THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

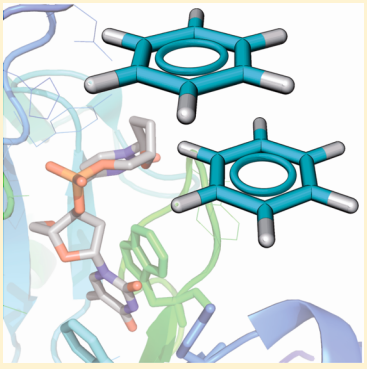
Aromatic Interactions Modulate the 5′-Base Selectivity of the DNA-Binding Autoantibody ED-10

Yi An, Rajesh K. Raju, Tongxiang Lu, and Steven E. Wheeler*

Department of Chemistry, Texas A&M University, College Station, Texas 77843, United States

Supporting Information

ABSTRACT: We present detailed computational analyses of the binding of four dinucleotides to a highly sequence-selective single-stranded DNA (ssDNA) binding antibody (ED-10) and selected point mutants. Anti-DNA antibodies are central to the pathogenesis of systemic lupus erythematosus (SLE), and a more complete understanding of the mode of binding of DNA and other ligands will be necessary to elucidate the role of anti-DNA antibodies in the kidney inflammation associated with SLE. Classical molecular mechanics based molecular dynamics simulations and density functional theory (DFT) computations were applied to pinpoint the origin of selectivity for the 5′-nucleotide. In particular, the strength of interactions between each nucleotide and the surrounding residues were computed using MMGBSA as well as DFT applied to a cluster model of the binding site. The results agree qualitatively with experimental binding free energies, and indicate that π-stacking, CH/π, NH/π, and hydrogen-bonding interactions all contribute to 5′-base selectivity in ED-10. Most importantly, the selectivity for dTdC over dAdC arises primarily from differences in the strength of π-stacking and XH/π interactions with the surrounding aromatic residues; hydrogen bonds play little role. These data suggest that a key Tyr residue, which is not present in other anti-DNA antibodies, plays a key role in the 5′-base selectivity, while we predict that the mutation of a single Trp residue can tune the selectivity for dTdC over dAdC.



I. INTRODUCTION

Systemic lupus erythematosus (SLE) is an autoimmune disease whose underlying cause is not fully understood.[1] Anti-DNA antibodies (Abs) often play a crucial role in the disease pathogenesis by triggering kidney inflammation via antibody−DNA complex deposition.[2] Several non-DNA ligands have been identified that bind to anti-DNA Abs, including both small molecules and macro biomolecules.[3,4] Molecules that bind pathogenic anti-DNA Abs may prevent tissue damage and are considered potential therapeutic agents for the treatment of SLE.[5] Thus, it is crucially important to elucidate the mechanisms by which anti-DNA Abs recognize diverse ligands in order to understand the roles that anti-DNA Abs play in autoimmune disease pathology, and for inhibitor design.[6]

DNA-binding antibodies can also serve as models of DNA and RNA recognition by proteins,[7] which underlie many key biological phenomena.[8−21] These recognition phenomena depend on the interplay of sundry noncovalent interactions, and a complete elucidation of the general mechanism of base recognition by proteins has remained elusive. Selective recognition of DNA and RNA bases has long been attributed to specific hydrogen-bonding interactions.[22−26] On the other hand, the role of aromatic amino acid residues [tyrosine (Tyr), tryptophan (Trp), phenylalanine (Phe), and histidine (His)] in these recognition processes has been a somewhat contentious issue, with previous attempts to quantify their role in DNA and RNA recognition based on statistical analyses of known protein−DNA structures leading to mixed conclusions.[27−31] However, because the strength of π-stacking interactions varies depending on the identity of the aromatic amino acid side chain and nucleobase,[32−36] these interactions will necessarily impact base selectivity. What remains to be quantified, however, is the extent to which π-stacking interactions can impact base selectivity in specific DNA and RNA binding proteins.

Anti-DNA antibodies are classified as either single-stranded (ss) or double-stranded (ds) DNA antibodies depending on what type of DNA they bind.[37−39] So far, only a few crystal structures of ss-DNA−antibody complexes are available.[40−46] Among these structures, ED-10, BV04-01, and DNA-1 are the best examples, with detailed information about affinity, structure, and sequence specificity.[40−46] On the basis of available structural data, Tanner and co-workers identified a conserved and fundamental structural element responsible for the recognition of ssDNA, termed the "ssDNA-antibody recognition module" (D-ARM).[47] The D-ARM consists of a tyrosine residue that stacks with the base and a glycine that forms a hydrogen bond with the base.

Sanguineti et al.[48] investigated the binding of a number of oligonucleotides to monoclonal Ab ED-10, which exhibits subnanomolar binding affinities. Some of their reported relative binding free energies are shown in Table 1. They showed via fluorescence binding experiments and thermodynamic meas-

Received: February 27, 2014
Revised: May 4, 2014
Published: May 6, 2014

dx.doi.org/10.1021/jp502069a | J. Phys. Chem. B 2014, 118, 5653−5659

Table 1. Experimental/MMGBSA Predicted Binding Free Energies/Enthalpies (kcal mol^{-1}), Relative to the Native Antigen, 5′-dTdC, for Four 18mer Oligonucleotides/Dinucleotides Bound to the Native Antibody (ED-10) and Three Point Mutants

	ED-10		point mutants[b]		
	Exp[a]	MMGBSA	AlaH50	PheH50	TyrH50
5′-dTdC	0.0	0.0	0.0	0.0	0.0
5′-dAdC	5.2	8.7	2.8	6.5	5.4
5′-dGdC	5.1	0.7			
5′-dCdC	4.9	14.7			

[a]Experimental binding free energies are for 18-mer oligonucleotides starting with the listed sequence. [b]For point mutants, enthalpies are calculated instead of free energies.

urements that the 5′-end of ssDNA has the highest binding affinity,[48] and that only the first two bases from the 5′-end (dTdC in the case of the native antigen) interact strongly with the antibody. Moreover, it was found that the binding affinity is insensitive to the oligonucleotide length, as long as the sequence 5′-dTdC is present. On the other hand, if the 5′-base is anything other than thymine, the binding free energy is about 5 kcal mol^{-1} less favorable.

