Molecular Dynamics Free-Energy Simulations of the Binding Contribution to the Fidelity of T7 DNA Polymerase

Jan Florián,*,# Arieh Warshel,# and Myron F. Goodman#,†

Department of Chemistry and Department of Biological Sciences—Hedco Molecular Biology Laboratories, University of Southern California, Los Angeles, California 90089-1062

Received: March 22, 2002

The relative stability of Watson−Crick and mismatched dNTP·template base pairs in the active site of T7 DNA polymerase (pol T7) was examined using linear-response (LRA) molecular dynamics simulations. The LRA method approximates the relative binding free energies by evaluating average electrostatic and van der Waals energies. The results of these computer simulations and the previous simulations of the human DNA polymerase β (pol β) ternary complex (Florian, J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem. B,* **2002**, *106*, 5739) carried out at the same theoretical level were compared. It was found that pol T7 provides a larger binding contribution to the replication fidelity than pol β by discriminating more effectively against mismatches that are most probable to occur in the neutral anti−anti configurations. The selectivity of the pol T7 in the binding step is largely determined by the template−dNTP interactions. These interactions are strengthened by the preorganized protein active site and the contribution of the crystallographic water molecule positioned in the minor groove of the newly forming base pair. The pol T7 active site limits the vertical displacements of dNTP base relative to the base-pair plane more effectively than the pol β active site. On the other hand, the pol T7 active site provides greater flexibility for the dNTP base to be displaced in the major-groove direction.

Introduction

The reproduction and evolution of life depends on the accurate replication of the genome, which is facilitated by DNA polymerases.[1] The synthesis of new DNA molecules would be impossible without these enzymes because they increase the rate of the phosphodiester bond formation about 10^{16} times compared to the corresponding reaction in water.[2−4] Besides achieving great rate enhancement, DNA polymerases are capable of selecting the right deoxyribonucleoside triphosphate (dNTP) substrate so that Watson−Crick base pairs are preferentially formed. The resulting fidelity of DNA synthesis is several orders of magnitude larger than would be expected from the energetics of Watson−Crick base pairing in the end of DNA helix in water.[5] The efforts to establish the source of such remarkable fidelity (e.g., see refs 6−9 and references therein) have been greatly aided by solving the crystal structures of several DNA polymerases.[10,11] Among them, the crystal structures of ternary complexes showing DNA and dNTP substrates bound to human DNA polymerase β (pol β),[12] DNA polymerase from bacteriophage T7 (pol T7),[13,14] and the HIV reverse transcriptase[15] had a profound impact on understanding the catalytic mechanism of DNA replication. However, these studies did not reveal a source of the polymerase fidelity, in part because the structures of the active sites interacting with the base moiety of dNTP were found to be significantly different in various polymerases.

The probability of inserting a "right" versus a "wrong" nucleotide in the growing DNA strand (primer) is affected by both the relative probability of binding of the two dNTP substrates in the enzyme active site and the relative rate constants for the PO bond formation (chemical step). Our goal is to scrutinize both of these steps by calculating the corresponding free-energy changes using atomic-scale computer simulations. In the first part of this project, we evaluated binding contributions to the fidelity of the polymerase β, including the assessment of the role of the key active site amino acid residues.[16] Given the intriguing fact that these residues are not conserved in different polymerase families,[10,11] it would be interesting to compare the structural and energetic features of dNTP binding in different polymerases. Thus, in the present paper, we focus on the binding event in pol T7 and point out the similarities and differences between dNTP binding in pol β and pol T7.

Computational Methods

In this section, we outline key features of our computational approach (for a more detailed description, see the accompanying paper[16]). We also describe the procedure that was used to generate starting configurations for molecular dynamics (MD) simulations.

MD trajectories of the ternary complex of polymerase T7, template−primer DNA, and incoming dNTP in a water environment were calculated in an 18 Å simulation sphere centered on the base moiety of dNTP (Figure 1) using the program Q.[17] Positions of atoms lying beyond the 18 Å sphere were fixed at their crystallographic positions, and their nonbonded interactions with the atoms within the simulation sphere were turned off. The computer simulations were carried out for all combinations of four DNA bases (A, T, C, G) forming neutral dNTP·template base pairs with the exception of the dGTP·G mispair.[16] Only anti−anti conformers were considered. The electrostatic (ES) and van der Waals (vdW) interaction energies between the base moiety of dNTP (S) and its protein + DNA + water environment (s) were sampled and used to determine relative binding

Department of Chemistry.
† Department of Biological Sciences.

10.1021/jp020791m CCC: $22.00
Published on Web 05/14/2002

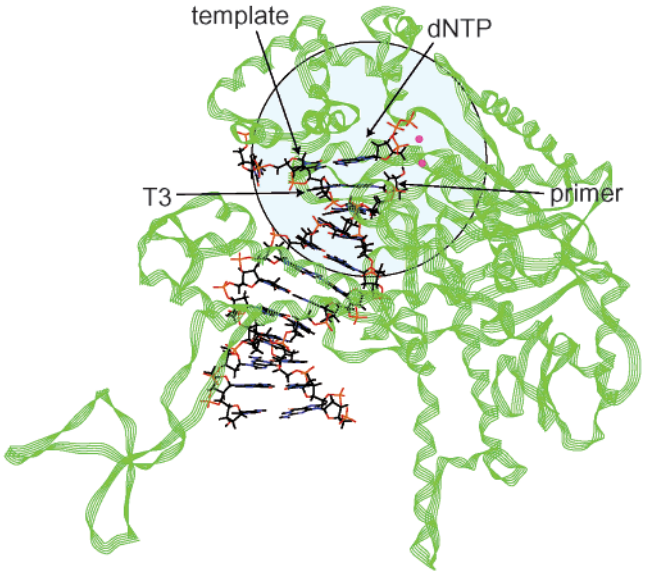

Figure 1. Ternary complex of human DNA polymerase T7 (green), DNA, and dNTP substrates (atom-based colors). The size of the simulation sphere is indicated by the light blue color. Positions of Mg^{2+} ions are indicated by purple spheres. Throughout this paper, the "template" and "primer" denote the template base opposite to the dNTP and the 3′-terminal base of the newly synthesized DNA strand.

