 min, as an average of 25 scans using a Beacon 2000 Fluorescence Polarization System (Invitrogen, Carlsbad, CA). Fluorescein and tetramethylrhodamine anisotropies for all ^{F}11F8·**1**, 11F8·^{R}**1**, and ^{F}11F8·^{R}**1** complexes were ≤0.20, validating the assumption of free fluorophore rotation upon complex formation as a basis for using Förster theory in our trFRET distance analysis (*25*).

Time-Resolved Fluorescence Resonance Energy Transfer. Intermolecular distances in the various 11F8·**1** equilibrium complexes were measured by trFRET between ^{F}11F8 and ^{R}**1** in 20 mM Tris and 150 mM NaCl (pH 8.0) at 25 °C. Each complex (100 μL; 0.5 μM 11F8 species, 2.5 μM **1** species to ensure saturation of the mAb based upon equilibrium fluorescence experiments**)** was incubated at 25 °C for 5 min to ensure that the complex was formed prior to the collection of time-resolved emission profiles of the fluorescein donor using time-correlated single-photon counting as previously described (*26*). Briefly, a frequency-doubled Nd:YVO_4 laser (Spectra-Physics Millenia Xs-P, operated at 8.9 W) pumped a frequency-doubled mode-locked Ti:sapphire laser (Spectra-Physics Tsunami, operated at 1 W) that excited fluorescein at 490 nm by 2 ps wide pulses picked down to 4 MHz. Under magic angle polarizer conditions, isotropic emission was detected at 520 nm (10 nm bandpass interference filter; Chroma, Rockingham, VT) in 4096 sampling channels with a time increment of 12 ps/channel to ≥40000 peak counts. Time-resolved fluorescence decays for ^{F}11F8 mutants were measured with and without the tetramethylrhodamine acceptor to determine the donor−acceptor distance. The effect of tetramethylrhodamine on the decay of fluorescein emission in the doubly labeled complex was then used to extract a three-dimensional Gaussian distance distribution between the two fluorophores as previously described (*26*−*28*). In all cases, 11F8 data were fit well by a single-distance distribution with χ^2 values of ≤1.3, with evenly distributed residuals. To extract absolute distances, 55 Å was used as the Förster distance (R_0) of the fluorescein and tetramethylrhodamine pair (*26*), and a value of $^2/_3$ was assumed for the orientation factor as experimentally supported by the low fluorophore anisotropies (*29*).

*Structural Coordinates for 11F8 and **1***. Structural coordinates for the 11F8 antigen binding fragment (Fab) were obtained from a homology model constructed by the methods of Chothia and Lesk (*9*, *19*, *30*). The general computational approach has been described elsewhere (*31*). A model of **1** was generated using ERNA-3D (*32*, *33*). Previous spectroscopic and footprinting experiments revealed that the selected ligand for 11F8 is a hairpin with a four-base loop, separated from a bulge by 3 bp as shown in Figure 1 (*18*). The primary sequence of **1** together with the base pair designations of Figure 1 was used as the starting point for modeling. Sodium counterions and a 10 Å water solvation layer were added to the resultant hairpin with the Builder module of Insight II (Accelrys Inc.), and the structure of **1** was energy minimized to a root-mean-square deviation (rmsd) of 0.01 Å.

*Docking of the 11F8·**1** Complex*. High-ambiguity-driven protein−protein docking (HADDOCK 1.3) was utilized to model docking of 11F8 and **1** (*34*). Two types of experimental data were implemented to guide docking in the following way. trFRET distances served as unambiguous restraints (discrete distance restraints), and important binding site residues identified by site-directed mutagenesis for both 11F8 and **1** were incorporated as ambiguous interacting restraints (known interacting residues but unknown contacts). All residues used as ambiguous restraints were >45% solvent accessible as determined with the Solvation module of Insight and either exhibited evidence of the direct involvement in the complex (active) or were solvent exposed neighbors to such residues (passive). $^{R31}V_H$, $^{L97}V_H$, $^{R98}V_H$, $^{Y100}V_H$, and $^{Y32}V_L$ were chosen as active 11F8 residues (*19*). G7, C8, C9, T10, T11, and C12 were selected as active **1** residues (*18*, *23*). trFRET-determined distances for donor−acceptor pairs in the equilibrium complexes were incorporated as unambiguous restraints in a NOE-formatted restraint file both independently and in conjunction with ambiguous mutagenesis-derived restraints. Stacking restraints (4 ± 2 Å) between binding site aromatic residues and thymine bases were incorporated to refine each model (*35*, *36*). During the final simulation, distance, stacking, and mutagenesis restraints

were considered, and resultant structures were energetically penalized (E_{dist}) for unsatisfied restraints.

Each docking procedure consisted of three steps. First, 11F8 and **1** were rotated around their center of mass to minimize intermolecular energy, and molecules were docked by rigid body energy minimization. More than 1000 structures were sorted according to the sum of van der Waals energy (E_{vdw}), electrostatic energy (E_{elec}), distance restraint energy (E_{dist}), and buried surface area (BSA). At least 200 lowest-energy solutions were refined in the second stage where they were subjected to three stages of semirigid simulated annealing (1000 steps, 2000 to 50 K, 8 fs time steps; 4000 steps, 2000 to 50 K, 4 fs time steps; and 1000 steps, 500 to 50 K, 2 fs time steps) followed by water refinement with an 8 Å shell of TIP3P water molecules to optimize the interface. Throughout semirigid body docking and simulated annealing, all 11F8 CDRs ($^{31-35}$HCDR1, $^{50-66}$HCDR2, $^{99-108}$HCDR3, $^{24-34}$LCDR1, $^{50-56}$LCDR2, and $^{89-97}$LCDR3) and all of **1** were allowed full backbone and side chain flexibility to most effectively represent native docking. Allowing side chain and backbone flexibility during simulated annealing was found to optimize the orientation and interface of 11F8 and **1**. The stereochemical quality of the resultant models was analyzed with PROCHECK, and interface hydrogen bond and nonbonding interactions were assessed by HADDOCK using standard CNS analysis scripts (*37*). Complexes were ranked on the basis of total energy (E_{tot}).

*Docking of the ^{R31S}11F8·**1** Complex*. Docking was used to model the effects of ^{S31R}11F8 somatic mutation upon ssDNA recognition. $^{R31}V_H$ was mutated from arginine to serine in the 11F8 homology model and docked with **1** as described for wild-type 11F8. Because $^{W33}V_H$ does not participate in ssDNA recognition in other anti-ssDNA autoantiboides, stacking restraints for $^{W33}V_H$ were not included.

Steady-State Fluorescence Resonance Energy Transfer (FRET). The orientation of 11F8 with respect to **1** was cross-validated by steady-state FRET. In all experiments, 11F8 (50 nM) (either unlabeled or fluorescein-labeled cysteine mutant) was equilibrated at 25 °C with **1** (250 nM) (either unlabeled or tetramethylrhodamine-labeled amine variant) to ensure that >98% of 11F8 was bound to **1**. For each measurement, binding buffer [20 mM Tris-HCl, 150 mM NaCl, and 0.1 mM TCEP (pH 8.0)] in a final volume of 1300 μL was utilized. Three samples were analyzed for each of the pairs of labeling sites (nine mAb mutants and six variants of **1** = 54 pairs): a sample with donor only (F_D, 50 nM ^{F}11F8 with 250 nM **1**), a sample with a donor in the presence of a saturating amount of acceptor (F_{DA}, 50 nM ^{F}11F8 with 250 nM ^{R}**1**), and a sample with acceptor alone (F_A, 50 nM 11F8 cysteine mutant with 250 nM ^{R}**1**). In all cases, the fluorescence spectra of the three samples were measured with a λ_{ex} of 496 nm, a 2 nm bandpass, at 1 nm intervals from 500 to 700 nm to observe both the fluorescein and tetramethylrhodamine fluorescence. The fluorescence spectrum of the acceptor alone was subtracted from the spectrum of the donor in the presence of a saturating amount of acceptor to correct for background tetramethylrhodamine fluorescence. The donor only spectrum was compared to the corrected spectrum of the double-labeled material to extract the quench in fluorescein fluorescence. The emission intensity at the wavelength of maximal donor emission (525 nm) was used to calculate the efficiency of energy transfer for 54 donor−acceptor complexes according to eq 1 (*38*−*41*). An average of three emission traces was used for each experiment, and three independent assays were averaged to calculate the energy transfer efficiency (E_T) for all pairs.

$$E_T = 1 - F_{DA}/F_D \quad (1)$$

Table 1: Equilibrium Affinity of 11F8 Fluorescein-Labeled Cysteine Mutants for **1**[a]

protein	K_d with **1**(55) (nM)	relative K_d
11F8	5.6 ± 2.4	1.0
$^{S17CF}V_H$	2.9 ± 0.3	0.5
$^{T30CF}V_H$	7.0 ± 0.5	1.3
$^{S53CF}V_H$	5.5 ± 1.6	1.0
$^{Q105CF}V_H$	7.2 ± 1.2	1.3
$^{S10CF}V_L$	2.4 ± 0.7	0.4
$^{S43CF}V_L$	2.8 ± 0.8	0.5
$^{S60CF}V_L$	7.5 ± 1.7	1.3
$^{S67CF}V_L$	6.4 ± 2.4	1.1
$^{T94CF}V_L$	6.5 ± 0.6	1.2

[a] Relative K_d values are normalized against K_d of 11F8 for **1**. Errors in K_d are standard deviations based on a minimum of three independent assays.

RESULTS

*Accessing Fluorescent Variants of 11F8 and **1***. Previous characterization of the 11F8·**1** complex was insufficient for determining the orientation of 11F8 with respect to **1** or for describing interface interactions, prompting us to model the sequence-specific interaction (*18*, *19*, *23*, *42*, *43*). Discrete distance restraints were obtained from trFRET and used along with mutagenesis restraints to guide complex docking, while steady-state FRET experiments were utilized to validate the model. The FRET studies necessitated site-specific covalent attachment of fluorescent reporters to both 11F8 and **1**. Specific labeling required mutation of nonconserved amino acid residues to cysteine for 11F8 and alkylamine modification of nucleobases for **1**. Modification and fluorophore conjugation must not disrupt the 11F8·**1** complex, and all conjugated fluorophores need to be solvent-exposed and freely rotating (*44*).