Sanguineti et al.[48] also reported a 1.89 Å resolution crystal structure with a truncated 6-mer nucleotide bound. We note that, in this crystal structure, the 6-mer has been cleaved, leaving 5′-dTdC-3′ bound, with the C3′ hydroxyl group missing from the cytosine nucleotide. The binding site of ED-10 from this crystal structure is depicted in Figure 1. Cytosine is

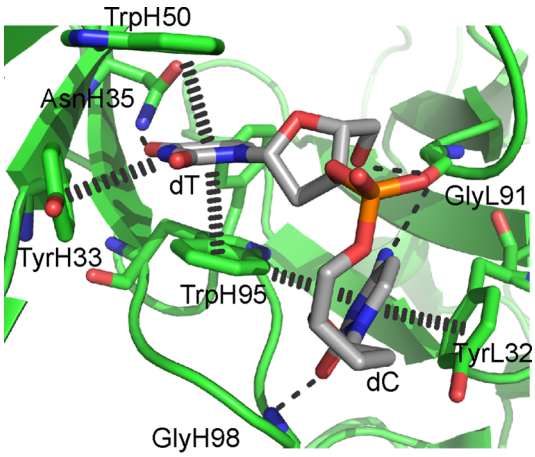


Figure 1. 5′-dTdC-3′ binding site in ED-10. The dC binding pocket is the D-ARM of Tanner and co-workers.[47]

located in the D-ARM, while thymine is bound in a similarly structured binding pocket. π-Stacking interactions from the light chain complementarity-determining region (LCDR) 1 and heavy chain complementarity-determining region (HCDR) 3 as well as hydrogen-bonding interactions from LCDR3 are observed in the D-ARM. The cytosine ring engages in π-stacking interactions with TyrL32 and is roughly perpendicular to the TrpH95 ring system in the tip of HCDR3. Cytosine also forms a hydrogen bond with the main-chain carbonyl from GlyL91. Missing from this binding pocket is the hydrogen-bonding interaction with the L91 side chain, as seen in other examples of the D-ARM.[40−46] Finally, cytosine also accepts a hydrogen bond from the main-chain amide of GlyH98. Sanguineti et al.[48] showed that TyrH97 played an integral role in the binding of this dinucleotide, rearranging with large conformational changes to allow subnanomolar binding affinity.

The binding site for thymine is more extensive, featuring π-stacking interactions with TrpH50 and TrpH95, an edge-to-face interaction with TyrH33, and a hydrogen bond donated from AsnH35 (see Figure 1). The presence of three flanking aromatic amino acid residues in the dT binding pocket appears to be unique among ss-DNA binding antibodies for which detailed structural information is available.[40−46] As shown below, edge-to-face XH/π interactions with TyrH33 are vital for the 5′-dT selectivity exhibited by this antibody.[48]

To date, there have been no detailed computational studies of the origin of base selectivity in anti-ssDNA binding antibodies. Herein, we present computational analyses of the binding of four dinucleotides to ED-10, as well as selected point mutants. We utilize classical molecular mechanics based molecular dynamics (MD) simulations[49] and density functional theory (DFT) computations applied to a cluster model of the 5′-nucleotide binding sites.[50−54] The primary aim is to identify the noncovalent interactions that give rise to the pronounced selectivity for 5′-dT exhibited by ED-10, and hence shed light on the origin of sequence selectivity and mode of antigen recognition in this anti-ssDNA antibody.

II. THEORETICAL METHODS

We have explored the ssDNA antibody complex (ED-10) with the native antigen 5′-dTdC bound as well as complexes with 5′-dAdC, 5′-dCdC, and 5′-dGdC. The crystal structure of ED-10 was extracted from PDB file 2OK0, in which the 5′-thymidine presents in an *anti* conformation. The 5′-nucleobase was mutated from thymine to adenine, cytosine, and guanine using Pymol,[55] with each mutated 5′-nucleoside kept as an *anti* conformer. The starting coordinates of each complex were generated by the AMBER package with hydrogens added.[56−58] Each complex was solvated in a 100.0 Å × 74.2 Å × 82.3 Å rectangular box of TIP3P water molecules.[59] Na^+ ions were added to neutralize the system.

With the atoms of the ssDNA antibody complex fixed, the energy of each system was first minimized for 10 000 steps using steepest descent and conjugate gradient algorithms, followed by another 10 000 energy minimization steps without constraints. After energy minimization, each system was heated from 0 to 300 K under constant pressure for a total of 300 ps. A three-staged density equilibration was applied, involving two initial steps with weak restraining forces and a third step without restraints for a total of 2 ns. A 10 ns simulation was then carried out for each protein−DNA complex using the Amber03 force field in an isothermal−isobaric (constant NPT) ensemble after the above 2 ns equilibration.[60] The coordinates of each complex were saved every 2 ps and the temperature maintained at 300 K via Langevin dynamics.[61] The constant pressure of 1 atm was kept using a Nosé−Hoover barostat.[62] The SHAKE algorithm was used to constrain bonds to hydrogen.[63] The particle mesh Ewald method was employed to treat the long-range electrostatic interactions.[64]

Binding enthalpies were estimated using MMGBSA (molecular mechanics/generalized Born model and surface area) for each ssDNA antibody complex based on 1000 frames recorded during the above-described 10 ns trajectories.[42] The single trajectory approach was used to calculate the binding enthalpy of dTdC, dAdC, dGdC, and dCdC in ED-10. Due to the computational cost of normal-mode analyses, the entropic

dx.doi.org/10.1021/jp502069a | J. Phys. Chem. B 2014, 118, 5653−5659

contribution to the relative binding free energy was estimated on the basis of normal-mode analyses applied to 40 frames. The multiple trajectory approach to MMGBSA was also employed, and confirmed the trend of binding energy obtained using the single-trajectory approach.[65,66] Finally, MMGBSA was also used to compute pairwise interactions between the 5′-base and key binding site residues by only considering interactions between pairs of atoms contained within these interacting species.

