HHS Public Access

Author manuscript

Proteins. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

Proteins. 2016 November ; 84(11): 1616–1624. doi:10.1002/prot.25104.

The control of the discrimination between dNTP and rNTP in DNA and RNA polymerase

Hanwool Yoon and **Arieh Warshel**

Department of Chemistry, University of Southern California, Los Angeles, California, 90089-1062

Abstract

Understanding the origin of discrimination between rNTP and dNTP by DNA/RNA polymerases is important both for gaining fundamental knowledge on the corresponding systems and for advancing the design of specific drugs. This work explores the nature of this discrimination by systematic calculations of the transition state (TS) binding energy in RB69 DNA polymerase (gp43) and T7 RNA polymerase. The calculations reproduced the observed trend, in particular when they included the water contribution obtained by the water flooding approach. Our detailed study confirms the idea that the discrimination is due to the steric interaction between the 2′OH and Tyr416 in DNA polymerase, while the electrostatic interaction is the source of the discrimination in RNA polymerase.

Keywords

DNA polymerase; RNA polymerase; sugar selectivity; enzyme catalysis; free energy perturbation; internal water in proteins

I. Introduction

Gene replication and transcription play a key role in the development of biological systems. These processes are catalyzed by DNA/RNA polymerases with high fidelity that maintains very low error rates. Such high fidelity is essential for the stability of reproduction and cellular activity but the occasional errors are also important source of evolution. The mechanism that DNA polymerases adopt to selectively choose the right Watson-Crick base pair between the template strand and the incoming nucleotide substrate has received a great attention in many studies. [1–15] Another important selective behavior of DNA and RNA polymerases is the discrimination against rNTP and dNTP, respectively. It is believed that DNA was evolved from RNA to stabilize earlier genetic system by giving a specific role to DNA as an information carrier in cells [16, 17]. One of the evidence that supports the RNA-to-DNA transition hypothesis is the analogous active site that uses two divalent ions mechanism, as well as the finding that rNTP is reduced to dNTP by reductases. [18, 19]

Understanding the rNTP and dNTP discrimination should help in advancing the fundamental description of replication processes and can also provide a powerful way for developing selective drugs that act against viruses and cancers. [20, 21]

In trying to understand the above selectivity, it is useful to consider a highly conserved tyrosine (e.g. Y416 and Y639 in RB69 DNA polymerase and T7 RNA polymerase, respectively). This residue is thought to be the key residue that contributes to the sugar selectivity, since it is the only residue near the 2′ position of the ribose sugar (see Figure 1) and [10, 22, 23]). Yang et al. [24] has shown that the mutant, Y416A, of RB69 DNA polymerase incorporates rNTP much more efficiently than its wild type. (Table 1) Similarly, but in an opposite way, Sousa et al. [25] mutated the corresponding tyrosine residue of T7 RNA polymerase, Y639, to phenylalanine (since in RNA polymerase, the extra 2′OH is probably preferred by its hydrogen bonds with the hydroxyl in Y639, and the mutant behaved as a DNA polymerase). (Table 4) These experiments were aimed at exploring the assumption [26, 27] that the tyrosine in DNA polymerase acts as a steric gate to the ribose of the incoming rNTP, while the corresponding tyrosine residue in RNA polymerase stabilizes the incoming rNTP by making a hydrogen bond with the hydroxyl group of the ribose.

The above hypothesis cannot be established uniquely by experimental studies, since it is hard to distinguish experimentally between electrostatic and steric contributions. Fortunately, computational approaches can help to determine what are the actual energy contributions that lead to the selectivity. Our previous studies have successfully elucidated sources of the high replication fidelity and the catalytic mechanism of DNA polymerase β [4–6, 13] Thus we apply in this paper the same approach to explore the origin of the sugar selectivity by RB69 DNA polymerase, using our empirical valence bond/free energy perturbation (EVB/FEP) calculations [28–31] and pKa calculations by the semi macroscopic version of the protein dipoles Langavin dipoles (LDLD/S-LRA) approach. [32–34]

II. Methods

Our task is to evaluate the barriers that correspond to k_{pol} and/or the TS binding energies, which are related to k_{pol}/K_D through (See [35])

$$\Delta G^{TS}_{bind} = \Delta g^{\ddagger}_{protein} - \Delta g^{\ddagger}_{water} = -RT\, ln\frac{k_{pol}}{K_D} + RT\, ln k_{wat} \quad (1)$$

The evaluation of this relationship should be done with a sufficiently reliable strategy and the approaches used towards this aim are mentioned below.