Antibody variable regions have four cysteine residues that form two disulfide bonds required for proper folding. A fifth solvent-exposed cysteine for conjugating fluorescein to the folded protein had to be introduced into 11F8 by mutation. For mutation, amino acid side chains were selected, first, on the basis of a relative solvent accessible surface area (SASA) of >50%. 11F8 has a low-nanomolar affinity for **1**, which is ~7-fold tighter than the affinities observed for 9F11·**1** and 15B10·**1** complexes (*45*). All three mAbs have significantly lower affinities for other ssDNA ligands [an ~50-fold lower affinity is observed for thymine-rich ssDNA, and an ~250-fold lower affinity is observed for nonspecific ssDNA ligands (*23*, *45*)]. A binding affinity within 2-fold of that of the wild-type complex was therefore used as an indication that a particular substitution did not alter binding affinity (*19*, *23*). On the basis of these criteria, nine amino acids throughout the 11F8 variable region could be mutated to cysteine and further conjugated to fluorescein: $^{S17C}V_H$, $^{T30C}V_H$, $^{S53C}V_H$, $^{Q105C}V_H$, $^{S10C}V_L$, $^{S43C}V_L$, $^{S60C}V_L$, $^{S67C}V_L$, and $^{T94C}V_L$ (Table 1).

Table 2: Equilibrium Affinity of Tetramethylrhodamine-Labeled **1** Variants for 11F8[a]

ssDNA	K_d with 11F8 (nM)	relative K_d
1(55)	5.6 ± 2.4	1.0
^{T4R}**1**(55)	11.2 ± 0.3	2.0
^{C8R}**1**(55)	11.3 ± 3.9	2.0
^{C9R}**1**(55)	6.1 ± 0.2	1.1
^{T10R}**1**(55)	5.8 ± 1.9	1.0
^{C12R}**1**(55)	2.5 ± 0.8	0.4
1	1.5 ± 0.2	1.0
$^{T5'R}$**1**	2.8 ± 0.6	1.9
$^{T3'R}$**1**	2.8 ± 1.1	1.9

[a] Relative K_d values are normalized against K_d for 11F8. Errors in K_d are standard deviations based on a minimum of three independent assays. All binding assays except those with the 3′- and 5′-labeled hairpins were completed with 55-mer DNA sequences that bind 11F8 via the same mechanism as the 21-mer (*42*).

Seven 21-base oligonucleotides were synthesized by replacing cytosine or thymine at T4, C8, C9, T10, and C12 as well as the two positions flanking the 5′ and 3′ ends of the hairpin with a C5 alkylamine-modified nucleobase so the FRET acceptor could walk around **1**. Each site afforded tetramethylrhodamine-labeled amine variants with an affinity for 11F8 similar to that of wild-type **1** (Table 2). The various fluorescein and tetramethylrhodamine locations are shown on our 11F8 and **1** models in Figure 2. Initial binding studies were completed with the 55-mer sequence [**1** (*55*)] that was identified with binding site selection experiments (*18*). When the majority of binding affinities had been determined, we decided that a shortened 21-mer hairpin sequence flanked with two thymine residues on each side rather than the original PCR primers would prohibit interactions between 11F8 and the original PCR primers during FRET experiments. Tables 1 and 2 show mutants whose binding affinity was determined before we decided to use the 21-mer of **1**; however, since we have shown that binding properties for the full-length and shortened hairpin sequences are similar (*42*), relative binding affinities with respect to the wild-type complex were not re-determined. All FRET and anisotropy studies were completed with the 21-mer analogues of **1**. Fluorescein and tetramethylrhodamine anisotropies for all ^{F}11F8·**1**, 11F8·^{R}**1**, and ^{F}11F8·^{R}**1** complexes were ≤0.20, validating the use of Förster theory for E_T and distance calculations (*25*) (Supporting Information).

Time-Resolved FRET-Determined Distances in the 11F8·1 Complex. Analysis of the fluorescein emission decay in the absence (^{F}11F8·**1**) and presence (^{F}11F8·^{R}**1**) of a tetramethylrhodamine acceptor yields a Gaussian donor−acceptor distance distribution at equilibrium. Each measurement with 11F8 and **1** resulted in a single distance distribution, which suggests that 11F8 binds **1** in a unique and specific orientation. A labeled amino acid residue was chosen from similar locations in the heavy and light chains (^{S17}V$_H$ and ^{S10}V$_L$) for measurement of distances to five common locations (T5′, C8, C9, T10, and C12) on **1**, chosen to include both terminal and loop nucleotides. Each intermolecular distance and accompanying parameters describing the fitted Gaussian distribution are listed in Table 3.

Modeling the 11F8·1 Complex. Data-guided docking was pursued to determine the orientation of **1** with respect to 11F8 and to describe the intermolecular interface in the sequence-specific complex. The HADDOCK docking program provided a flexible platform for inputting experimental structural restraints while allowing backbone and side chain flexibility during the entire docking process. Approximately five nucleotides of **1** are involved in the complex with 11F8 (*18*). Previous studies suggest that C8, C9, T10, and T11 as well as the G7·C12 loop closing base pair might contact 11F8 (*18*, *23*). Six 11F8 amino acids, ^{R31}V$_H$, ^{W33}V$_H$, ^{L97}V$_H$, ^{R98}V$_H$, ^{Y100}V$_H$, and ^{Y32}V$_L$, contribute 80% of the binding free energy during complex formation, do not exhibit cooperativity, and are likely involved in the complex (*19*). All 11F8 residues except ^{W33}V$_H$ (33% SASA) together with nucleotides 7−12 of **1** have a SASA of >45% and were designated as active mutagenesis restraints (*34*). Solvent accessible neighbors within two amino acids of active residues of 11F8 and the remaining nucleotides in **1** were set as passive residues. Active mutagenesis restraints are satisfied when any atom of an active residue in one molecule is within 3.0 ± 2 Å of any active or passive residue in the partner molecule, thus ensuring that the binding site includes critical residues. Therefore, 11F8 mutagenesis restraints are satisfied if a single contact occurs between each active residue and any hairpin nucleotide, while mutagenesis restraints for **1** are satisfied if a contact occurs between the loop nucleotides and a 11F8 binding site residue.

Characterization of many autoantibody·antigen complexes, including anti-DNA autoantibodies and others, indicates that hydrophobic interactions are the primary driving force for recognition (*35*, *36*, *46*−*48*). Our data suggest that 11F8 also recognizes ssDNA in part through interactions between thymine residues and aromatic side chains in the binding pocket (^{Y32}V$_L$, ^{Y100}V$_H$, and ^{W33}V$_H$) (*19*, *43*). Other anti-ssDNAs utilize binding site aromatic residues to bind two thymine nucleotides (*35*, *36*). Given the three binding site aromatic residues and two loop thymine residues involved in 11F8·**1** binding, two orientations and various registers are possible for the complex (Figure 3). Initial docking experiments driven by only mutagenesis restraints resulted in a diverse array of complexes docked in both orientations and sampling different registers. Next, intermolecular distances measured by trFRET were incorporated as the sole restraints to guide docking of 11F8 and **1**. Interestingly, 11F8·**1** complexes docked on the basis of trFRET distances as well as trFRET distances in conjunction with mutagenesis restraints all exclusively suggested a 3′ to 5′ orientation. This nomenclature refers to the ssDNA orientation where the 3′-most thymine residue (T11) is associated with the 11F8 light chain and the 5′-most thymine residue (T10) is associated with the 11F8 heavy chain. All resultant complexes suggested a single binding site with T11 stacked between ^{Y32}V$_L$ and ^{Y100}V$_H$ and T10 stacked between ^{Y100}V$_H$ and ^{W33}V$_H$. Thus, to further refine the 11F8·**1** complex, stacking restraints (4 ± 2 Å) based upon related anti-ssDNA crystal structures were incorporated between aromatic residues and thymine nucleotides such that **1** was oriented in the 3′ to 5′ orientation. The lowest-energy ensemble resulting from trFRET, mutagenesis, and stacking restraints contained 23 complexes. The 10 lowest-energy structures are overlaid in Figure 4A and suggest that 11F8 utilizes hydrophobic stacking, π-cation, and sequence-specific hydrogen bond interactions to recognize **1**.

Further Support for the 3′ to 5′ Orientation. Equilibrium trFRET measurements support the 3′ to 5′ orientation of **1**.

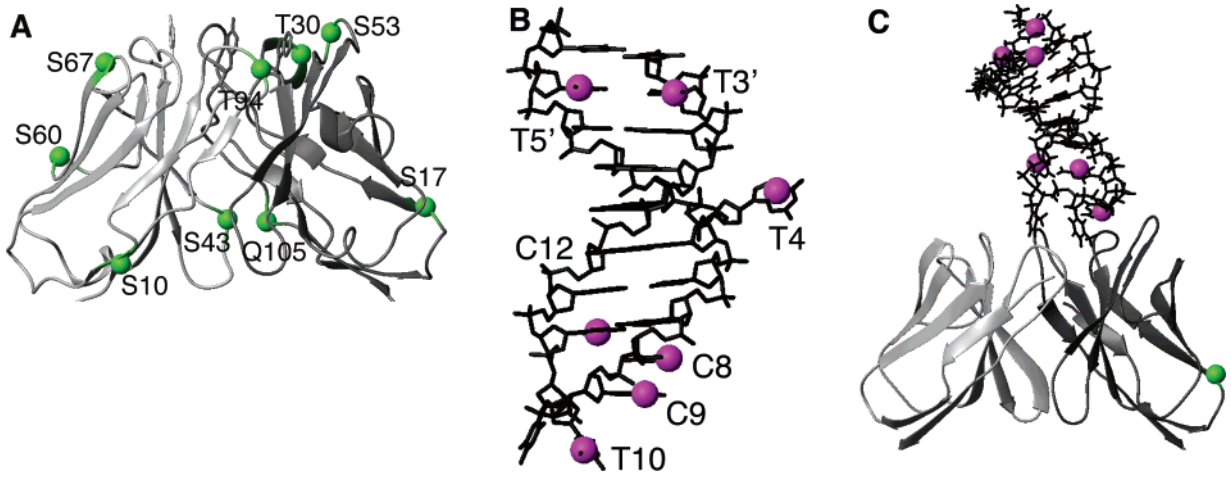

FIGURE 2: Ribbon diagrams of 11F8 and **1** illustrating the location of the fluorescent probe. The light chain of 11F8 is colored light gray, the heavy chain dark gray, and the DNA black. (A) Positions for ^{F}11F8 are shown as green spheres labeled with the amino acid number. (B) Positions for ^{R}**1** are shown as pink spheres labeled with the nucleotide number. (C) 11F8·**1** complex illustrating a single series of six steady-state FRET measurements from a single ^{F}11F8 ($^{S17C}V_H$) to each ^{R}**1**.