These MD simulations were complemented by density functional theory (DFT) computations on cluster models of the 5′-binding site at the B97-D/TZV(2d,2p) level of theory.[67−69] This was done primarily to quantify the noncovalent interactions between ssDNA and the binding site amino acid residues. The B97-D functional provides accurate predictions of dispersion-dominated noncovalent interactions, including the π-stacking, CH/π, and NH/π interactions that are germane to this study.[70−73] Binding site amino acids were defined as all the residues within 5 Å of the 5′-nucleobase. The sugar ring and phosphate of the nucleotides were removed, and open valences were capped with hydrogen atoms. To model an amino acid or groups of linked amino acids in the binding site, the terminal peptide bonds were cut. The open valence on NH was capped with hydrogen and a methyl group was added to the carbonyl. The positions of all of the hydrogen atoms were optimized with the remainder of the atoms fixed in space. For the native antigen, dTdC, we performed DFT computations on both the crystal structure binding site geometry as well as the average structure from the MD simulations. The DFT computations for the other dinucleotides were based on the average MD structure, followed by 10 000 steps of energy minimization. Both the interaction energy of the overall binding pocket and the interaction energies between the 5′-base and key residues were evaluated. The interaction energy was computed as the difference in energy between the bound complex and the separated interacting species, all at the complex geometry. All DFT predicted interaction energies correspond to gas-phase computations, which should provide a reasonable model for the interactions within the hydrophobic core of the protein. However, we also evaluated solution-phase interaction energies for the 5′-bases with the entire 5′-binding site, to gauge the effect of solvation. The resulting interaction energies are in qualitative agreement with the gas-phase computations (see the Supporting Information). All DFT computations were performed using Gaussian 09 and utilized density fitting techniques.[74]

III. RESULTS AND DISCUSSION

A. Binding of the Native Antigen (dTdC). First, we discuss the binding of the native antigen, dTdC, by ED-10. In the crystal structure geometry (Figure 1), B97-D applied to a model of the binding site indicates that the strongest interactions involving thymine are the π-stacking interactions with TrpH95 and TrpH50, followed by the hydrogen-bonding interaction with AsnH35 and the NH/π interaction with TyrH33 (see Table 2). It is notable that the hydrogen-bonding interaction with AsnH35 provides less stabilization of this complex than the π-stacking interactions with Trp residues. As seen below, hydrogen bonding also plays only a minor role in the selectivity for the 5′-base. For cytosine, the strongest interaction stems from the loop containing TrpH95, ThrH96, TyrH97, GlyH98, and SerH99, which together contribute about −13.8 kcal mol^{-1} to the overall binding energy. This is complemented by strong hydrogen-bonding interactions with GlyL91 and the π-stacking interaction between cytosine and TyrL32 (Table S1 in the Supporting Information). Overall, the DFT-predicted interaction energy between dT and the 5′-binding pocket residues in ED-10 is −35.0 kcal mol^{-1}.

Table 2. B97-D Total and Pairwise Interaction Energies (kcal mol^{-1}) between the 5′-Base and Key Amino Acid Residues for Four Dinucleotides with ED-10[a]

	B97-D			
res.	dTdC	dAdC	dGdC	dCdC
TrpH50	−8.6 (−8.4)	−6.0	−9.6	−5.6
TrpH95	−7.7 (−8.9)	−8.0	−9.9	−3.5
AsnH35	−7.8 (−6.7)	−9.2		0.4
TyrH33	−8.9 (−6.3)	−2.9	−17.8[c]	−3.1
GlyL91	−2.5 (−2.6)	−3.6	−2.5	−3.2
TyrH97	−0.3 (−2.6)	−0.3	−3.8	−7.9
total[b]	−38.8 (−35.0)	−32.7	−45.2	−22.8

[a]For dTdC, B97-D interaction energies for the crystal structure geometry are provided in parentheses. [b]Total B97-D interaction energy of the 5′-base with the binding site of ED-10. [c]Interaction of G with TyrH33-MetH34-AsnH35.

The binding site of ED-10 from the averaged structures over the 10 ns trajectory showed some qualitative deviations from the crystal structure. These differences could arise in part due to the presence of the C3′ hydroxyl group on the cytosine nucleotide, which was absent in the reported crystal structure. During the dynamics simulation, the conformation of the dinucleotide does not undergo significant changes but is accompanied by movement of PheL89, TrpH47, SerH56, IleL94 toward dT. The result is a slight compression of the binding site. However, the primary intermolecular interactions depicted in Figure 1 are maintained throughout the trajectory. B97-D interaction energies based on the average MD structure indicate some reordering of the important noncovalent interactions relative to those in the crystal structure geometry, and an overall slight enhancement of the predicted binding of thymine to −38.8 kcal mol^{-1}. In particular, on the basis of the average structure, DFT predicts that the NH/π interaction between thymine and TyrH33 is now stronger than the stacking interactions with tryptophan.

The total MMGBSA predicted binding enthalpy for dTdC in ED-10 is −42.7 kcal mol^{-1}. The main favorable noncovalent interactions involving thymine are π-stacking interactions with TrpH50 and TrpH95, hydrogen-bonding interactions with AsnH35 and GlyL91, and an NH/π interaction with TyrH33 (see Table 3), which is generally consistent with the DFT results for the binding site cluster model. For cytosine, the most favorable interaction is the hydrogen-bonding interaction between NH on HisL27 and the phosphate group on cytosine. The interaction between the hydroxyl group on TyrH97 and the deoxyribose−phosphate backbone of 3′ cytosine is also highly favorable. Other key interactions are the stacking interaction with TyrL32 and hydrogen-bonding interactions with GlyH98 (Table S1 in the Supporting Information). Accounting for entropic effects, the predicted binding free energy for dTdC is −20.3 kcal mol^{-1}.

B. Origin of 5′-Base Selectivity. Experimentally, Sanguineti et al. showed that altering the 5′-nucleobase drastically reduces the binding affinity (see Table 1).[48] Below, we consider the binding of 5′-dTdC-3′ and three other dinucleotides (dAdC, dGdC, and dCdC) to ED-10, in order to quantify the interactions that underlie the strong selectivity for 5′-dT. Table

dx.doi.org/10.1021/jp502069a | J. Phys. Chem. B 2014, 118, 5653−5659

Table 3. MMGBSA Pairwise Interaction Enthalpies between the 5′-Base and Key Amino Acid residues, as Well as Total Binding Ethalpies (ΔH) and Binding Free Energies (ΔG) for Four Dinucleotides with ED-10[a]

	MMGBSA			
res.	dTdC	dAdC	dGdC	dCdC
TrpH50	−4.8	−2.5	−5.5	−2.9
TrpH95	−4.6	−4.7	−4.7	−1.9
AsnH35	−3.5	−3.7	−3.8	0.0
TyrH33	−3.2	−1.4	−4.3	−2.2
GlyL91	−2.8	−1.8	−2.6	−1.1
TyrH97	−1.1	−1.9	−1.2	−2.8
ΔH	−42.7	−33.4	−41.8	−25.7
ΔG	−20.3	−11.6	−19.6	−5.6

[a]All values in kcal mol^{-1}.