free energies, $\Delta\Delta G_{bind}$, using the linear response approximation LRA/α,[16,18]

$$\Delta\Delta G_{bind} = \Delta\Delta G_Q^{ES} + \Delta\Delta G_0^{ES} + \Delta\Delta G_{bind,0} \quad (1)$$

Here, $\Delta\Delta G_{bind,0}$ is approximated as the difference in average vdW interaction energies (U^{vdW}) between S and s containing templates A or B,

$$\Delta\Delta G_{bind,0} = \alpha[\langle U_{Ss}^{vdW}(B)\rangle_Q - \langle U_{Ss}^{vdW}(A)\rangle_Q] \quad (2)$$

where $\langle\rangle$ denotes averaging over a trajectory generated by the molecular dynamics calculations, and $\alpha = 0.56$.[16] The terms $\Delta\Delta G_Q^{ES}$ and $\Delta\Delta G_0^{ES}$ in eq 1 are relative electrostatic energies,

$$\Delta\Delta G_Q^{ES} = \beta[\langle U_{Ss}^{ES}(B)\rangle_Q - \langle U_{Ss}^{ES}(A)\rangle_Q] \quad (3a)$$

$$\Delta\Delta G_0^{ES} = \beta[\langle U_{Ss}^{ES}(B)\rangle_0 - \langle U_{Ss}^{ES}(A)\rangle_0] \quad (3b)$$

where the coefficient $\beta = 0.5$, which is implied by the LRA/α method,[16] was used.

Expressions 3a and 3b differ in the method of determining the trajectories (geometries) for which the electrostatic energies between S and s are evaluated. In $\Delta\Delta G_Q^{ES}$, the trajectories are determined with a full set of charges on S. On the other hand, to evaluate the contribution of structural preorganization of the protein active site, $\Delta\Delta G_0^{ES}$, S was kept uncharged when the trajectories denoted by the subscript 0 were calculated. The electrostatic contributions of a protein or DNA residue (R) to $\Delta\Delta G_{bind}$ of eq 1 were estimated as

$$\Delta\Delta G_{bind}^{R} = 0.5[\langle U_{SR}^{ES}(B)\rangle_Q - \langle U_{SR}^{ES}(A)\rangle_Q + \langle U_{SR}^{ES}(B)\rangle_0 - \langle U_{SR}^{ES}(A)\rangle_0] \quad (4)$$

where subscript SR denotes that only the interactions between the residue R and the nucleobase part of dNTP were included. For R corresponding to a template or primer residue, only the contribution from its nucleobase moiety (i.e., the base + C1′H atoms of the sugar) was included in the U_{SR}^{ES} energy.

Starting structures for MD simulations were generated from the crystal structure of DNA polymerase T7 complexed with the DNA and ddGTP.[14] The actual crystal structure that we used was newly refined to the resolution of 1.85 Å and had the same template/primer/incoming nucleoside triphosphate composition as the originally reported structure (1T7P).[14,19] The 2′,3′-dideoxyribose moieties of ddGTP (incoming nucleotide) and A1022 (3′-terminal nucleotide of the primer strand) in the crystal structure were modified by adding 3′-OH groups. The pol T7 structure was further modified by removing the distant thioredoxin subunit. The histidine residues 506, 704, 607, 683, 417, and 707 were ionized (protonated), and all other histidine residues were kept in their neutral forms. The protonation states of other ionizable residues were determined from their pK_a constants in water. That is, all glutamate and aspartate residues and DNA phosphate groups were negatively charged (−1), whereas all lysine and arginine residues carried charge +1. This selection of charged groups resulted in overall electroneutral systems within the 7 and 10 Å spheres centered on the base of dGTP, and the total −2 charge inside the 18 Å simulation sphere. The base of the incoming nucleotide and the opposite templating base were "mutated" to form all 16 possible base pairs in the pol T7 active site, while keeping glycosidic bonds in their anti conformations corresponding to standard Watson−Crick structures. These structural modifications were done using the "mutate residue" command in the SYBYL 6.6 program,[20] and were followed by manual adjustment of selected torsional angles to remove large steric clashes. The water molecules were added into the simulated system by immersing the simulation sphere into the sphere of bulk water molecules. Those water molecules that were not sterically overlapping with the atoms present in the crystal structure were retained in the starting structure for the MD simulations.