Table 3: trFRET-Based Distance Distributions between $^{S17CF}V_H$ and $^{S10CF}V_L$ of 11F8 in Complex with $^{T5'R}$**1**, ^{C8R}**1**, ^{C9R}**1**, ^{T10R}**1**, and ^{C12R}**1**[a]

	mean distance (Å)/fwhm (Å)/χ^2				
donor	$^{T5'R}$**1**	^{C8R}**1**	^{C9R}**1**	^{T10R}**1**	^{C12R}**1**
$^{S17CF}V_H$	56.6 ± 0.2/17.6 ± 0.4/1.1	45.6 ± 0.5/22.4 ± 1.3/1.2	41.4 ± 0.1/23.0 ± 0.4/1.2	41.1 ± 0.1/26.5 ± 0.2/1.2	38.8 ± 0.4/32.5 ± 1.0/1.1
$^{S10CF}V_L$	57.0 ± 0.2/16.1 ± 0.3/1.3	41.2 ± 0.1/23.9 ± 0.7/1.2	42.3 ± 0.1/24.0 ± 0.2/1.2	43.7 ± 0.2/22.5 ± 0.2/1.2	45.3 ± 0.5/25.3 ± 1.0/1.2

[a] Mean donor−acceptor distances and their full widths at half-maximum (fwhm) were measured in 20 mM Tris-HCl (pH 8.0) and 0.1 mM TCEP at 25 °C. Errors are standard deviations from at least two independent measurements. χ^2 is the reduced χ^2 value as a measure of fit quality.

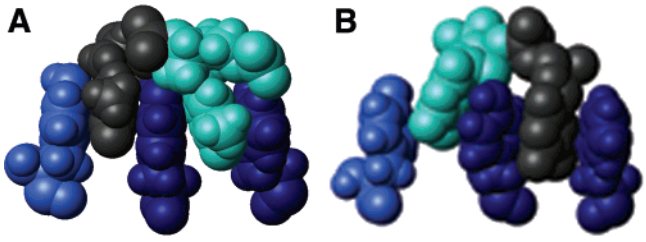

FIGURE 3: Possible orientations of the 11F8·**1** complex. The light chain residue ($^{Y32}V_L$) is colored light blue, and heavy chain residues ($^{Y100}V_H$-middle and $^{W33}V_H$-right) are colored navy. The loop thymine residue closest to the 5′ end of **1** (T10) is colored teal, and the loop thymine residue closest to the 3′ end (T11) is colored gray. (A) Schematic illustrating loop thymine residues stacked 3′ to 5′ with respect to 11F8 V_L on the left. (B) Schematic illustrating loop thymine residues stacked 5′ to 3′ with respect to 11F8 V_L on the left. In theory, there are also additional registers possible in which the thymine aromatic stacking is shifted to the left or right in either orientation.

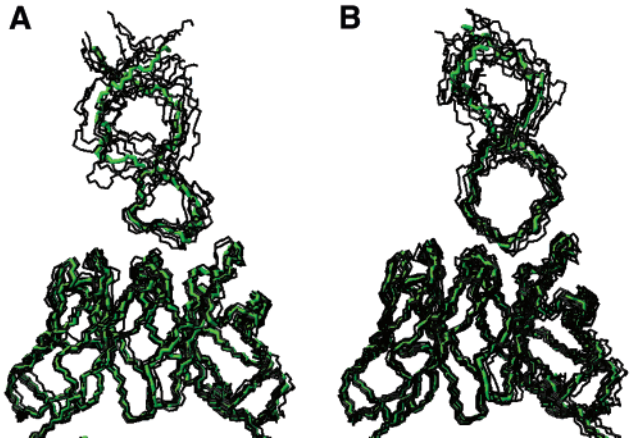

FIGURE 4: Models of the 11F8·**1** and ^{R31S}11F8·**1** complexes. (A) Overlaid backbone atoms for the 10 lowest-energy 11F8·**1** complexes (lowest-energy structure colored green). (B) Overlaid backbone atoms for the 10 lowest-energy ^{R31S}11F8·**1** complexes.

To further validate the likelihood for a 3′ to 5′ versus 5′ to 3′ oriented 11F8·**1** complex, two new docking simulations were completed in which complexes were docked with the opposite orientation using equal but opposite stacking restraints. These complexes were generated using only stacking restraints. Specifically, aromatic carbon atoms of tyrosine side chains $^{Y32}V_L$ and $^{Y100}V_H$ were restrained to 4 ± 2 Å with respect to thymine aromatic carbon and nitrogen atoms of T10 and T11 of **1**. A 4 Å π-stacking distance was chosen on the basis of stacking interactions observed in the two published anti-ssDNA·T crystal structures (*35*, *36*). Restraints between $^{Y32}V_L$ and ^{T11}**1**, ^{T11}**1** and $^{Y100}V_H$, and $^{Y100}V_H$ and ^{T10}**1** were used for a bias toward the 3′ to 5′ orientation, while restraints between $^{Y32}V_L$ and ^{T10}**1**, ^{T10}**1** and $^{Y100}V_H$, and $^{Y100}V_H$ and ^{T11}**1** were used for a bias toward the 5′ to 3′ orientation.

The lowest-energy resultant complex in each orientation was compared to equilibrium FRET measurements. Intermolecular distances were measured from the γ-carbon of each of the nine cysteine donors to the C5 position of six nucleotide base acceptors. Due to the size of 11F8 and the location of the DNA binding site at the center of the molecule, 39% of the distances differed by <5 Å between the oppositely oriented complexes, and only 33% of the distances differed by >10 Å. Because the differences between the two orientations are subtle, energy transfer efficiencies for transfer from each donor position to six of

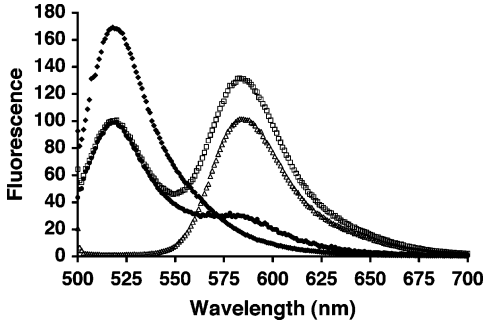

FIGURE 5: Representative fluorescence spectra for the ^{S17CF}11F8·^{C8R}**1** complex. The fluorescence spectra for acceptor only [11F8·^{C8R}**1** (△)], donor in the presence of a saturating amount of acceptor [^{S17CF}11F8·^{C8R}**1** (□)], and donor only [^{S17CF}11F8·**1** (◆)] are shown. The donor-only spectrum was compared to the corrected spectrum of the double-labeled material (●) to extract the quench in fluorescein fluorescence and calculate the FRET efficiency. The transfer efficiency, E_T, was calculated for 54 donor−acceptor pairs according to eq 1.

the seven acceptor molecules were measured by steady-state FRET for cross-validation (Figure 2C).

The energy transfer efficiency (E_T) between fluorescein and tetramethylrhodamine in solution is distance-dependent, where E_T in steady-state FRET is related to the mean distance between probes through the relationship $1/R^6$ as shown in eq 2.

$$E = R_o^6/(R_o^6 + R^6) \quad (2)$$

When $1/E_T$ as determined from steady-state FRET is plotted against R^6 derived from the 3′ to 5′ and 5′ to 3′ models of the 11F8·**1** complex, a better correlation should be observed for the model that is docked in the correct orientation. The energy transfer efficiency for each donor−acceptor pair was determined from the difference in the background-corrected fluorescence emission spectra of complexes formed in the absence (^{F}11F8·**1**) and presence (^{F}11F8·^{R}**1**) of acceptor as shown in Figure 5. Fifty-four transfer efficiencies and corresponding model distances from the two models were used to compare the 3′ to 5′ and 5′ to 3′ orientations (Table 4). Unlike trFRET, steady-state FRET reports only a single mean distance and is sensitive to differences in labeling efficiencies (ranging from 60 to 90%) that are caused by dissimilar thiol reactivities. We therefore compared energy transfer efficiencies among the same cysteine labeling sites and plotted the observed inverse FRET efficiencies for each set of donor−acceptor pairs against the intermolecular distances in the 3′ to 5′ and then 5′ to 3′ models to the sixth power. Representative plots for the four 11F8 V_H mutants are shown in Figure 6. Linear correlation coefficients for each donor were determined for all data points and are listed in Table 5. Seven of the nine series exhibit a stronger correlation with the 3′ to 5′ orientation, and overall differences in the correlation between energy transfer and model distances for the two data sets are statistically significant ($p < 0.04$). Collectively, these experiments demonstrate that FRET-based distance measurements can be used to distinguish between two structurally relatively similar models.

*Effects of $^{R31}V_H$ Residue Identity on the 11F8−**1** Interface.* Two binding site residues that have been shown to mediate sequence specificity in 11F8, $^{R31}V_H$ and $^{Y100}V_H$, are acquired during antibody maturation (*49*). The related non-sequence-specific mAbs 9F11 and 15B10 contain $^{S31}V_H$ and $^{F100}V_H$ instead. We have shown that arginine at V_H residue 31 and aromaticity at V_H residue 100 are requirements for specificity (*19*). Docking simulations were pursued also in the presence of the $^{R31S}V_H$ mutation to investigate how a lack of R31 may mediate a loss of sequence specificity. First, docking was modeled between ^{R31S}11F8 and **1** with loose stacking restraints (6 ± 2 Å) around $^{Y100}V_H$ to allow **1** to bind in either orientation. In 70 of 100 lowest-energy ^{R31S}11F8·**1** complexes, **1** was oriented 3′ to 5′ with the same register ($^{Y32}V_L$-^{T11}**1**-$^{Y100}V_H$-^{T10}**1**) as the wild-type 11F8·**1** model. These results suggest a preference for the 3′ to 5′ orientation is not mediated by $^{R31}V_H$ alone, but rather the asymmetrical binding site topology resulting from the contributions of multiple side chains. Second, the model of the ^{R31S}11F8·**1** complex was refined with 4 ± 2 Å stacking restraints to further favor the 3′ to 5′ orientation. The lowest-energy ensemble contained 27 structures, 10 of which are overlaid in Figure 4B.