3 shows the total MMGBSA binding free energies and enthalpies for these four dinucleotides, as well as the pairwise interactions between the 5′-base and its surrounding residues. Predicted binding enthalpies relative to dTdC are listed in Table 1. MMGBSA results also show that, for all four dinucleotides, the interactions involving the 3′ base (cytosine) are relatively unchanged (see Table S1 in the Supporting Information). That is, the base selectivity at the 5′-end stems from changes in interactions with the 5′-base itself, not secondary effects impacting the 3′ binding pocket.

The MMGBSA data indicate that dAdC, dGdC, and dGdC are all bound less strongly than the native antigen (dTdC), as observed experimentally.[48] However, the magnitude of the change in binding observed experimentally (about 5 kcal mol^{-1}, regardless of the base)[49] is overestimated in the case of dCdC and underestimated for dGdC. In the former case, this is potentially due to rearrangements of TyrH97 and TrpH95 in the binding site upon binding of dCdC[48] that were not captured by the MD simulations. Regardless, these results should still provide qualitative insight into the origin of 5′-dT selectivity in ED-10.

For dAdC, MMGBSA predicts a binding enthalpy that is 9.3 kcal mol^{-1} less favorable than dTdC. This arises primarily from the reduced strength of the π-stacking and edge-to-face interactions with TrpH50 and TyrH33. The decreased strength of these interactions is reflected in the average distances between relevant groups (see Table 4), which show a considerably larger separation between adenine and both TrpH50 and TyrH33 compared with dTdC. The hydrogen bond to GlyL91, which primarily involves the phosphate group, is also reduced slightly with dAdC bound compared to dTdC, which is again reflected in both the MMGBSA pairwise interaction energy and the average distance. Notably, the other major interactions in the binding pocket show little change. In particular, the hydrogen-bonding interaction with AsnH35 is at least as strong in dAdC as in the native antigen, dTdC. Overall, MMGBSA predicts that 2.3 kcal mol^{-1} of the preferential binding of dT over dA arises from differences in π-stacking interactions with TrpH50, while differences between the edge-to-face interactions of the nucleobase with TyrH33 account for 1.8 kcal mol^{-1} of the difference in binding enthalpy. After inclusion of the entropic contribution, the predicted binding free energy for dAdC is 8.7 kcal mol^{-1} less favorable than that for dTdC.

These findings are further supported by B97-D interaction energies, which show that the interaction of adenine with the binding site is 6.1 kcal mol^{-1} less favorable than that of thymine. As seen in the MMGBSA results, this drop is due primarily to reductions in the strength of the π-stacking interactions with TrpH50 and the edge-to-face interaction with TyrH33. In the latter case, the NH/π interaction between thymine and TyrH33 (−8.9 kcal mol^{-1}) is replaced with a much weaker CH/π interaction between adenine and TyrH33 (−2.9 kcal mol^{-1}), contributing significantly to the preferential binding of dTdC over dAdC. Similarly, B97-D predicts that the π-stacking interaction between adenine and TrpH50 is 2.6 kcal mol^{-1} less favorable than the corresponding interaction involving thymine. Thus, even though some of the other interactions operative in the binding site are stronger for adenine (most notably the H-bond to AsnH35), these enhanced interactions are overshadowed by the reduced π-stacking and XH/π interactions with aromatic amino acid side chains.

For dCdC, both B97-D and MMGBSA predict a precipitous drop in binding strength compared to dTdC, due to a reduction in the strength of nearly all interactions with the binding site. Indeed, the strongest interaction with the 5′-dC is the π-stacking interaction with TrpH50, which is predicted to be 2 kcal mol^{-1} weaker than that for dTdC. Given the lack of a suitable H-bond acceptor on cytosine, the H-bonding interaction with AsnH35 observed for dTdC and dAdC is not present with dCdC bound. Indeed, B97-D predicts that the interaction between AsnH35 and the 5′-cytosine is slightly repulsive. These significantly reduced interaction energies are offset slightly by the presence of TyrH97, which provides a favorable interaction with 5′-dC that is not present with any of the other bound dinucleotides. Overall, the present results indicate that selectivity for 5′-dT over dC arises from a combination of hydrogen-bonding, edge-to-face, and π-stacking interactions.

For dGdC, MMGBSA predicts only a 0.9 kcal mol^{-1} reduction in binding enthalpy, and a 0.7 kcal mol^{-1} reduction in binding free energy, relative to dTdC. Moreover, B97-D predicts a more favorable overall interaction between guanine and the binding pocket compared to thymine. In particular, the π-stacking interactions of guanine with TrpH50 and TrpH95 and the NH/π interaction with TyrH33 are both predicted to

Table 4. Mean Distances (in Å) for Key Non-Covalent Interactions[a]

res.	interaction	dTdC	dAdC	dGdC	dCdC
TrpH50	π-stacking	3.6 (98%)	4.5 (0%)	3.6 (99%)	3.8 (79%)
TrpH95	π-stacking	3.6 (98%)	3.6 (97%)	3.7 (94%)	5.4 (0%)
TyrH33	XH/π	2.8	3.5	2.4	2.6
AsnH35	H-bond	1.9 (79%)	2.2 (13%)	2.3 (53%)	>4 (0%)

[a]Values in parentheses are the percent of the MD snapshots in which the centroid−centroid distance was less than 4 Å for π-stacking interactions, and the H···X distance was less than 2 Å for hydrogen bonds.

dx.doi.org/10.1021/jp502069a | J. Phys. Chem. B 2014, 118, 5653−5659

be more favorable compared to the analogous interactions with thymine, and the π-stacking interactions present in the case of guanine are predicted by B97-D to be the most favorable of any of the dinucleotides. These favorable effects for dGdC are somewhat compensated for by other effects. However, overall neither MMGBSA nor B97-D reproduces the experimentally observed 5 kcal mol^{-1} drop in binding affinity for dGdC compared to dTdC.