Results

The calculated relative binding free energies ($\Delta\Delta G_{bind}$) and their electrostatic ($\Delta\Delta G_Q^{ES}$), preorganization ($\Delta\Delta G_0^{ES}$), and van der Waals ($\Delta\Delta G^{vdW}$) components for pol T7 are presented in Table 1. The last column shows $\Delta\Delta G_{bind}$ values that were calculated (using the identical computational methodology) for binding in the active site of polymerase β.[16] We would like to emphasize that all of the data presented in Table 1 correspond to neutral anti−anti dNTP−template geometries.[21] Among other configurations, the neutral syn−anti A·A and A·G mispairs, as well as the hemiprotonated A(anti)·C(anti), C(anti)·C(anti), and A(anti)·G(syn) mismatches, are feasible alternatives to neutral anti−anti pairs. The significance of these alternative mismatches is in accord with large magnitudes of $\Delta\Delta G_{bind}$ calculated for the neutral A·A, A·G, A·C, and C·C anti−anti mismatches in pol β,[16] the mispair geometries observed in the X-ray and NMR structures of duplex DNA,[22−24] and small energies required for the protonation of N3 and O2 atoms of cytosine and N1 nitrogen of adenine in aqueous solution.[25,26] The protonated or syn/anti forms or both are less likely for the dCTP·T, dTTP·C, dTTP·T, dTTP·G, and dGTP·T mispairs. For these mispairs, the pol T7 was found to provide larger binding contribution to the polymerization fidelity than pol β. In particular, $\Delta\Delta G_{bind}$ for the C, G, and T opposite to dTTP substrate has increased, respectively, from 1.6, 0.9, and 2.6 kcal/mol to 4.2, 4.2, and 3.5 kcal/mol upon going from pol β to pol T7 (Table 1).

TABLE 1: The Calculated Relative Binding Free Energies ($\Delta\Delta G_{bind}$) for the Protein−DNA−Substrate Complexes Involving the Formation of Mispairs[a]

		pol T7				pol β
substrate	templ	$\Delta\Delta G_Q^{ES}$	$\Delta\Delta G_0^{ES}$	$\Delta\Delta G_{bind,0}$	$\Delta\Delta G_{bind}$	$\Delta\Delta G_{bind}$
dATP	T	0	0	0	0	0
	A[b]	1.0	1.4	1.0	3.4	11.9
	C[b]	3.1	1.8	0.4	5.3	12.7
	G	0.4	2.5	0.7	3.6	1.6
dTTP	A	0.0	0.0	0.0	0.0	0
	C[b]	0.5	2.7	1.0	4.2	1.6
	G	−0.3	4.8	−0.2	4.2	0.9
	T[b]	0.3	2.4	0.8	3.5	2.6
dGTP	C	0.0	0.0	0.0	0.0	0
	A[b]	5.3	3.7	0.5	9.5	7.4
	T[b]	4.7	7.7	0.6	13	10.4
dCTP	G	0.0	0.0	0.0	0.0	0
	C[b]	5.1	1.9	0.8	7.8	3.5
	A[b]	6.1	1.4	0.6	8.1	11.1
	T[b]	3.2	0.5	1.5	5.2	6.1

[a] The results (in kcal/mol) are given relative to the Watson−Crick pair that involves the same dNTP. The relationship between the predicted equilibrium binding constants (K) at 298 K and calculated magnitudes of $\Delta\Delta G_{bind}$ is $1.36(\log K_{WC} - \log K_{mispair}) = -\Delta\Delta G_{bind}$(mispair). $\Delta\Delta G_Q^{ES}$ and $\Delta\Delta G_0^{ES}$ denote the electrostatic components of $\Delta\Delta G_{bind}$ calculated by averaging over trajectories evaluated for the charged and uncharged nucleobase part of dNTP, respectively (eq 3). $\Delta\Delta G_{bind,0}$ denotes the contribution from the binding of an uncharged ligand (eq 2). [b] Average of the results for ternary complexes with and without an extra water molecule inserted in the minor grove (Table 2).

TABLE 2: The Comparison of $\Delta\Delta G_{bind}$ and Its Components[a] Calculated from Trajectories Generated in the Presence or Absence of a Water Molecule in the Initial Structure of the dNTP Binding Site of pol T7

substrate	templ[b]	$\Delta\Delta G_Q^{ES}$	$\Delta\Delta G_0^{ES}$	$\Delta\Delta G_{bind,0}$	$\Delta\Delta G_{bind}$
dATP	A	1.1	1.9	1.0	3.9
	A(w)	0.9	0.9	0.9	2.8
	C	2.4	1.7	0.1	4.2
	C(w)	3.8	2.0	0.6	6.4
dTTP	A	0.0	0.0	0.0	0.0
	A(w)	2.3	2.5	0.0	4.9
	C	2.3	2.1	0.5	4.9
	C(w)	−1.3	3.3	1.4	3.4
	T	0.3	2.3	0.2	2.8
	T(w)	0.2	2.6	1.3	4.1
dGTP	C	0.0	0.0	0.0	0.0
	C(w)	0.6	4.7	−0.6	4.7
	A	6.3	4.2	0.7	11.2
	A(w)	4.4	3.2	0.2	7.8
	T	6.0	7.4	0.4	13.8
	T(w)	3.3	8.1	0.7	12.1
dCTP	A	8.2	1.5	−0.4	9.3
	A(w)	3.9	1.4	1.5	6.8
	C	8.4	1.1	0.0	9.5
	C(w)	1.7	2.8	1.5	6.0
	T	2.8	0.4	1.7	4.9
	T(w)	3.6	0.6	1.4	5.6

[a] The results (in kcal/mol) are given relative to the Watson−Crick pair that involves the same dNTP. $\Delta\Delta G_Q^{ES}$ and $\Delta\Delta G_0^{ES}$ denote the electrostatic components of $\Delta\Delta G_{bind}$ calculated by averaging over trajectories evaluated for the charged and uncharged nucleobase part of dNTP, respectively (eq 3), and $\Delta\Delta G_{bind,0}$ denotes the contribution from binding of an uncharged ligand (eq 2). [b] Symbols A(w), C(w), and T(w) denote structures in which an extra water molecule was present in the minor groove of the dNTP·A, dNTP·C, and dNTP·T base pairs, respectively.