*Assessment of the Lowest-Energy 11F8·**1** and ^{R31S}11F8·**1** Models.* The quality of the lowest-energy modeled 11F8·**1** and ^{R31S}11F8·**1** structures was determined in two ways. Pairwise backbone rmsds were determined between the α-carbon coordinates for each protein and nucleic acid residue in the lowest-energy 11F8·**1** or ^{R31S}11F8·**1** complex and the additional nine lowest-energy structures in each ensemble, respectively, to ensure that the lowest-energy structure was representative of a "typical" complex. The 10 lowest-energy 11F8·**1** complexes had an average energy of −420 ± 38 kcal/mol and an average backbone rmsd of 1.6 ± 0.3 Å. The ^{R31S}11F8·**1** complexes had an average energy of −410 ± 34 kcal/mol and an average backbone rmsd of 1.6 ± 0.4 Å. In both cases, visual inspection indicated that experimental restraints were sufficient to repeatedly converge on a single binding site, and observed contacts were representative of the ensemble. Inspection of the per-residue rmsd values showed that the backbone deviations were small for protein side chains and relatively high (>5 Å) for ssDNA residues not directly involved in the complex. This observation indicated the flexibility afforded to ssDNA and mAb CDRs during docking. The lowest-energy 11F8·**1** and ^{R31S}-11F8·**1** models were analyzed with PROCHECK to verify that the complexes contained reasonable backbone and side chain conformations. The benchmarks for structural quality of both models are listed in Table 6.

*Features of the 11F8·**1** Model*. Comparison of our 11F8·**1** model with the crystal structures of anti-DNA BV04-01 and DNA-1 bound to thymine sequences revealed that thymine nucleotides bind similarly at the N-termini of the V_H and V_L subunits with the 3′ end associated with the light chain (*35*, *36*). As observed in these structures, our model suggests that the bases of **1** are buried within the protein and phosphate groups are relatively solvent exposed. The predominant ssDNA-binding motif of $^{Y32}V_L$, a hydrogen-bonding residue at $^{91}V_L$, and an aromatic residue at the tip of the third CDR in the heavy chain participate in the 11F8 complex where loop thymine bases form an intimate parallel stack with binding site aromatic side chains (Figure 7). Specifically, ^{T11}**1** inserts between $^{Y32}V_L$ and $^{Y100}V_H$ and ^{T10}**1** stacks between $^{Y100}V_H$ and $^{W33}V_H$ to form several <4 Å contacts. Additionally, our model revealed contacts on the right side of the binding pocket between nonpolar ($^{L97}V_H$)

Table 4: Steady-State Transfer Efficiencies Compared to Model Distances for Each ^{F}11F8·^{R}**1** Complex[a]

complex	E_T	3′ to 5′ distance (Å)	5′ to 3′ distance (Å)	complex	E_T	3′ to 5′ distance (Å)	5′ to 3′ distance (Å)
^{S17C}11F8·$^{T3'}$**1**	0.34 ± 0.07	55.3	53.8	^{S10C}11F8·$^{T3'}$**1**	0.47 ± 0.01	51.0	53.0
^{S17C}11F8·^{T4}**1**	0.61 ± 0.05	41.5	37.6	^{S10C}11F8·^{T4}**1**	0.61 ± 0.01	31.8	43.0
^{S17C}11F8·^{C8}**1**	0.63 ± 0.02	37.2	37.6	^{S10C}11F8·^{C8}**1**	0.58 ± 0.02	41.5	36.4
^{S17C}11F8·^{C9}**1**	0.69 ± 0.03	28.2	39.9	^{S10C}11F8·^{C9}**1**	0.57 ± 0.03	40.9	34.2
^{S17C}11F8·^{C12}**1**	0.61 ± 0.03	40.8	33.7	^{S10C}11F8·^{C12}**1**	0.66 ± 0.04	37.0	39.1
^{S17CF}11F8·$^{T5'}$**1**	0.39 ± 0.03	57.1	40.1	^{S10CF}11F8·$^{T5'}$**1**	0.43 ± 0.04	53.2	53.0
^{T30C}11F8·$^{T3'}$**1**	0.38 ± 0.03	30.3	27.5	^{S43C}11F8·$^{T3'}$**1**	0.56 ± 0.02	38.3	42.3
^{T30C}11F8·^{T4}**1**	0.45 ± 0.00	25.5	37.9	^{S43C}11F8·^{T4}**1**	0.64 ± 0.01	18.0	40.0
^{T30C}11F8·^{C8}**1**	0.58 ± 0.02	13.7	25.8	^{S43C}11F8·^{C8}**1**	0.58 ± 0.01	26.0	30.7
^{T30C}11F8·^{C9}**1**	0.63 ± 0.03	8.1	26.5	^{S43C}11F8·^{C9}**1**	0.55 ± 0.03	24.2	29.8
^{T30C}11F8·^{C12}**1**	0.66 ± 0.01	19.8	18.2	^{S43C}11F8·^{C12}**1**	0.71 ± 0.03	20.2	28.6
^{T30CF}11F8·$^{T5'}$**1**	0.45 ± 0.02	31.9	34.4	^{S43CF}11F8·$^{T5'}$**1**	0.54 ± 0.04	39.5	45.2
^{S53C}11F8·$^{T3'}$**1**	0.50 ± 0.03	32.0	24.5	^{S60C}11F8·$^{T3'}$**1**	0.41 ± 0.06	31.3	52.4
^{S53C}11F8·^{T4}**1**	0.63 ± 0.01	30.2	34.8	^{S60C}11F8·^{T4}**1**	0.75 ± 0.01	13.1	50.9
^{S53C}11F8·^{C8}**1**	0.59 ± 0.01	13.5	24.4	^{S60C}11F8·^{C8}**1**	0.59 ± 0.01	30.8	34.7
^{S53C}11F8·^{C9}**1**	0.64 ± 0.02	7.0	26.1	^{S60C}11F8·^{C9}**1**	0.58 ± 0.04	35.9	31.0
^{S53C}11F8·^{C12}**1**	0.62 ± 0.05	20.8	16.5	^{S60C}11F8·^{C12}**1**	0.69 ± 0.3	23.6	34.7
^{S53CF}11F8·$^{T5'}$**1**	0.42 ± 0.04	34.9	30.6	^{S60CF}11F8·$^{T5'}$**1**	0.48 ± 0.02	31.8	54.6
^{Q105C}11F8·$^{T3'}$**1**	0.33 ± 0.05	43.1	42.1	^{S67C}11F8·$^{T3'}$**1**	0.30 ± 0.07	32.2	38.6
^{Q105C}11F8·^{T4}**1**	0.66 ± 0.02	23.4	48.9	^{S67C}11F8·^{T4}**1**	0.42 ± 0.04	27.7	24.1
^{Q105C}11F8·^{C8}**1**	0.40 ± 0.08	29.6	31.8	^{S67C}11F8·^{C8}**1**	0.56 ± 0.02	28.0	18.3
^{Q105C}11F8·^{C9}**1**	0.41 ± 0.05	25.6	31.5	^{S67C}11F8·^{C9}**1**	0.19 ± 0.02	32.5	14.2
^{Q105C}11F8·^{C12}**1**	0.45 ± 0.01	29.9	27.5	^{S67C}11F8·^{C12}**1**	0.47 ± 0.03	21.5	22.9
^{Q105C}11F8·$^{T5'}$**1**	0.33 ± 0.01	43.4	43.3	^{S67CF}11F8·$^{T5'}$**1**	0.35 ± 0.03	36.7	38.6
				^{T94C}11F8·$^{T3'}$**1**	0.27 ± 0.03	36.2	25.7
				^{T94C}11F8·^{T4}**1**	0.31 ± 0.07	28.7	20.9
				^{T94C}11F8·^{C8}**1**	0.44 ± 0.01	19.9	13.1
				^{T94C}11F8·^{C9}**1**	0.45 ± 0.05	16.4	14.4
				^{T94C}11F8·^{C12}**1**	0.37 ± 0.01	20.3	14.2
				^{T94CF}11F8·$^{T5'}$**1**	0.25 ± 0.03	40.6	26.6

[a] The transfer efficiency (E_T) between ^{F}11F8 and ^{R}**1** compared with the respective intermolecular distances measured from the Cγ atom of each cysteine mutant to C5 of each nucleotide base in the respective 3′ to 5′ or 5′ to 3′ oriented 11F8·**1** complex. Reported errors for FRET efficiency are the standard deviations of at least three independent assays.

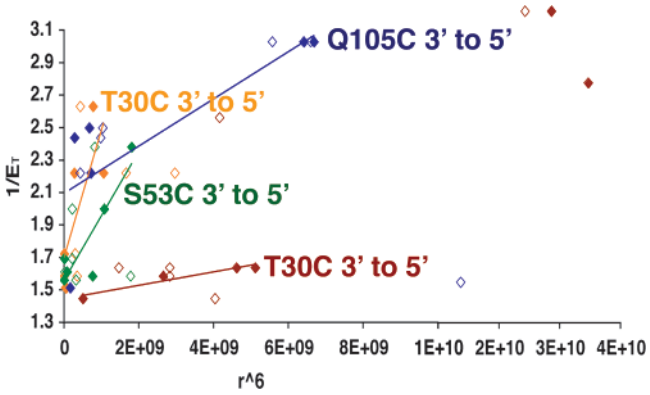

FIGURE 6: Cross-validation of the 11F8·**1** model. Representative plot of the inverse of FRET efficiency for each heavy chain donor series vs the intermolecular distances in the 3′ to 5′ (◆) and 5′ to 3′ (◇) models to the sixth power. Because the relationship of 1/*E* vs r^6 is linear, the orientation with the best linear correlation is supported.

and basic (^{R31}V$_H$, ^{R96}V$_H$, and ^{R98}V$_H$) side chains that are not utilized by other autoantibodies to interact with ssDNA. The guanidinium group of ^{R98}V$_H$ contacts the pyrimidine ring of cytosine ^{C12}**1**, creating a π-cation interaction in which the cationic side chain likely stabilizes the bound ssDNA (Figure 8). Two bidentate hydrogen bond interactions are also observed at the binding site of the 11F8·**1** model. ^{R31}V$_H$ forms a bidentate hydrogen bond between the guanidinium group and O2 and N3 of loop cytosine ^{C9}**1**, and ^{R96}V$_H$ forms a bidentate hydrogen bond with the G7·C12 loop closing base pair of **1** between the guanidinium group and O6 and N7 of ^{G7}**1** (Figure 8).