We have investigated two potential sources of this discrepancy. First, we note that, in all of the above complexes, the 5′-nucleosides (dT, dA, dC, and dG) adopt an *anti* conformation throughout the simulations. We performed analogous simulations with the 5′-nucleosides with a *syn* conformation, to test whether the non-native antigens, and dGdC in particular, bind in a different conformation than dTdC. However, MMGBSA predicted binding enthalpies were less favorable for the *syn* conformer, compared to the anti conformer, for each of the dinucleotides considered. Thus, it does not appear that the 5′-base of the non-native antigens adopts a different conformation than in dTdC. Second, we carried out multiple trajectory MMGBSA computations to investigate whether differences in conformations of the bound and free dinucleotides were contributing to the experimental observation of reduced binding free energies for the other dinucleotides. As seen with the single trajectory results, these multiple trajectory simulations (see Table S2 in the Supporting Information) show that, as 5′-dT is replaced with either dA or dC, the binding enthalpy drops by a significant amount. However, dGdC is predicted to bind slightly more strongly than the native antigen dTdC. That is, both multiple and single trajectory approaches predict a similar binding trend for ssDNA.

C. Impact of TrpH50 Mutations. The above results demonstrated that π-stacking interactions between the 5′-nucleobase and TrpH50 contribute strongly to the preferential binding of dTdC over dAdC. To probe the importance of this interaction further, we considered mutation of TrpH50 to Ala, Phe, or Tyr (see Table 5).

For all three point mutants, the binding enthalpy of dTdC becomes less favorable, while the binding of dAdC is predicted to be more favorable. In other words, although these point mutants are still predicted to be selective for 5′-dT, these mutations reduce the selectivity compared to the native antibody ED-10. The largest change in selectivity is predicted for the AlaH50 mutant, for which dTdC is predicted to be bound by only 2.8 kcal mol^{-1} more than dAdC, compared to the 9.3 kcal mol^{-1} difference for native ED-10.

The origin of this reduced selectivity can be seen by again examining the individual interactions within the binding pockets. For all three point mutants, the interactions in the cytosine pocket are unperturbed relative to those in the native antibody. Moreover, most of the interactions present in the binding site change very little, and the overall changes in binding enthalpy are due primarily to differences in the interaction of the 5′-base with the H50 residue as well as secondary effects on the interaction with TyrH33. In particular, for the AlaH50 mutant, the reduced selectivity for dTdC over dAdC arises from changes in the interactions involving GlyL91 and TyrH33.

Table 5. MMGBSA Predicted Total Binding Enthalpy and Pairwise Interactions (kcal mol^{-1}) between the 5′-Base and Key Amino Acid Residues for TrpH50 Mutants and the dTdC and dAdC Dinucleotides

dTdC	TrpH50 (native)	AlaH50	PheH50	TyrH50
ResH50	−4.8	−0.9	−3.7	−3.8
TrpH95	−4.6	−4.2	−4.5	−4.5
AsnH35	−3.5	−3.5	−3.5	−3.5
TyrH33	−3.2	−3.5	−3.5	−3.2
GlyL91	−2.8	−2.9	−2.7	−2.8
TyrH97	−1.1	−1.0	−1.4	−1.3
total	−42.7	−40.4	−41.5	−41.3
dAdC	TrpH50 (native)	AlaH50	PheH50	TyrH50
ResH50	−2.5	−0.6	−1.6	−1.3
TrpH95	−4.7	−4.5	−4.8	−4.6
AsnH35	−3.7	−3.9	−3.8	−3.1
TyrH33	−1.4	−1.2	−1.4	−0.9
GlyL91	−1.8	−2.3	−2.4	−2.9
TyrH97	−1.9	−0.9	−1.6	−2.1
total	−33.4	−37.6	−35.0	−35.9

IV. SUMMARY AND CONCLUDING REMARKS

Binding free energies and enthalpies for four dinucleotides (dAdC, dGdC, dCdC, and the native antigen, dTdC) by ED-10 were computed using MMGBSA, and further analyzed using DFT applied to cluster models of the binding sites. Overall, MMGBSA provides relative binding enthalpies that are in qualitative agreement with experimental data for two of the three non-native dinucleotides (dAdC and dCdC), while, for dGdC, MMGBSA appears to overestimate the binding affinity relative to dTdC. The MMGBSA data show that the affinity of ED-10 for the 3′ cytosine is relatively constant across the four dinucleotides, and the strong selectivity for dTdC arises primarily through differences in interactions within the 5′-binding site.

Analyses of the non-covalent interactions operative in the 5′-binding site reveal a number of key interactions. In particular, MMGBSA results show that, for dTdC and dGdC, stacking interactions of the 5′-base with TrpH50 and TrpH95 contribute a total of about −10 kcal mol^{-1} to the binding enthalpy. However, the other two dinucleotides (dAdC and dCdC) show weaker overall binding, and π-stacking interactions of the 5′-base with TrpH50 are reduced significantly. Hence, π-stacking interactions with TrpH50 not only contribute to dinucleotide binding but are also central to the 5′-base selectivity of ED-10. Overall, the strong selectivity of ED-10 for dTdC over dAdC arises primarily from differential π-stacking interactions with TrpH95 and differences in XH/π interactions between the nucleobase and TyrH33. We note that this Tyr residue is not present in other ssDNA-binding autoantibodies, suggesting that these other antibodies will not exhibit the same strong 5′-base selectivity as ED-10. Similarly, MMGBSA computations predict that mutations of TrpH50 will provide measurable reductions in the selectivity for dTdC over dAdC, providing an experimental means of quantifying the role of π-stacking interactions in the 5′-base selectivity of this DNA-binding autoantibody.

Overall, both MMGBSA and B97-D data show that π-stacking, CH/π, NH/π, and hydrogen-bonding interactions all contribute to the binding of the 5′-base in complexes of dinucleotides with ED-10. However, these results demonstrate that the strong selectivity for 5′-dT sequences is primarily a result of differences in aromatic interactions (XH/π and π-stacking interactions), not hydrogen-bonding interactions.

dx.doi.org/10.1021/jp502069a | J. Phys. Chem. B 2014, 118, 5653−5659

■ ASSOCIATED CONTENT

Supporting Information

Additional tables and Cartesian coordinates. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: wheeler@chem.tamu.edu. Phone: 979-862-3241.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported in part by the Welch Foundation, Grant A-1775, and we also acknowledge the Texas A&M Supercomputing Facility for providing computational resources. R. L. Stanfield is thanked for fruitful discussions.