The base-pair geometries that contain variable template bases opposite a common dNTP substrate are given in Figures 2−5. For the presentation in Figures 2−5, the geometries of these

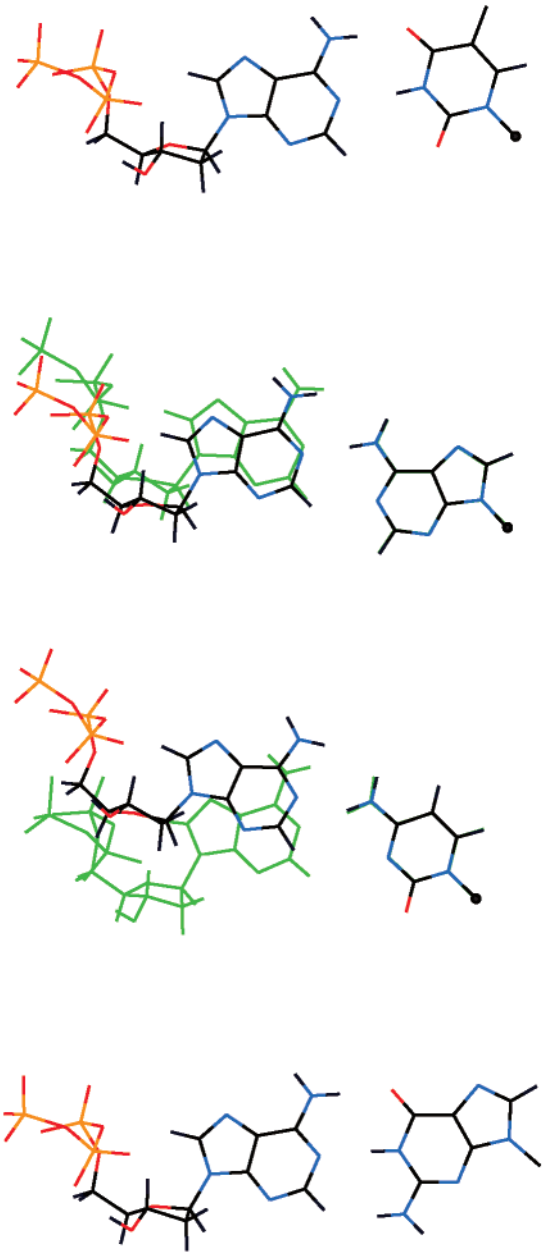

Figure 2. Average structures of the Watson−Crick and mismatched base pairs in the polymerase T7 active site for dATP binding opposite to T, A, C, and G (from top to bottom). The structures obtained for the simulations in the presence and absence of an extra water molecule in the minor groove are drawn using green and atom-type-based colors, respectively.

base pairs were determined using the structure of the pol T7 ternary complex averaged over a 200 ps trajectory. Because crystallographic information was missing for mismatched dNTP·template complexes, we examined a possibility that these structures might be more hydrated than Watson−Crick pairs. This was done by simulating the system in which only the crystallographic water molecule (H_2O698) was present in the minor groove of the dNTP·template base pair (atom-based colored structures in Figures 2−5), as well as by simulating the system with an extra water molecule present in the minor groove (green structures in Figures 2−5). The binding energies for the simulations with and without this additional water molecule are compared in Table 2. The added water molecule destabilizes the dGTP·C base pair by as much as 4.7 kcal/mol. This is an expected result because the crystal structure of the

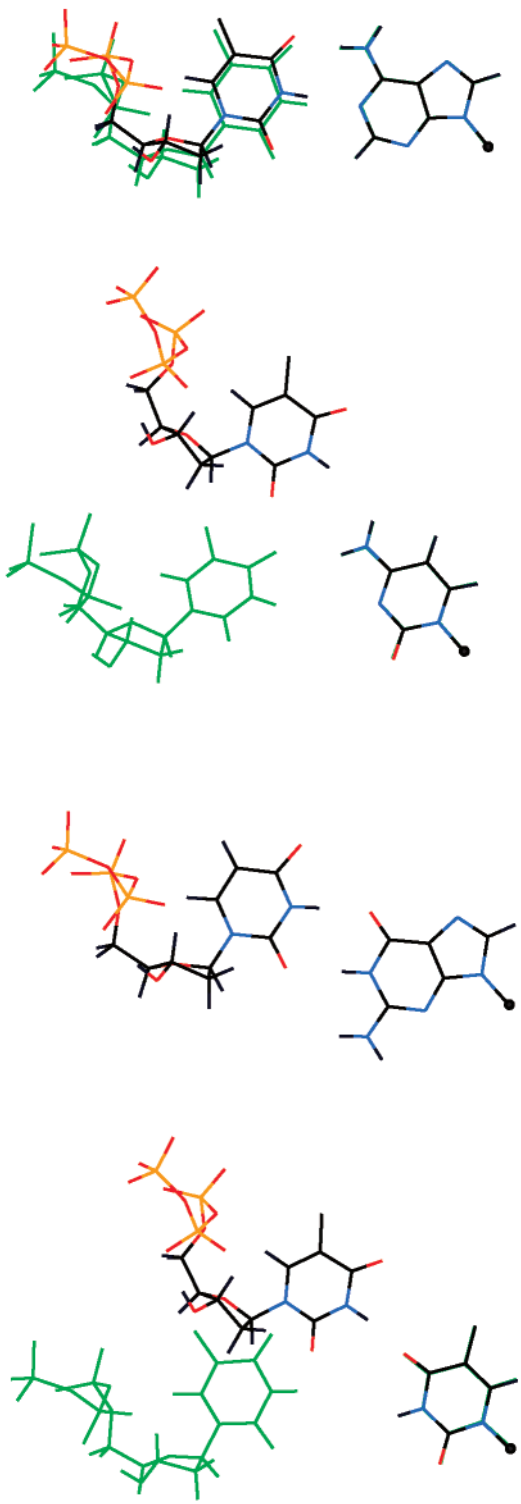

Figure 3. Average structures of the Watson−Crick and mismatched base pairs in the polymerase T7 active site for dTTP binding opposite to A, C, G, and T (from top to bottom). The structures obtained for the simulations in the presence and absence of an extra water molecule in the minor groove are drawn using green and atom-type-based colors, respectively.