Table 5: Strength of the Correlation between Equilibrium Data and Intermolecular Distance in the Lowest-Energy 11F8·**1** Complexes in the 3′ to 5′ vs 5′ to 3′ Orientation[a]

donor series	3′ to 5′ (r^2)	5′ to 3′ (r^2)
^{S17CF}V$_H$	0.89	0.62
^{T30CF}V$_H$	0.63	0.21
^{S53CF}V$_H$	0.80	0.00
^{Q105CF}V$_H$	0.66	0.17
^{S10CF}V$_L$	0.95	0.83
^{S43CF}V$_L$	0.38	0.22
^{S60CF}V$_L$	0.17	0.29
^{S67CF}V$_L$	0.14	0.00
^{T94CF}V$_L$	0.79	0.88
		$p = 0.036$

[a] The quality of the linear fit (r^2) was determined for each series for the inverse of transfer efficiency plotted vs the sixth power of the intermolecular distance (*R*) from either the 3′ to 5′ or 5′ to 3′ oriented model.

*Contrasting the 11F8·**1** and ^{R31S}11F8·**1** Models.* While ^{C9}**1** in the 11F8·**1** model is flipped out to hydrogen bond with ^{R31}V$_H$, serine does not contact ssDNA in the ^{R31S}11F8·**1** model. In the absence of arginine, the ^{R31S}11F8·**1** model predicts a slightly different series of stacking interactions involving interactions with ^{Y32}V$_L$ and ^{Y100}V$_H$, but none with ^{W33}V$_H$ as observed in the 11F8·**1** model. Nevertheless, the average buried surface area (BSA) for the 11F8·**1** complex is 1180 ± 50 Å^2, which is slightly smaller than but not significantly different from the average BSA for the ^{R31S}-11F8·**1** complex (1205 ± 40 Å^2), suggesting that hydrophobic shape complementarity is important for recognition, but not sufficient for 11F8 sequence specificity. Different binding

Table 6: PROCHECK Benchmarks for Stereochemical Quality of the 11F8•**1** and ^{R31S}11F8•**1** Models[a]

General Parameters		
	11F8	^{R31S}11F8
total no. of residues	429	429
no. of residues other than Gly or Pro	373	373
no. of residues in the Ramachandran plot		
most favored regions	311 (83%)	308 (83%)
additional allowed regions	51 (14%)	58 (16%)
generously allowed regions	4 (1%)	2 (1%)
disallowed regions	7 (2%)	5 (1%)
Main Chain Parameters		
ω angle ± standard deviation	421 ± 0.9°	421 ± 0.9°
ζ angle ± standard deviation	399 ± 1.2°	399 ± 1.1°
steric clashes per 100 residues	0	0
H-bond energy standard deviaiton	261 ± 0.9	260 ± 0.9
overall *G* factor	429 (0.2)	429 (0.2)

[a] The program evaluated how a normal number of residue properties are in the model with respect to a set of reference structures. The *G* factor provides an overall measure of the combined stereochemical properties, and values above zero indicate that backbone and side chain atoms are in reasonable conformations.

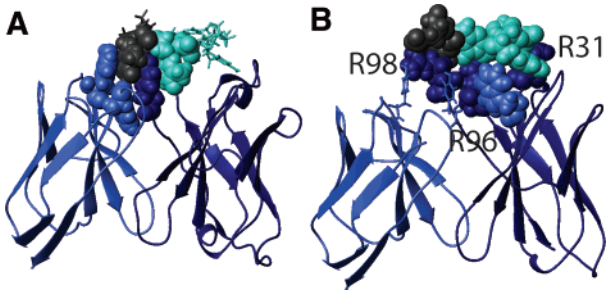

FIGURE 7: Schematics of interface residues in the 11F8•**1** model. Residues are shown as a line, ribbon, or CPK. The variable light chain residues are colored light blue, heavy chain residues navy, T11 and residues to the 5′ end of **1** gray, and T10 and residues to the 3′ end of **1** teal. (A) The conserved ssDNA recognition motif of $^{Y32}V_L$, $^{H91}V_L$ (bottom of binding site), and an aromatic residue at the tip of HCDR3 ($^{Y100}V_H$) stacked with two thymine residues is present in the 11F8•**1** model. (B) Three additional arginine residues ($^{R31}V_H$, $^{R96}V_H$, and $^{R98}V_H$) and an aromatic residue ($^{W33}V_H$) that are not utilized by other autoantibodies to contact ssDNA make contacts in the sequence-specific complex. The nature of the contacts is described further in the text and Figure 8. $^{W33}V_H$ is colored light blue to differentiate it from the three arginine side chains.

site aromatic and basic residues can be involved in stacking to afford nearly equivalent buried surface areas demonstrating that complementarity can be achieved through two deep stacking interactions in the case of the ^{R31S}11F8•**1** complex or by a more extensive surface in the 11F8•**1** complex. These contrasting models suggest that when arginine is present at the periphery of the binding pocket, ssDNA orders to interact with amino acid side chains that are otherwise not utilized in complexes bearing serine at $^{31}V_H$.

DISCUSSION

Hydrophobic Stacking in 11F8 ssDNA Recognition. A common feature of sequence-specific antibody−antigen and protein−ssDNA interactions is binding sites that contain grooves to accommodate nucleotide bases between aromatic and basic residues (*7*, *48*, *50*−*53*). A major contribution to the free energy of bimolecular association in antibody−antigen and protein−ssDNA recognition originates from the

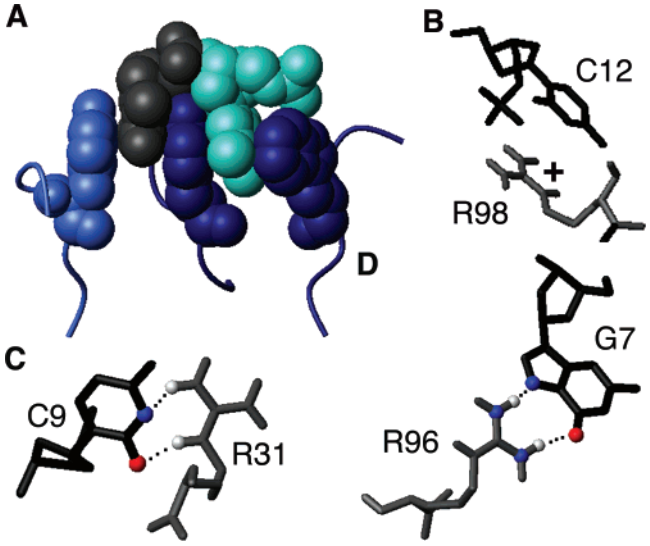

FIGURE 8: Hydrophobic and hydrogen bond interactions from the 11F8•**1** model. Residues are shown as a line, ribbon, or CPK. (A) Stacking interactions at the core of the complex between 11F8 aromatic side chains and T_{10} and T_{11}. The variable light chain residues are colored light blue and heavy chain residues navy, and T11 is colored gray and T10 teal. (B) π-cation interaction between $^{R98}V_H$ and ^{C12}**1**. (C) Sequence-specific bidentate hydrogen bond interaction between $^{R31}V_H$ and ^{C9}**1**. (D) Sequence-specific bidentate interaction between $^{R96}V_H$ and ^{G7}**1**.

release of water molecules associated with nonpolar side chains during the transition from an aqueous environment in the unbound state to a packed hydrophobic environment in the complex called the nonclassical hydrophobic effect (*54*−*58*). Thermodynamic, kinetic, and mutagenesis characterization of 11F8−**1** recognition suggests hydrophobic stacking and concurrent desolvation also provide the driving force for ssDNA recognition (*19*, *23*, *43*). The crystal structure of a highly sequence conserved anti-ssDNA autoantibody, DNA-1, shows that thymine residues bind in the 3′ to 5′ orientation. Significantly, each experiment and docking simulation reported herein support the analogous model in which the thymine residues in **1** stack 3′ to 5′ with 11F8 aromatic residues $^{Y32}V_L$, $^{Y100}V_H$, and $^{W33}V_H$ (Figure 8A). According to this model, the 11F8•**1** complex is stabilized by a multitude of noncovalent interactions and indicates that along with aromatic stacking, diverse contacts between the guanidinium group of binding site arginine residues and ssDNA bases provide a second prominent feature of the 11F8•**1** interface.

Arginine is the most common amino acid at protein−nucleic acid and anti-DNA−DNA interfaces (*59*, *60*). The concentration of arginine residues in contact regions exceeds the frequency of occurrence in other protein regions, which is likely an evolutionary consequence of the propensity for the guanidinium group to form stabilizing interactions. Specifically, arginine can form π-cation interactions between the cationic guanidinium group to stabilize a bound aromatic ligand as well as single or bidentate hydrogen bonds to afford specificity to a variety of substrates (*61*−*66*). π-cation interactions are both specific and strong when fully exposed to water and play an important role in protein−DNA recognition and complex stability (*67*, *68*). For example, π-cation interactions are necessary for recognition and intron splicing by the tRNA-splicing endonuclease in all eukaryotic organisms and govern the specificity of transcription factor

Ndt80 in *Saccharomyces cerevisiae* (*69*, *70*). 11F8 has four arginine residues in the CDR loops: $^{R24}V_L$, $^{R31}V_H$, $^{R96}V_H$, and $^{R98}V_H$. With the exception of $^{R24}V_L$, which is not located in the vicinity of **1**, the model predicts that all arginine residues contact DNA. The model suggests that $^{R98}V_H$ participates in the stacked core of the complex through an energetically favorable π-cation interaction with ^{C12}**1** (Figure 8B).