■ REFERENCES

(1) Tan, E. M.; Cohen, A. S.; Fries, J. F.; Masi, A. T.; McShane, D. J.; Rothfield, N. F.; Schaller, J. G.; Talal, N.; Winchester, R. J. The 1982 Revised Criteria for the Classification of Systemic Lupus Erythematosus. *Arthritis Rheum.* **1982**, *25*, 1271−1277.

(2) Jang, Y. J.; Stollar, B. D. Anti-DNA Antibodies: Aspects of Structure and Pathogenicity. *Cell. Mol. Life Sci.* **2003**, *60*, 309−320.

(3) Blatt, N. B.; Glick, G. D. Anti-DNA Autoantibodies and Systemic Lupus Erythematosus. *Pharmacol. Ther.* **1999**, *83*, 125−139.

(4) Sibille, P.; Ternynck, T.; Nato, F.; Buttin, G.; Strosberg, D.; Avrameas, A. Mimotopes of Polyreactive Anti-DNA Antibodies Identified Using Phage-Display Peptide Libraries. *Eur. J. Immunol.* **1997**, *27*, 1221−1228.

(5) Amoura, Z.; Koutouzov, S.; Piette, J. G. The Role of Nucleosomes in Lupus. *Curr. Opin. Rheumatol.* **2000**, *12*, 369−373.

(6) Schuermann, J. P.; Prewitt, S. P.; Davies, C.; Deutscher, S. L.; Tanner, J. J. Evidence of Structural Plasticity of Heavy Chain Complementarity-Determining Region 3 in Antibody-Ssdna Recognition. *J. Mol. Biol.* **2005**, *347*, 965−978.

(7) Stevens, S. Y.; Glick, G. D. Evidence for Sequence-Specific Recognition of DNA by Anti-Single-Stranded DNA Autoantibodies†. *Biochemistry* **1998**, *38*, 560−568.

(8) Braddock, D. T.; Baber, J. L.; Levens, D.; Clore, G. M. Molecular Basis of Sequence-Specific Single-Stranded DNA Recognition by Kh Domains: Solution Structure of a Complex between Bnrnp K Kh3 and Single-Stranded DNA. *EMBO J.* **2002**, *21*, 3476−3485.

(9) Price, C. M.; Cech, T. R. Telomeric DNA-Protein Interactions of *Oxytricha* Macronuclear DNA. *Genes Dev.* **1987**, *1*, 783−793.

(10) Gottschling, D. E.; Zakian, V. A. Telomere Proteins: Specific Recognition and Protection of the Natural Termini of *Oxytricha* Macronuclear DNA. *Cell* **1986**, *47*, 195−205.

(11) Lin, J.-J.; Zakian, V. A. The *Saccharomyces Cdc13* Protein Is a Single-Strand Tg_{1-3} Telomeric DNA-Binding Protein in Nitro That Affects Telomere Behavior in Vivo. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 13760−13765.

(12) Baumann, P.; Cech, T. R. Pot1, the Putative Telomere End-Binding Protein in Fission Yeast and Humans. *Science* **2001**, *292*, 1171−1175.

(13) Anderson, E. M.; Halsey, W. A.; Wuttke, D. S. Site-Directed Mutagenesis Reveals the Thermodynamic Requirements for Single-Stranded DNA Recognition by the Telomere-Binding Protein Cdc13. *Biochemistry* **2003**, *42*, 3751−3758.

(14) Bochkarev, A.; Bochkareva, E. From Rpa to Brca2: Lessons from Single-Stranded DNA Binding by the Ob-Fold. *Curr. Opin. Struct. Biol.* **2004**, *14*, 36−42.

(15) Nugent, C. I.; Hughes, T. R.; Lue, N. F.; Lundblad, V. Cdc13p: A Single-Strand Telomeric DNA-Binding Protein with a Dual Role in Yeast Telomere Maintenance. *Science* **1996**, *274*, 249−252.

(16) Lohman, T. M.; Ferrari, M. E. Escherichia Coli Single-Stranded DNA-Binding Protein: Multiple DNA-Binding Modes and Cooperativities. *Annu. Rev. Biochem.* **1994**, *63*, 527−570.

(17) Wold, M. S. Replication Protein A: A Heterotrimeric, Single-Stranded DNA-Binding Protein Required for Eukaruotic DNA Metabolism. *Annu. Rev. Biochem.* **1997**, *66*, 61−92.

(18) Swamynathan, S. K.; Nambiar, A.; Guntaka, R. V. Role of Single-Stranded DNA Regions and Y-Box Proteins in Transcriptional Regulation of Viral and Cellular Genes. *FASEB J.* **1998**, *12*, 515−522.

(19) Raveh, S.; Vinh, J.; Rossier, J.; Agou, F.; Veron, M. Peptidic Determinants and Structural Model of Human Ndp Kinase B (Nm23-H2) Bound to Single-Stranded DNA. *Biochemistry* **2001**, *40*, 5882−5893.

(20) Doe, R. C.; Bonanno, J. B.; Sonenberg, N.; Burley, S. K. Recognition of Polyadenylate Rna by the Poly(A)-Binding Protein. *Cell* **1999**, *98*, 835−845.

(21) Varani, G.; Nagai, K. Rna Recognition by Rnp Proteins During Rna Processing. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 407−445.

(22) 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.* **1976**, *73*, 804−808.

(23) Cheng, A. C.; Chen, W. W.; Fuhrmann, C. N.; Frankel, A. D. Recognition of Nucleic Acid Bases and Base-Pairs by Hydrogen Bonding to Amino Acid Side-Chains. *J. Mol. Biol.* **2003**, *327*, 781−796.

(24) Allers, J.; Shamoo, Y. Structure-Based Analysis of Protein-Rna Interactions Using the Program Entangle. *J. Mol. Biol.* **2001**, *311*, 75−86.