T7 ternary complex that contains the dGTP·C base pair[14] does not show any other water molecule in the minor groove besides H_2O698. Thus, the overall stability of the dGTP·C and other Watson−Crick base pairs presented in Table 1 corresponds to simulations without extra water. On the other hand, the extra water molecule was found to stabilize several mispairs (Table 2) and to bring their geometries closer to Watson−Crick shape

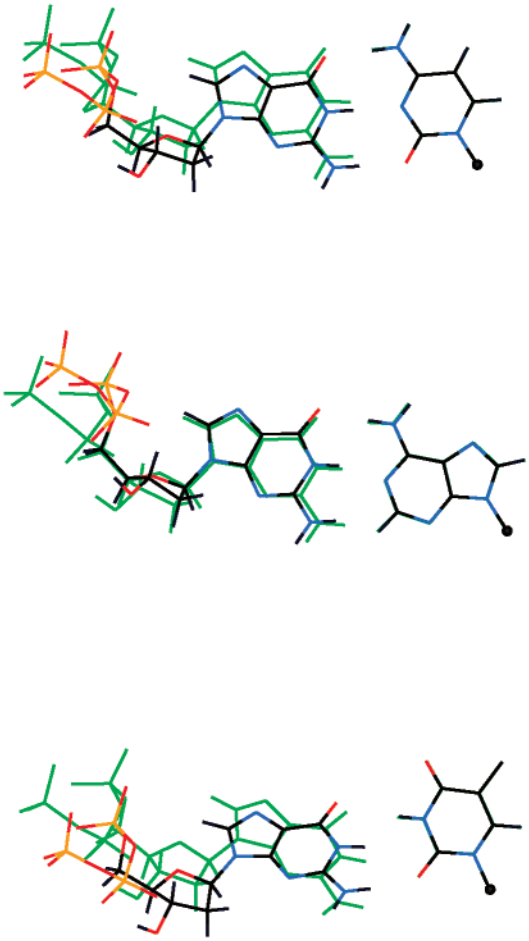

Figure 4. Average structures of the Watson−Crick and mismatched base pairs in the polymerase T7 active site for dGTP binding opposite to C, A, and T (from top to bottom). The structures obtained for the simulations in the presence and absence of an extra water molecule in the minor groove are drawn using green and atom-type-based colors, respectively.

(Figures 2−5). Therefore, we considered in the first approximation configurations of mispairs with and without extra water to be equally probable, and consequently obtained data in Table 1 as arithmetic averages of energies presented in Table 2. Note that the results with extra water molecule for dATP·G and dTTP·G base pairs are not presented in Table 2 and Figure 4. This is because these mispairs were very unstable because of the steric clashes between the guanine amino group and the inserted water molecule. However, the calculated $\Delta\Delta G_{bind}$ values for other templates show remarkable stability with regard to large changes of simulation conditions caused by the insertion of an extra water molecule in the active site.

The protein residues surrounding dNTP·template base pairs are shown in Figure 6. Polymerases T7 and β share some general features of the organization of their nucleobase binding pockets, such as the kinked template backbone that exposes the top surface of the base pair, a hydrogen-bonding network of polar protein residues in the minor groove, or the presence of bulk water along the major groove edge of the dNTP·template base pair. However, all of the specific structural details of the nucleobase binding sites in the two proteins are completely different. The top surface of the base pair seems to face a more nonpolar environment in pol T7 than in pol β. In particular, there are two tyrosine residues (Tyr526 and Tyr530) that are stacked above the base pair. The central element of the minor groove environment in the pol T7 active site is the water