Hydrogen Bonding in 11F8−ssDNA Recognition. While hydrophobic interactions contribute to the stability of protein·nucleic acid complexes, hydrogen bonds between amino acid side chains and nucleic acid bases often confer sequence specificity (*71*). Due to their role in DNA recognition, arginine residues in the CDRs of 11F8 were systematically mutated to explore the role of specific residues in ssDNA recognition (*19*). Conservative $^{R96K}V_H$ mutation drastically decreased the affinity for **1** (63-fold); however, the mutant protein was less stable than wild-type 11F8, and affinity differences could not be attributed to the loss of contacts at the interface (*19*). Therefore, the role of $^{R96}V_H$ in recognition could not be experimentally determined. Our 11F8·**1** model suggests that $^{R96}V_H$ forms a bidentate hydrogen bond with the G7·C12 loop closing base pair of **1** between the guanadinium group of $^{R96}V_H$ and O6 and N7 of ^{G7}**1** (Figure 8C). Arginine−guanine interactions are by far the most prevalent amino acid−base interaction and have been observed in hundreds of diverse crystal structures, including *Oxytricha nova* telomere end binding protein, HIV-1, nucleolin, Zif268 zinc finger, and RNA synthetases (*52*, *71*−*75*). An arginine residue in U1A spliceosomal protein forms identical interactions with the C5·G16 loop closing base pair of its cognate RNA hairpin, and these interactions have been attributed to sequence specificity (*76*). The proposed interaction between $^{R96}V_H$ and **1** likely facilitates necessary contacts for orienting 11F8 and **1** such that loop thymines efficiently dock into the conserved preorganized thymine binding site.

We have shown that sequence-specific recognition by 11F8 is mediated primarily by $^{R31}V_H$ and $^{Y100}V_H$ and that the aromaticity of $^{100}V_H$ is essential for sequence specificity rather than the hydroxyl group of tyrosine (*19*, *23*, *43*). $^{R31}V_H$ is predicted to form extensive contacts with **1** that are likely critical for both positioning and locking **1** in the correct orientation in the 11F8 binding pocket. Our 11F8·**1** model suggests a bidentate hydrogen bond interaction between the guanidinium group and O2 and N3 of loop cytosine ^{C9}**1** (Figure 8D). Other protein−DNA interactions utilize this same bidentate contact between arginine and cytosine to afford specificity. For example, glutamyl-tRNA synthetases discriminate between the Glu (^{34}YUC^{36}) and Gln (^{34}YUG^{36}) anticodons in higher organisms through a bidentate hydrogen bond between Arg358 and O2 and N3 of C36 (*77*, *78*). Ligand specificities of three nucleic acid aptamers are also partially mediated by a planar bidentate interaction between arginine and the Watson−Crick edge of a cytosine base (*79*). While they do not directly mediate specificity, exploration of other protein·nucleotide complexes in the Protein Data Bank reveals the presence of the bidentate arginine−cytosine contact in many complexes, including DNA glycosylases, ribosome subunits, and additional tRNA synthetases (*75*, *80*, *81*).

Sequence Specificity in Autoantibody−ssDNA Recognition. Identifying the minimum determinants for sequence-specific ligand binding is necessary and interesting for 11F8 since both specificity for **1** and pathogenicity in lupus are consequences of 11F8 affinity maturation. Mutagenesis, thermodynamic, kinetic, and footprinting studies sufficiently identified binding site residues in the 11F8·**1** complex but were unable to differentiate the specific role for each residue (*18*, *19*, *23*, *43*). Our data-driven model of the 11F8·**1** complex provides insight into the structural determinants of specificity, in particular, roles for binding site arginine residues $^{R31}V_H$, $^{R96}V_H$, and $^{R98}V_H$. By comparing our models (for both 11F8·**1** and ^{R31S}11F8·**1**) with related crystal structures, we can identify possible structural determinants of sequence specificity.

Two crystal structures of anti-ssDNA bound to thymine (DNA-1 and BV04-01) reveal that three conserved side chains ($^{Y32}V_L$, a hydrogen-bonding residue at $^{91}V_L$, and an aromatic at the tip of HCDR3) facilitate recognition of two consecutive thymine nucleotides (*35*, *36*). These three side chains ($^{Y32}V_L$, $^{H91}V_L$, and $^{Y100}V_H$) are conserved in 11F8 and two clonally related nonpathogenic mAbs that exhibit high affinity for thymine (*18*). Perhaps this combination of binding site residues predisposes mAbs to thymine intercalation, affording a preference for thymine. Our 11F8·**1** model involves the same recognition hot spot but also suggests additional interactions are involved in recognition beyond those utilized by other anti-ssDNAs, namely, heavy chain residues $^{R31}V_H$, $^{W33}V_H$, $^{R96}V_H$, and $^{R98}V_H$ (Figure 7B) (*35*). Both 11F8 and DNA-1 contain the heavy chain J558 element, yet there is no similarity in the four additional residues utilized by 11F8 to recognize **1**. Sequence differences arise from a combination of the following three sources. First, analysis of the germline gene sequence reveals an $^{S31R}V_H$ somatic mutation in 11F8, while DNA-1 contains the germline serine. Second, $^{W33}V_H$ is a highly conserved germline residue in the J558 family, but DNA-1 has a valine mutation at that position (*82*). Finally, $^{R96}V_H$ and $^{R98}V_H$ result from $D_{Q52}-D_{SP16.2}$ fusion and N-addition at the VDJ junction, respectively (*49*, *83*). Rare events of DD fusions and unusual D reading frames resulting in arginine-rich heavy chain CDR3 have been observed often in anti-DNA (*84*−*86*). With the exception of the $^{R31}V_H$ somatic mutation, all four residues are conserved among clonally related mAbs 9F11 and 15B10. Because 9F11 and 15B10 are neither sequence-specific nor pathogenic, contacts among $^{W33}V_H$, $^{R96}V_H$, and $^{R98}V_H$ may be necessary for binding ssDNA with high affinity but are not sufficient for a sequence-specific complex.

$^{S31R}V_H$ and $^{F100Y}V_H$ are two germline mutations that arise in the 11F8 CDRs during affinity maturation and differ from 9F11 and 15B10. We have shown that sequence specificity and thermodynamic and kinetic parameters associated with recognition of **1** are essentially the same in the presence of $^{F100}V_H$ and $^{Y100}V_H$, suggesting that $^{R31}V_H$ is solely responsible for the differences observed for 11F8 (*45*). $^{R31}V_H$ is located on the periphery of the binding pocket, removed from the binding site intercalation observed in related autoantibody crystal structures as well as the ^{R31S}11F8·**1** model. The observation that a residue at the periphery of the 11F8 binding site such as $^{R31}V_H$ affects specificity is supported by structural analysis of the location of somatic mutations in known antibody crystal structures. Naturally occurring function-altering somatic mutations occur more frequently at sites not directly interacting with the antigen (*87*). The

same trend was observed in anti-HEL in vitro, where a higher substrate affinity was achieved by mutations at the periphery of the antigen-binding site rather than residues directly interacting with the antigen (*88*). Additionally, efforts toward directed evolution of new protein function by hypermutation have shown that while direct mutation to the nonessential side chains at the recognition interface often has little effect on the native interaction, mutations to peripheral underutilized parts of the protein can have significant effects on recognition and subsequent reactivity (*89*). This observation is probably due to the fact that peripheral residues are not part of the general recognition interface or contribute an important structural element and instead have greater conformational flexibility to impart new properties without affecting primary function.

Our observations help explain the effects of somatic mutation in 11F8. In the crystal structure of DNA-1 with thymine as well as the ^{R31S}11F8·**1** model, no contacts are observed between $^{31}V_H$ and ssDNA. Perhaps 11F8 evolved sequence specificity by extending the binding interface to include protein side chains that are not utilized in other anti-DNA·DNA complexes. The combination of $^{R31}V_H$ and $^{R96}V_H$ at the opposite end of the binding site could work in concert to orient the ligand through sequence-specific hydrogen bonds, to maximally interact with the available side chains in the 11F8 binding pocket. Our ^{R31S}11F8·**1** model suggests that recognition in the presence of a single binding site arginine (or no arginine as in DNA-1) is characterized primarily by enthalpically favorable aromatic stacking in a binding site preorganized for thymine intercalation. On the basis of intermolecular contacts in our 11F8·**1** and ^{R31S}11F8·**1** models as well as related anti-ssDNA crystal structures, we hypothesize that $^{R31}V_H$ mediates sequence specificity both directly through bidentate hydrogen bond contacts and indirectly by inducing a slight conformational change to involve additional side chains in the binding interface. Such a hypothesis can now experimentally be tested. 11F8 illustrates evolution of sequence specificity via heavy chain somatic mutation to afford a slowly dissociating protein·ssDNA complex (*43*).