(25) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. Amino Acid-Base Interactions: A Three-Dimensional Analysis of Protein-DNA Interactions at an Atomic Level. *Nucleic Acids Res.* **2001**, *29*, 2860−2874.

(26) Cheng, A. C.; Frankel, A. D. Ab Initio Interaction Energies of Hydrogen-Bonded Amino Acid Side Chain-Nucleic Acid Base Interactions. *J. Am. Chem. Soc.* **2003**, *126*, 434−435.

(27) Luscombe, N. M.; Laskowski, R. A.; Thornton, J. M. On the Molecular Discrimination between Adenine and Guanine by Proteins. *Nucleic Acids Res.* **2001**, *29*, 4294−4309.

(28) Ahmad, S.; Gromiha, M. M.; Sarai, A. Analysis and Prediction of DNA-Binding Proteins and Their Binding Residues Based on Composition, Sequence and Structural Information. *Bioinformatics* **2004**, *20*, 477−486.

(29) Baker, C. M.; Grant, G. H. Role of Aromatic Amino Acids in Protein-Nucleic Acid Recognition. *Biopolymers* **2007**, *85*, 456−470.

(30) Lejeune, D.; Delsaux, N.; Charloteaux, B.; Thomas, A.; Brasseur, R. Protein−Nucleic Acid Recognition: Statistical Analysis of Atomic Interactions and Influence of DNA Structure. *Proteins: Struct., Funct., Bioinf.* **2005**, *61*, 258−271.

(31) Mandel-Gutfreund, Y.; Margalit, H. Quantitative Parameters for Amino Acid-Base Interaction: Implications for Prediction of Protein-DNA Binding Sites. *Nucleic Acids Res.* **1998**, *26*, 2306−2312.

(32) Rutledge, L. R.; Wetmore, S. D. The Assessment of Density Functionals for DNA-Protein Stacked and T-Shaped Complexes. *Can. J. Chem.* **2010**, *88*, 815−830.

(33) Rutledge, L. R.; Durst, H. F.; Wetmore, S. D. Evidence for Stabilization of DNA/Rna-Protein Complexes Arising from Nucleobase-Amino Acid Stacking and T-Shaped Structures. *J. Chem. Theory Comput.* **2009**, *5*, 1400−1410.

(34) Rutledge, L. R.; Campbell-Verduyn, L. S.; Wetmore, S. D. Characterization of the Stacking Interactions between DNA or Rna Nucleobases and the Aromatic Amino Acids. *Chem. Phys. Lett.* **2007**, *444*, 167−175.

(35) Rutledge, L. R.; Campbell-Verduyn, L. S.; Hunter, K. C.; Wetmore, S. D. Characterization of Nucleobase-Amino Acid Stacking Interactions Utilized by a DNA Repair Enzyme. *J. Phys. Chem. B* **2006**, *110*, 19652−19663.

(36) Copeland, K. L.; Anderson, J. A.; Farley, A. R.; Cox, J. R.; Tschumper, G. S. Probing Phenylalanine/Adenine Pi-Stacking Interactions in Protein Complexes with Explicitly Correlated and Ccsd(T) Computations. *J. Phys. Chem. B* **2008**, *112*, 14291−14295.