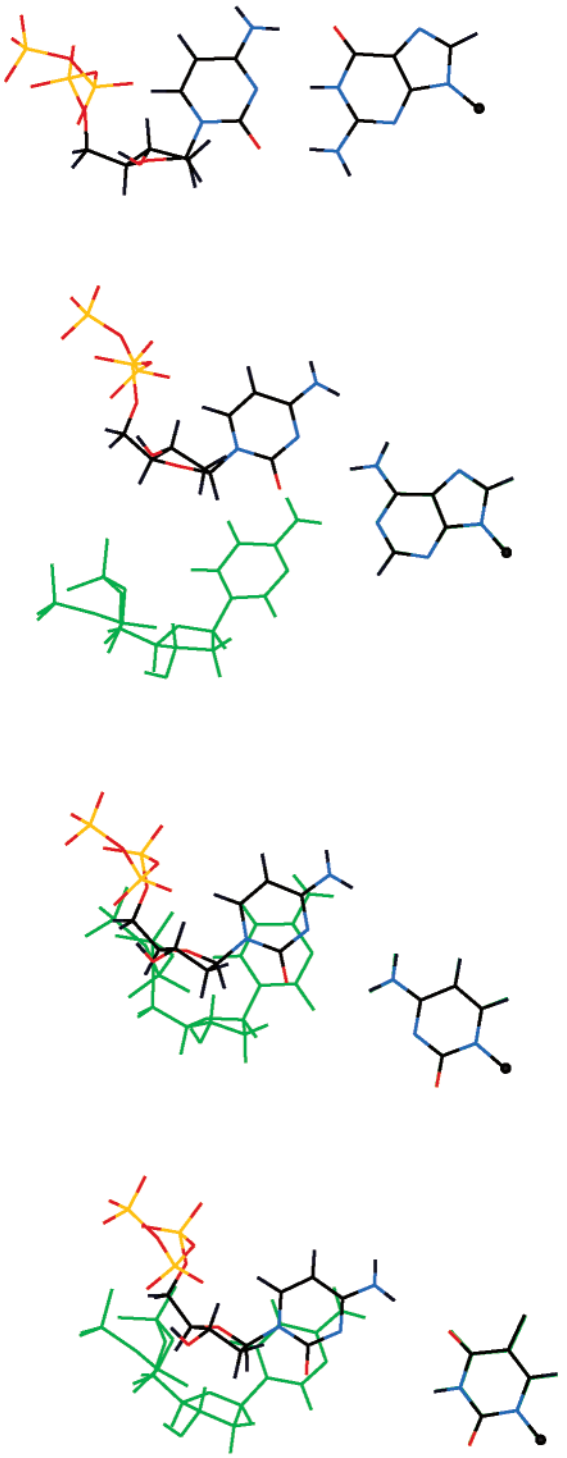

Figure 5. Average structures of the Watson−Crick and mismatched base pairs in the polymerase T7 active site for dCTP binding opposite to G, A, C, and T (from top to bottom). The structures obtained for the simulations in the presence and absence of an extra water molecule in the minor groove are drawn using green and atom-type-based colors, respectively.

molecule observed in the crystal structure (H_2O698). This molecule is kept in its place by hydrogen bonding interactions with Glu480, Asn611, and Gln615 residues. In all of our simulations, the position and contacts of H_2O698 with these amino acid side chains are conserved. The average positions of protein side chains for various mispairs are not shown in Figures 2−5. This is because we believe that the contributions of the nearby protein residues to the stability of a given mispair

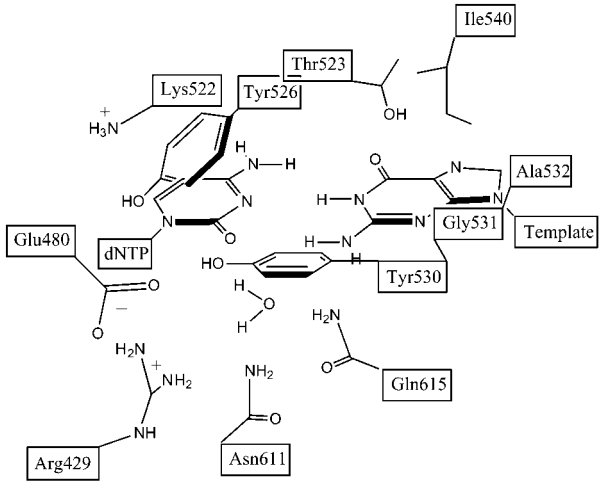

Figure 6. Schematic representation of the protein residues interacting with the base moieties of the template and dNTP substrate in the active site of pol T7. The numbering of the protein residues used in the crystal structure of Doublie et al.[14] was retained. The water molecule shown corresponds to the crystallographic water H_2O698.

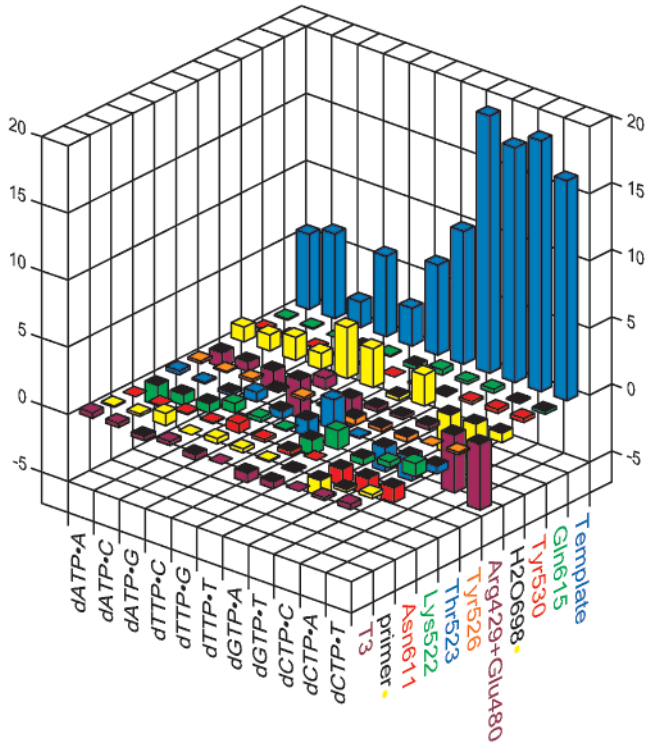

Figure 7. Group contributions to $\Delta\Delta G_{bind}$. For approximate positions of the studied groups in the active site of pol T7, see Figure 6. The negative contributions are indicated by black top of the corresponding bar.