SUPPORTING INFORMATION AVAILABLE

Fluorescence anisotropies of fluorescein and tetramethylrhodamine probes determined individually for each fluorophore. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Bochkarev, A., and Bochkareva, E. (2004) From RPA to BRCA2: Lessons from single-stranded DNA binding by the OB-fold, *Curr. Opin. Struct. Biol. 14*, 36−42.
2. Kur, J., Olszewski, M., Dlugolecka, A., and Filipkowski, P. (2005) Single-stranded DNA-binding proteins (SSBs): Sources and applications in molecular biology, *Acta Biochim. Pol. 52*, 569−574.
3. Desveaux, D., Allard, J., Brisson, N., and Sygusch, J. (2002) A new family of plant transcription factors displays a novel ssDNA-binding surface, *Nat. Struct. Biol. 9*, 512−517.
4. Kusakabe, T., Hine, A. V., Hyberts, S. G., and Richardson, C. C. (1999) The Cys4 zinc finger of bacteriophage T7 primase in sequence-specific single-stranded DNA recognition, *Proc. Natl. Acad. Sci. U.S.A. 96*, 4295−4300.
5. Stassen, A. P., Folmer, R. H., Hilbers, C. W., and Konings, R. N. (1994) Single-stranded DNA binding protein encoded by the filamentous bacteriophage M13: Structural and functional characteristics, *Mol. Biol. Rep. 20*, 109−127.
6. Myong, S., Rasnik, I., Joo, C., Lohman, T. M., and Ha, T. (2005) Repetitive shuttling of a motor protein on DNA, *Nature 437*, 1321−1325.
7. Mitton-Fry, R. M., Anderson, E. M., Theobald, D. L., Glustrom, L. W., and Wuttke, D. S. (2004) Structural basis for telomeric single-stranded DNA recognition by yeast Cdc13, *J. Mol. Biol. 338*, 241−255.
8. Theobald, D. L., and Schultz, S. C. (2003) Nucleotide shuffling and ssDNA recognition in *Oxytricha nova* telomere end-bidning protein complexes, *EMBO J. 22*, 4314−4324.
9. Chothia, C., and Lesk, A. M. (1987) Canonical Structures for the Hypervariable Regions of Immunoglobulins, *J. Mol. Biol. 196*, 901−917.
10. Waldman, M., and Madaio, M. P. (2005) Pathogenic autoantibodies in lupus nephritis, *Lupus 14*, 19−24.
11. Tsao, B. P., Ebling, F. M., Roman, C., Panisian-Sahakian, N., Calame, K., and Hahn, B. H. (1990) Structural characterization of the variable regions of immunogloulin genes encoding a pathogenic autoantibody in murine lupus, *J. Clin. Invest. 85*, 530−540.
12. Ravirajan, C. T., Rahman, M. A., Papadaki, L., Griffiths, M. H., Kalsi, J. K., Martin, A. C. R., Ehrenstein, M. R., Latchman, D. S., and Isenberg, D. A. (1998) Genetic, structural and functional properties of an IgG DNA-binding monoclonal antibody from a lupus patient with nephritis, *Eur. J. Immunol. 28*, 339−350.
13. Foster, M. H., Cizman, B., and Madaio, M. P. (1993) Biology of Disease. Nephritogenic autoantibodies in systemic lupus erythematosus: Immunochemical properties, mechanisms of immune deposition, and genetic origins, *Lab. Invest. 69*, 494−507.
14. Koffler, D., Carr, D., Agnello, V., Thiburn, R., and Kunkel, H. G. (1971) Antibodies to polynucleotide in sera: Antigenic specificity and relation to disease, *J. Exp. Med. 134*, 294−312.
15. Katz, J. B., Limpanasithikul, W., and Diamond, B. (1994) Mutational analysis of an autoantibody: Differential binidng and pathogenicity, *J. Exp. Med. 180*, 925−932.
16. Swanson, P. C., Ackroyd, C., and Glick, G. D. (1996) Ligand recognition by anti-DNA autoantibodies. Affinity, specificity, and mode of binding, *Biochemistry 35*, 1624−1633.
17. Swanson, P. C., Yung, R. L., Blatt, N. B., Eagan, M. A., Norris, J. M., Richardson, B. C., Johnson, K. J., and Glick, G. D. (1996) Ligand recogition by murine anti-DNA antibodies. II. Genetic analysis and pathogenicity, *J. Clin. Invest. 97*, 1−13.
18. Stevens, S. Y., and Glick, G. D. (1999) Evidence for sequence-specific recognition of DNA by anti-single-stranded DNA autoantibodies, *Biochemistry 38*, 560−568.
19. Cleary, J., and Glick, G. D. (2003) Mutational Analysis of a Sequence-Specific ssDNA Binding Lupus Autoantibody, *Biochemistry 42*, 30−41.
20. Bendich, A. J., and Bolton, E. T. (1968) The DNA-Agar Procedure, *Methods Enzymol. 12B*, 635−640.
21. Mach, H., Middaugh, C. R., and Lewis, R. V. (1992) Statistical determination of the average values of the extinction coefficients of trypotphan and tyrosine in native proteins, *Anal. Biochem. 200*, 74−80.
22. Haugland, R. P. (2005) *Handbook of Fluorescent Probes and Research Products*, 9th ed., Molecular Probes, Eugene, OR.
23. Ackroyd, P. C., Cleary, J., and Glick, G. D. (2001) Thermodynamic basis for sequence-specific recognition of ssDNA by an autoantibody, *Biochemistry 40*, 2911−2922.
24. Walter, N. G. (2002) Probing RNA structural dynamics and function by fluorescence resonance energy transfer (FRET), *Curr. Protoc. Nucleic Acid Chem. 11.10*, 11.10.11−11.10.23.
25. Haas, E., Katchalski-Katzir, E., and Steinberg, I. Z. (1978) Effect of the Orientation of Donor and Acceptor on the Probability of Energy Transfer Involving Electronic Transitions of Mixed Polarization, *Biochemistry 17*, 5064−5070.
26. Walter, N. G., Burke, J. M., and Millar, D. P. (1999) Stability of hairpin ribozyme tertiary structure is governed by the interdomain junction, *Nat. Struct. Biol. 6*, 544−549.
27. Pereira, M. J. B., Harris, D. A., Rueda, D., and Walter, N. G. (2002) Reaction Pathway of the Trans-Acting Hepatitis Delta Virus Ribozyme: A Conformational Change Accompanies Catalysis, *Biochemistry 41*, 730−740.
28. Rueda, D., Wick, K., McDowell, S. E., and Walter, N. G. (2003) Diffusely bound Mg^{2+} ions slightly reoirent stems I and II of the hammerhead ribozyme to increase the probability of formation of the catalytic core, *Biochemistry 42*, 9924−9936.