(37) Cygler, M.; Boodhoo, A.; Lee, J. S.; Anderson, W. F. Crystallization and Structure Determination of an Autoimmune Anti-Poly(Dt) Immunoglobulin Fab Fragment at 3.0-a Resolution. *J. Biol. Chem.* **1987**, *262*, 643−648.
(38) Herron, J. N.; He, X. M.; Ballard, D. W.; Blier, P. R.; Pace, P. E.; Bothwell, A. L. M.; Voss, E. W.; Edmundson, A. B. An Autoantibody to Single-Stranded-DNA - Comparison of the 3-Dimensional Structures of the Unliganded Fab and a Deoxynucleotide Fab Complex. *Proteins: Struct., Funct., Genet.* **1991**, *11*, 159−175.
(39) Mol, C. D.; Muir, A. K. S.; Cygler, M.; Lee, J. S.; Anderson, W. F. Structure of an Immunoglobulin Fab Fragment Specific for Triple-Stranded DNA. *J. Biol. Chem.* **1994**, *269*, 3615−3622.
(40) Ou, Z.; Bottoms, C. A.; Henzi, M. T.; Tanner, J. J. Impact of DNA Hairpin Folding Energetics on Antibody-Ssdna Association. *J. Mol. Biol.* **2007**, *374*, 1029−1040.
(41) Swanson, P. C.; Ackroyd, C.; Glick, G. D. Ligand Recognition by Anti-DNA Autoantibodies. Affinity, Specificity, and Mode of Binding. *Biochemistry* **1996**, *35*, 1624−1633.
(42) Lee, J. S.; Dombroski, D. F.; Mosmann, T. R. Specificity of Autoimmune Monoclonal Fab Fragments Binding to Single-Stranded Deoxyribonucleic-Acid. *Biochemistry* **1982**, *21*, 4940−4945.
(43) Tan, E. M. Antinuclear Antibodies - Diagnostic Markers for Autoimmune-Diseases and Probes for Cell Biology. *Adv. Immunol.* **1989**, *44*, 93−151.
(44) Rumbley, C. A.; Denzin, L. K.; Yantz, L.; Tetin, S. Y.; Voss, E. W. Construction, Characterization, and Selected Site-Specific Mutagenesis of an Anti-Single-Stranded DNA Single-Chain Autoantibody. *J. Biol. Chem.* **1993**, *268*, 13667−13674.
(45) Komissarov, A. A.; Calcutt, M. J.; Marchbank, M. T.; Peletskaya, E. N.; Deutscher, S. L. Equilibrium Binding Studies of Recombinant Anti-Single-Stranded DNA Fab - Role of Heavy Chain Complementarity-Determining Regions. *J. Biol. Chem.* **1996**, *271*, 12241−12246.
(46) Yokoyama, H.; Mizutani, R.; Satow, Y.; Komatsu, Y.; Ohtsuka, E.; Nikaido, O. Crystal Structure of the 64m-2 Antibody Fab Fragment in Complex with a DNA Dt(6−4)T Photoproduct Formed by Ultraviolet Radiation. *J. Mol. Biol.* **2000**, *299*, 711−723.
(47) Tanner, J. J.; Komissarov, A. A.; Deutscher, S. L. Crystal Structure of an Antigen-Binding Fragment Bound to Single-Stranded DNA. *J. Mol. Biol.* **2001**, *314*, 807−822.
(48) Sanguineti, S.; Centeno Crowley, J. M.; Lodeiro Merlo, M. F.; Cerutti, M. L.; Wilson, I. A.; Goldbaum, F. A.; Stanfield, R. L.; de Prat-Gay, G. Specific Recognition of a DNA Immunogen by Its Elicited Antibody. *J. Mol. Biol.* **2007**, *370*, 183−195.
(49) Karplus, M.; Petsko, G. A. Molecular Dynamics Simulations in Biology. *Nature* **1990**, *347*, 631−639.
(50) Churchill, C. D. M.; Wetmore, S. D. Noncovalent Interactions Involving Histidine: The Effect of Charge on Π-Π Stacking and T-Shaped Interactions with the DNA Nucleobases. *J. Phys. Chem. B* **2009**, *113*, 16046−16058.
(51) Rutledge, L. R.; Durst, H. F.; Wetmore, S. D. Evidence for Stabilization of DNA/Rna-Protein Complexes Arising from Nucleobase-Amino Acid Stacking and T-Shaped Interactions. *J. Chem. Theory Comput.* **2009**, *5*, 1400−1410.
(52) Rutledge, L. R.; Wetmore, S. D. The Assessment of Density Functionals for DNA-Protein Stacked and T-Shaped Complexes. *Can. J. Chem.* **2010**, *88*, 815−830.
(53) Rutledge, L. R.; Churchill, C. D. A.; Wetmore, S. D. A Preliminary Investigation of the Additivity of Π-Π or Π^{+}-Π Stacking and T-Shaped Interactions between Natural or Damaged DNA Nucleobases and Histidine. *J. Phys. Chem. B* **2010**, *114*, 3355−3367.
(54) Rutledge, L. R.; Navarro-Whyte, L.; Peterson, T. L.; Wetmore, S. D. Effects of Extending the Computational Model on DNA-Protein T-Shaped Interactions: The Case of Adenine-Histidine Dimers. *J. Phys. Chem. A* **2011**, *115*, 12646−12658.
(55) *The Pymol Molecular Graphics System*, version 1.3r1; Schrodinger, LLC: Portland, OR, 2010.
(56) Pearlman, D. A.; Case, D. A.; Caldwell, J. W.; Ross, W. S.; Cheatham, T. E., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. Amber, a Package of Computer Programs for Applying Molecular Mechanics, Normal Mode Analysis, Molecular Dynamics and Free Energy Calculations to Simulate the Structural and Energetic Properties of Molecules. *Comput. Phys. Commun.* **1995**, *91*, 1−41.
(57) Case, D. A.; Cheatham, T. E.; Darden, T.; Gohlke, H.; Luo, R.; Merz, K. M.; Onufriev, A.; Simmerling, C.; Wang, B.; Woods, R. J. The Amber Biomolecular Simulation Programs. *J. Comput. Chem.* **2005**, *26*, 1668−1688.
(58) Case, D.; Darden, T. A.; Cheatham, T. E.; Simmerling, C.; Wang, J.; Duke, R.; Luo, R.; Crowley, M.; Walker, R.; Zhang, W.; et al. *Amber 11*; University of California: San Francisco, CA, 2011.
(59) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−935.
(60) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; et al. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.* **2003**, *24*, 1999−2012.
(61) Loncharich, R. J.; Brooks, B. R.; Pastor, R. W. Langevin Dynamics of Peptides: The Frictional Dependence of Isomerization Rates of N-Acetylalanyl-N′-Methylamide. *Biopolymers* **1992**, *32*, 523−535.
(62) Martyna, G. J.; Klein, M. L.; Tuckerman, M. Nose-Hoover Chains: The Canonical Ensemble Via Continuous Dynamics. *J. Chem. Phys.* **1992**, *97*, 2635−2643.
(63) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. Numerical Integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of N-Alkanes. *J. Comput. Phys.* **1977**, *23*, 327−341.
(64) Essmann, U.; Perera, L.; Berkowitz, M. L.; Darden, T.; Lee, H.; Pedersen, L. G. A Smooth Particle Mesh Ewald Method. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(65) Kollman, P. A.; Massova, I.; Reyes, C.; Kuhn, B.; Huo, S. H.; Chong, L.; Lee, M.; Lee, T.; Duan, Y.; Wang, W.; et al. Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models. *Acc. Chem. Res.* **2000**, *33*, 889−897.
(66) Miller, B. R.; McGee, T. D.; Swails, J. M.; Homeyer, N.; Gohlke, H.; Roitberg, A. E. Mmpbsa.Py: An Efficient Program for End-State Free Energy Calculations. *J. Chem. Theory Comput.* **2012**, *8*, 3314−3321.
(67) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(68) Becke, A. Density-Functional Thermochemistry. V. Systematic Optimization of Exchange-Correlation Functionals. *J. Chem. Phys.* **1997**, *107*, 8554−8560.
(69) Schafer, A.; Huber, C.; Ahlrichs, R. Fully Optimized Contracted Gaussian Basis Sets of Triple Zeta Valence Quality for Atoms Li to Kr. *J. Chem. Phys.* **1994**, *100*, 5829−5835.
(70) Grimme, S. Semiempirical GGA-Type Density Functional Constructed with a Long-Range Dispersion Correction. *J. Comput. Chem.* **2006**, *27*, 1787−1799.
(71) Raju, R. K.; Bloom, J. W. G.; An, Y.; Wheeler, S. E. Substituent Effects on Non-Covalent Interactions with Aromatic Rings: Insights from Computational Chemistry. *ChemPhysChem* **2011**, *12*, 3116−3130.
(72) Wheeler, S. E. Local Nature of Substituent Effects in Stacking Interactions. *J. Am. Chem. Soc.* **2011**, *133*, 10262−10274.
(73) Bloom, J. W. G.; Raju, R. K.; Wheeler, S. E. Physical Nature of Substituent Effects in Xh/Π Interactions. *J. Chem. Theory Comput.* **2012**, *8*, 3167−3174.
(74) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; et al. *Gaussian 09*, revision B.01; Gaussian, Inc.: Wallingford, CT, 2009.