(relative to the corresponding Watson−Crick base pair) can be more rigorously examined using average interaction energies than geometric relationships. These energies are presented in Figure 7 and in Tables 1S−4S of Supporting Information. Here, a large contribution to the binding specificity was found to originate from the electrostatic interactions between incoming dNTP and the template base. H_2O698 plays a significant role in stabilizing dATP·T and dTTP·A Watson−Crick pairs. However, none of the protein residues contribute directly to the binding part of the pol T7 fidelity.

The deoxyribose moieties of dNTP substrates in all base pairs shown in Figures 2−5 have the C3′-endo conformation that is typical of sugars in A−DNA duplexes. The C3′-endo dNTP

conformations were previously found in the simulations of the pol β ternary complex.[16] However, several dNTP substrates that formed mispairs in the pol β active site were found to attain C2′-endo sugar conformations.[16] Also, in the pol T7 active site, the calculated conformations of the primer are C3′-endo, while the template sugars attain exclusively C2′-endo conformations.

There are significant qualitative similarities between geometries of the same mispairs in the pol β and T7 active sites. For example, no lateral base displacement is observed in the average geometries of the dATP·G or dGTP·A base pair in either protein. Also, in both proteins, dATP is found to be displaced into the major groove when it mispairs with A or C in the template. More quantitatively, the dNTP moiety shows considerably larger major groove displacements in pol T7 than in pol β. Significant major groove displacements of the dNTP base occur in the dATP·C (Figure 2), dTTP·C and dTTP·T (Figure 3), and dCTP·C and dCTP·T (Figure 5) mispairs. However, the appearance of the large changes in the dNTP position in these structures (compared with its position in Watson−Crick base pairs) reflects the fact that Figures 2−5 present mispair geometries with approximately parallel glycosidic bonds for template bases in all base pairs. Actually, both the template and dNTP bases in dATP·C, dTTP·C, dTTP·T, dCTP·C, and dCTP·T mismatches are bent away from their positions in the Watson−Crick pair. The presence of an extra water molecule in the minor groove limits the minor groove displacement of the templating base but does not affect the binding energetics, because the magnitudes of $\Delta\Delta G_{bind}$ energies calculated with and without the extra water molecule in the minor groove are quite similar (Table 2). Yet, a positive correlation between the mispair geometry and the misinsertion probability may develop when the contribution of the chemical step is taken into account.

In contrast to the pol T7's larger flexibility for the lateral displacements of dNTP and template, the buckles and propeller twists (i.e., the displacements in the vertical direction) of the dNTP·template base pairs are restricted more effectively by pol T7 than pol β. For example, the dATP·A base pair is less buckled in pol T7 than it is in pol β. As a result, pol T7 is expected to have a smaller tendency to generate frameshift mutations than pol β. The larger vertical rigidity of the dNTP base in pol T7 is probably due to the nonspecific stacking interactions between the dNTP base and the phenol rings of Tyr526 and Tyr530 residues, although different DNA sequences present in the pol β and T7 crystal structures might also contribute to this effect.

Discussion

The goal of the present and the preceding papers was to start building a bridge between the polymerase X-ray structure and its biochemical function. Our "building material" for this bridge was the energetics and dynamics of the ternary complex of pol T7, template−primer DNA, and dNTP, which was evaluated on the atomic scale of detail. Basically, we have assumed that the fidelity can be decomposed into two separable steps: binding of dNTPs (i.e., the probability of having matched versus mismatched dNTPs in the protein active site) and the chemical step. Both steps are known to make a significant contribution to fidelity,[27] but at this stage, we have studied just the ground-state binding contribution to fidelity. In the process, we predicted structures of the polymerase·DNA·dNTP complexes that contain various matched and mismatched dNTP·template base pairs. We found that the dATP, dTTP, and dCTP incoming nucleoside triphosphates tend to be displaced in the direction of the major groove when they are bound with a mismatched template. These mismatches are often stabilized by the presence of an additional water molecule in the minor groove. A rigorous theoretical prediction of the actual presence of water molecules in an enzyme active site requires the evaluation of the free energy of the water insertion process, which exceeds the scope of the present study. Nevertheless, the calculated structures and energetics of mismatched base pairs suggest that an additional water molecule should be present in the minor grove of the ternary complexes containing dTTP·C and dGTP·A mismatches.

Apart from the dominant role of the template−dNTP interactions discussed in the preceding paper,[16] the crystallographic water is the key element for fidelity of AT and TA base pair replication by pol T7. The crystallographic water is also the main reason for the higher calculated binding fidelity of pol T7 than pol β, although differences in lateral versus vertical flexibility of mismatched base pairs in pol T7 and pol β might also play an important role. Contributions of the latter indirect structural effects to polymerase fidelity are however difficult to evaluate quantitatively. On the basis of our results, the mutation of any active site residue that would lead to the destabilization of the crystallographic water or the change of its position/orientation should increase pol T7 misinsertion frequencies by decreasing the binding component of fidelity. For example, the mutation of the Asn611 residue to a nonpolar and slightly bulkier isoleucine seems to be a good candidate for creating a less accurate variant of pol T7. Of course, a quantitative analysis by computer simulations of dNTP binding to such a mutated protein would be a more accurate way of predicting the outcome of protein-engineering experiments. The LRA approach used in this and the previous paper is an efficient tool for such theoretical predictions.

The reliability of our structures cannot be presently judged by comparing them to crystallographic data because the crystal structures of ground-state ternary complexes containing non-Watson−Crick base pairs have not yet been solved. Should such comparison become feasible in the future, one should bear in mind that the force field and sampling limitations led us to neglect equilibria between the neutral, ionized, and tautomeric forms of the bases, and the dNTP and template conformers with syn glycosidic bond torsions.