29. Walter, N. G., and Burke, J. M. (2000) Fluorescence assays to study structure, dynamics, and function of RNA and RNA-ligand interactions, *Methods Enzymol. 317*, 409−440.
30. Chothia, C., Lesk, A. M., Levitt, M., Amit, A. G., Mariuzza, R. A., Phillips, S. E. V., and Poljak, R. J. (1986) The predicted structure of immunoglobulin D1.3 and its comparison with the crystal structure, *Science 233*, 755−758.
31. Eagan, M. A., Norris, J. M., Cooper, B. C., and Glick, G. D. (1995) Structural patterns in anti-DNA autoantibodies: A molecular modeling study, *Bioorg. Chem. 23*, 482−498.
32. Mueller, F., Doring, T., Erdemir, T., Greuer, B., Junke, N., Osswald, M., Rinke-Appel, J., Stade, K., Thamm, S., and Brimacombe, R. (1995) Getting closer to an understanding of the three-dimensional structure of ribosomal RNA, *Biochem. Cell Biol. 73*, 767−773.
33. Mueller, F., and Brimacombe, R. (1997) A New Model for the Three-dimensional Folding of *Escherichia coli* 16S Ribosomal RNA. I. Fitting the RNA to a 3D Electron Microscopic Map at 20 Å, *J. Mol. Biol. 271*, 524−544.
34. Dominguez, C., Boelens, R., and Bonvin, A. M. J. J. (2003) HADDOCK: A Protien-Protein Docking Approach Based on Biochemical or Biophysical Information, *J. Am. Chem. Soc. 125*, 1731−1737.
35. Tanner, J. J., Komissarov, A. A., and Deutscher, S. L. (2001) Crystal structure of an antigen-binding fragment bound to single-stranded DNA, *J. Mol. Biol. 314*, 807−822.
36. Herron, J. N., He, X. M., Ballard, D. W., Blier, P. R., Pace, P. E., Blothwell, A. L. M., Voss, E. W. J., and Edmundson, A. B. (1991) An autoantibody to single-stranded DNA: A comparison of the three-dimentional structures of the unliganded Fab and a deoxynucleotide-Fab complex, *Proteins 11*, 159−175.
37. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) AQUA and PROCHECK-NMR: Programs for checking the quality of protein structures solved by NMR, *J. Biomol. NMR 8*, 477−486.
38. Furey, W. S., Joyce, C. M., Osborne, M. A., Klenerman, D., Peliska, J. A., and Blasubramanian, S. (1998) Use of fluorescence resonance energy transfer to investigate the conformation of DNA substrates bound to the Klenow fragment, *Biochemistry 37*, 2979−2990.
39. Kohler, J. J., and Schepartz, A. (2001) Kinetic Studies of Fos*Jun*DNA Complex Formation: DNA Binding Prior to Dimerization, *Biochemistry 40*, 130−142.
40. Stryer, L. (1978) Fluorescence Energy Transfer as a Spectroscopic Ruler, *Annu. Rev. Biochem. 47*, 819−846.
41. DiPilato, L. M., Cheng, X., and Zhang, J. (2004) Fluorescent indicators of cAMP and Epac activation reveal differential dynamics of cAMP signaling within discrete subcellular compartments, *Proc. Natl. Acad. Sci. U.S.A. 101*, 16513−16518.
42. Beckingham, J. A., and Glick, G. D. (2001) Sequence specific recognition of ssDNA by a lupus autoantibody: Kinetics and mechanism of binding, *Bioorg. Med. Chem. 9*, 2243−2252.
43. Beckingham, J. A., Cleary, J., Bobeck, M. J., and Glick, G. D. (2003) Kinetic Analysis of Sequence-Specific Recognition of ssDNA by an Autoantibody, *Biochemistry 42*, 4118−4126.
44. Ratner, V., Kahana, E., Eichler, M., and Haas, E. (2002) A General Strategy for Site-Specific Double Labeling of Globular Proteins for Kinetic FRET Studies, *Bioconjugate Chem. 13*, 1163−1170.
45. Bobeck, M. J., Cleary, J., Beckingham, J. A., Ackroyd, P. C., and Glick, G. D. (2007) Effect of Somatic Mutation on DNA Binding Properties of Anti-DNA Autoantibodies, *Biopolymers* (in press).
46. Paula, S., Monson, N., and Ball, W. J. J. (2005) Molecular Modeling of Cardiac Glycoside Binding by the Human Sequence Monoclonal Antibody 1B3, *Proteins 60*, 382−391.
47. Kim, Y.-R., Kim, J.-S., Lee, S.-H., Lee, W.-R., Sohn, J.-N., Chung, Y.-C., Shim, H.-K., Lee, S.-C., Kwon, M.-H., and Kim, Y.-S. (2006) Heavy and Light Chain Variable Single Domains of an Anti-DNA Binding Antibody Hydrolyze Both Double- and Single-stranded DNAs without Sequence Specificity, *J. Biol. Chem. 281*, 15287−15295.
48. Pellequer, J.-L., Chen, S.-w. W., Keum, Y.-s., Karu, A. E., Li, Q. X., and Roberts, V. A. (2005) Structural basis for preferential binding of non-ortho-substituted polychlorinated biphenyls by the monoclonar antibody S2B1, *J. Mol. Recognit. 18*, 282−294.
49. Swanson, P. C. (1995) Anti-DNA autoantibodies from a lupus-prone mouse: Ligand binding properties, structure and pathogenicity, Ph.D. Thesis, University of Michigan, Ann Arbor, MI.
50. Lopez, M. M., Yutani, K., and Makhatadze, G. I. (2001) Interactions of the Cold Shock Protein CspB from *Bacillus subtilis* with Single-Stranded DNA, *J. Biol. Chem. 18*, 15511−15518.
51. Theobald, D. L., Mitton-Fry, R. M., and Wuttke, D. S. (2003) Nucleic Acid Recognition by OB-Fold Proteins, *Annu. Rev. Biophys. Biomol. Struct. 32*, 115−133.
52. Horvath, M. P. S., Bevilacqua, J. M., Ruggles, J. A., and Schultz, S. C. (1998) Crystal Structure of the *Oxytricha nova* Telomere End Binding Protein Complexed with Single Stranded DNA, *Cell 95*, 963−974.
53. Larkin, C., Datta, S., Harley, M. J., Anderson, B. J., Ebie, A., Hargreaves, V., and Schildbach, J. F. (2005) Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase, *Structure 10*, 1533−1544.
54. Liu, Y., Yang, Z., Utzat, C. D., Liu, Y., Geacintov, N. E., Basu, A. K., and Zou, Y. (2005) Interactions of human replication protein A with single-stranded DNA adducts, *Biochem. J. 385*, 519−526.
55. Wojciechowski, M., Fogolari, F., and Baginski, M. (2005) Thermodynamic and electrostatic properties of ternary *Oxytricha nova* TEBP-DNA complex, *J. Struct. Biol. 152*, 169−184.
56. Rouzina, I., Pant, K., Karpel, R. L., and Williams, M. C. (2005) Theory of Electrostatically Regulated Binding of T4 Gene 32 Protein to Single- and Double-Stranded DNA, *Biophys. J. 89*, 1941−1956.
57. Bugreeva, I. P., Bugreev, D. V., and Nevinsky, G. A. (2005) Formation of nucleoprotein RecA filament on single-stranded DNA: Analysis by stepwise increase in ligand complexity, *FEBS J. 272*, 2734−2745.
58. Stern, J. C., Aderson, B. J., Owens, T. J., and Schildbach, J. F. (2004) Energetics of the Sequence-specific Binding of Single-stranded DNA by the F Factor Relaxase Domain, *J. Biol. Chem. 279*, 29155−29159.
59. Hoffman, M. M., Khrapov, M. A., Cox, J. C., Yao, J., Tong, L., and Ellington, A. D. (2004) AANT: The Amino Acid-Nucleotide Interaction Database, *Nucleic Acids Res. 32*, D174−D181.
60. Haley, J., Mason, L. J., Nagl, S., Giles, I., Latchman, D. S., Isenberg, D. A., and Rahman, A. (2004) Somatic mutations to arginine residues affect the binding of human monoclonal antibodies to DNA, histones, SmD and Ro antigen, *Mol. Immunol. 40*, 745−758.
61. Wilkinson, T. A., Botuyan, M. V., Kaplan, B. E., Rossi, J. J., and Chen, Y. (2000) Arginine side-chain dynamics in the HIV-1 Rev-RRE complex, *J. Mol. Biol. 303*, 515−529.
62. Wintjens, R., Lievin, J., Rooman, M., and Buisine, E. (2000) Contribution of cation-π interactions to the stability of protein-DNA complexes, *J. Mol. Biol. 302*, 395−410.
63. Allain, F. H.-T., Howe, P. W. A., Neuhaus, D., and Varani, G. (1997) Structural basis of the RNA-binding specificity of human U1A protein, *EMBO J. 16*, 5764−5774.
64. GuhaThakurta, D., and Draper, D. E. (2000) Contribution of basic residues to ribosomal protein L11 recognition of RNA, *J. Mol. Biol. 295*, 569−580.
65. Luscombe, N. M., Austin, S. E., Barman, H. M., and Thorton, J. M. (2001) An overview of the structures of protein-DNA complexes, *Genome Biol. 1*, 1−37.
66. Izatt, R. M., Bradshaw, J. S., Pawlak, K., Bruening, R. L., and Tarbet, B. J. (1992) Thermodynamic and kinetic data for macrocycle interaction with neutral molecules, *Chem. Rev. 92*, 1261−1354.
67. Gallivan, J. P., and Dougherty, D. A. (1999) Cation-π interactions in structural biology, *Proc. Natl. Acad. Sci. U.S.A. 96*, 9459−9464.
68. Gromiha, M. M., Santhosh, C., and Ahmad, S. (2004) Structural analysis of cation-π interactions in DNA binding proteins, *Int. J. Biol. Macromol. 24*, 203−211.
69. Lamoureux, J. S., and Glover, J. N. (2006) Principles of protein-DNA recognition revealed in the structural analysis of Ndt80-MSE DNA complexes, *Structure 14*, 555−565.
70. Trotta, C. R., Paushkin, S. V., Patel, M., Li, H., and Peltz, S. W. (2006) Cleavage of pre-tRNAs by the splicing endonuclease requires a composite active site, *Nature 441*, 375−377.
71. Cheng, A. C., Chen, W. W., Fuhrmann, C. N., and Frankel, A. D. (2003) Recognition of Nucleic Acid Bases and Base-pairs by Hydrogen Bonding to Amino Acid Side Chains, *J. Mol. Biol. 327*, 781−796.
72. Xiaomei, Y., Ellington, G. A., and Patel, D. J. (1996) Deep penetration of an α-helix into a widened RNA major groove in

the HIV-1 rev peptide-RNA aptamer complex, *Nat. Struct. Biol. 3*, 1026−1033.
73. Allain, F. H.-T., Bouvet, P., Dieckmann, T., and Feigon, J. (2000) Molecular basis of sequence-specific recognition of pre-ribosomal RNA by nucleolin, *EMBO J. 19*, 6870−6881.
74. Jamieson, A. C., Wang, H., and Kim, S.-H. (1996) A zinc finger directory for high-affinity DNA recognition, *Proc. Natl. Acad. Sci. U.S.A. 93*, 12834−12839.
75. Moulinier, L., Elier, S., Eriani, G., Gangloff, J., Thierry, J.-C., Gabriel, K., McClain, W. H., and Moras, D. (2001) The structure of an AspRS-tRNAASP complex reveals a tRNA-dependent control mechanism, *EMBO J. 20*, 5290−5301.
76. Oubridge, C., Ito, H., Evans, P. R., Teo, C. H., and Nagai, K. (1994) Crystal structure at 1.92 Angstrom resolution of the RNA-binding domain of the U1A spliceosomal protein complexed with an RNA hairpin, *Nature 372*, 432−438.
77. Sekine, S., Nureki, O., Shimada, A., Vassylyev, D. G., and Yoloyama, S. (2001) Structural basis for anticodon recognition by discriminating glutamyl-tRNA synthetase, *Nat. Struct. Biol. 8*, 203−206.
78. Schulze, J. O., Masoumi, A., Nickel, D., Jahn, M., Jahn, D., Schubert, W.-D., and Heinz, D. W. (2006) Crystal Structure of a Non-discriminating Glutamyl-tRNA Synthetase, *J. Mol. Biol. 361*, 888−897.
79. Hermann, T., and Patel, D. J. (2000) Adaptive Recognition by Nucleic Acid Aptamers, *Science 287*, 820−825.
80. Norman, D. P. G., Bruner, S. D., and Verdine, G. L. (2001) Coupling of Substrate Recognitino and Catalysis by a Human Base-Excision DNA Repair Protein, *J. Am. Chem. Soc. 123*, 359−360.
81. Agalarov, S. C., Prasad, S., Funke, P. M., Stout, C. D., and Williamson, J. R. (2000) Structure of the S15,S6,S18-rRNA Complex: Assembly of the 30S Ribosome Central Domain, *Science 288*, 107−112.
82. Radic, M. Z., and Weigert, M. (1994) Genetic and structural evidence for antigen selection of anti-DNA antibodies, *Annu. Rev. Immunol. 12*, 487−520.
83. Blatt, N. B., Bill, R. M., and Glick, G. D. (1998) Characterization of a unique anti-DNA hybridoma, *Hybridoma 17*, 33−39.
84. Eilat, D. (1990) The role of germline gene expresssion and somatic mutation in the generation of autoantibodies to DNA, *Mol. Immunol. 27*, 203−210.
85. Shlomchik, M., Mascelli, H., Shan, H., Radic, M. Z., Pisetsky, D., Rothstein-Marshak, A., and Weigert, M. (1990) Anti-DNA antibodies from autoimmune mice arise by clonal expansion and somatic mutation, *J. Exp. Med. 171*, 265−297.
86. Radic, M. Z., Mascelli, M. A., Erikson, J., Shan, H., Shlomchik, M., and Weigert, M. (1989) Structural patterns in anti-DNA antibodies from MRL/*lpr* mice, *Cold Spring Harbor Symp. Quant. Biol. 54*, 933−946.
87. Ramirez-Benitez, M. D. C., and Almagro, J. C. (2001) Analysis of Antibodies of Known Structure Suggests a Lack of Correspondence Between the Residues in Contact With the Antigen and Those Modified by Somatic Hypermutation, *Proteins: Struct., Funct., Bioinf. 45*, 199−206.
88. Li, Y., Hongmin, L., Yang, F., Smith-Gill, S. J., and Mariuzza, R. A. (2003) X-ray snapshots of the maturation of an antibody response to a protein antigen, *Nat. Struct. Biol. 10*, 482−488.
89. Aharoni, A., Gaidukov, L., Khersonsky, O., Gould, S. M., Roodveldt, C., and Tawfik, S. (2005) The 'evolvability' of promiscuous protein functions, *Nat. Genet. 37*, 73−76.

BI700212S