There are also caveats regarding the correspondence between the calculated free energies and their biochemical counterparts—relative binding constants of matched and mismatched dNTPs. Here, additional biochemical studies of the relative binding of a given dNTP to polymerases containing variable template bases are needed to establish a connection between the theory and experiment. Although such assays are more demanding than the studies of different dNTPs competing for a single template, they can be carried out without difficulty. The availability of pertinent experimental data would benefit greatly the theory in that it would make the calculated magnitudes of $\Delta\Delta G_{bind}$ more accurate through the proper adjustment of the LRA parameter β (eq 3). In this paper, the parameter β was set to its ideal value of 0.5 based on the theory of linear response of the solvent (protein) to its interactions with a solute (dNTP, see also ref 16). Naturally, deviations from linearity may lead to β that differs from its theoretical value. For example, Åqvist and Hansson, who examined the validity of the linear response approximation for various solutes in aqueous solution, found that β for imidazole and *N*-methylacetamide is respectively 0.44 and 0.43.[28] Obviously, deviations from linearity may become larger in such a heterogeneous environment as a protein. Consequently, the predicted $\Delta\Delta G_{bind}$ values may be systemati-

cally overestimated by up to 30% because of uncertainty in the choice of the LRA parameter β. However, the large magnitudes of $\Delta\Delta G_{bind}$ calculated for dGTP·A and dGTP·T mispairs (Table 1) are most probably not caused by the deviation of β from its theoretical value of 0.5 but rather by an insufficient relaxation of the complex from its crystal geometry. Because the crystal structure that we used contained the dGTP·C base pair, the relative energies of dGTP-containing mispairs were expected to show the largest relaxation-related artifacts, which was indeed the case. To diminish these artifacts, the use of a simulation sphere larger than 18 Å should be considered in future studies. Such future work should reveal how much the energetics of the catalytic step of the polymerization reaction are affected by the departure of dNTP·template base pairs from their ideal Watson−Crick geometry.

Acknowledgment. We thank Silvie Doublie and Tom Ellenberger for providing us with their newly refined structure of pol T7. This work was supported by the NIH Grants GM21422 (to M.F.G.) and GM24492 (to A.W.).

Supporting Information Available: Tables 1S−4S containing group contributions to the electrostatic part of the relative binding free energy. This material is available free of charge via the Internet at http://pubs.acs.org.

References and Notes

(1) Kornberg, A.; Baker, T. A. *DNA Replication*; W. H. Freeman: New York, 1992.

(2) Estimated using the observed rate of hydrolysis of dimethyl phosphate in water[3] ($k_{wat} = 10^{-15}$ s^{-1}) and the catalytic rate for polymerase I[4] ($k_{cat} = 5$ s^{-1}) at 25 °C.

(3) Wolfenden, R.; Ridgway, C.; Young, G. *J. Am. Chem. Soc.* **1890**, *120*, 833.

(4) Bryant, F. R.; Johnson, K. A.; Benkovic, S. J. *Biochemistry* **1983**, *22*, 3537.

(5) Petruska, J.; Goodman, M. F.; Boosalis, M. S.; Sowers, L. C.; Cheong, C.; Tinoco, I. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 6252.

(6) Loeb, L. A.; Kunkel, T. A. *Annu. Rev. Biochem.* **1982**, *52*, 429.

(7) Echols, H.; Goodman, M. F. *Annu. Rev. Biochem.* **1991**, *60*, 477.

(8) Wong, I.; Patel, S. S.; Johnson, K. A. *Biochemistry* **1991**, *30*, 526.

(9) Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1.

(10) Steitz, T. A. *J. Biol. Chem.* **1999**, *274*, 17395.

(11) Doublie, S.; Sawaya, M. R.; Ellenberger, T. *Structure* **1999**, *7*, R31.

(12) Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry* **1997**, *36*, 11205.

(13) Doublie, S.; Ellenberger, T. *Curr. Opin. Struct. Biol.* **1998**, *8*, 704.

(14) Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature* **1998**, *391*, 251.

(15) Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. *Science* **1998**, *282*, 1669.

(16) Florián, J.; Goodman, M. F.; Warshel, A. *J. Phys. Chem. B* **2002**, *106*, 5739.

(17) Marelius, J.; Kolmodin, K.; Feierberg, I.; Åqvist, J. *J. Mol. Graphics Modell.* **1999**, *16*, 213.

(18) Sham, Y. Y.; Chu, Z. T.; Tao, H.; Warshel, A. *Proteins Struct. Funct. Genet.* **2000**, *39*, 393.

(19) Doublie, S. Personal communication.

(20) *SYBYL 6.6*; Tripos Inc.: St. Louis, MO, 1999.

(21) Note that throughout this paper we use the terms "primer" and "template" as abbreviations for "3-terminal nucleotide of the primer strand" and "the template base positioned opposite to the dNTP substrate", respectively.

(22) Kennard, O.; Hunter, W. N. *Q. Rev. Biophys.* **1989**, *22*, 327.

(23) Wang, C.; Gao, H.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1991**, *113*, 5486.

(24) Boulard, Y.; Cognet, J. A. H.; Gabarro-Arpa, J.; Le Bret, M.; Sowers, L. C.; Fazakerley, G. V. *Nucleic Acids Res.* **1992**, *20*, 1933.

(25) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.

(26) Florián, J.; Baumruk, V.; Leszczynski, J. *J. Phys. Chem.* **1996**, *100*, 5578.

(27) Johnson, K. A. *Annu. Rev. Biochem.* **1993**, *32*, 685.

(28) Åqvist, J.; Hansson, T. *J. Phys. Chem.* **1996**, *100*, 9512.