II.1 EVB/FEP calculations

To date, arguably the most effective way of evaluating activation free energies of enzymatic reactions is the empirical valence-bond (EVB) method described elsewhere. [28–31] Our previous studies with the EVB have successfully revealed the important factors for the high replication fidelity of DNA pol beta. [4–6, 13] Here, we applied the same method for exploring the selectivity of RB69 DNA polymerase with the following three different substrates ; the natural substrate (PDB: 4FK0) and the mismatched Watson-Crick base pair substrate (PDB: 4FK4), which are reported in Ref. [36]. We also explored the incorrect sugar with the 2′OH, which was obtained by substituting the 2′H by a hydroxyl group for the incorrect sugar structure and relaxing the structure by the same procedure used for the wild

type. The stepwise associative mechanism was considered, as it was found out in our previous studies [37–40] that it gives similar trend to that obtained by the concerted mechanism, but it makes it simpler to explore the TS energy (by evaluating the energy of the pentacovalent intermediate that is being formed between the nucleophilic attack and the bond breaking steps).

The starting structures were immersed in a 18 Å sphere of water molecules subject to the surface-constraint all-atom solvent (SCAAS) type boundary conditions [33]. Long range effects were treated by the local reaction field (LRF) method (which is equivalent to having no cut-off for the electrostatic interactions) [33]. The center of the simulations sphere was defined by the geometric center of the EVB reacting atoms (**defined in** figure S1). The parameters used in the EVB calculations are described in section SI.

The positions of all atoms beyond 18 Å from this center were fixed at their crystallographic positions and their nonbonded interactions with the atoms within the simulation sphere were turned off. Note that the SCAAS allows one to obtain reliable results using smaller explicit systems than those used by most other methods. The protonation states were determined by calculating the pK_a s with our coarse grained (CG) model[41] and then keeping the ionizable residues within 12 Å from the simulation center in their ionization states. The effect of the more distanced ionizable residues was determined macroscopically while using a large charge–charge dielectric [42]. The immersed structures were equilibrated by increasing the temperature from 10 K to 300 K during 200 ps run, while imposing on the protein atoms 100 kcal/mol $Å^2$ harmonic constraint. The temperature was decreased back to 10K over another 200 ps with the same harmonic constraint and the constraint was then released over 100 ps. The systems were reheated to 310 K over 200 ps with fully released constraint followed by 800 ps relaxation. A further relaxation was done for another200 ps to generate 5 different starting configurations for the EVB and FEP calculations by picking a structure at each 40ps. The resulting structures were used in starting the free energy calculations, and the corresponding final results were taken as the average of the 5 runs.

The EVB simulation used 51 frames with each 10 ps for moving the potential from the reactant to the product potential at 300 K with a time step of 1 fs. 10 kcal/mol $Å^2$ harmonic constraint was applied to two magnesium ions for both simulations in protein and reference system (water) which otherwise do not keep their positions close to the reactive site in the water simulation, which are used for the calibration. We note here that the effect of the Mg ions has been explored in many of our studies (e.g. [43, 44]) that include examination of the predictively of the binding of the second metal ions and the late binding of the second ion. More recent studies (e.g.[45, 46] also addressed related questions. However, when the Mg is treated classically without our six dummy atom model it is justified to use harmonic constraint for fixing the mg ions near their observed position.

The FEP simulation of the alchemical transformation used the intermediate state (see below) structures from the EVB simulations. The alchemical simulations were run with the same number of frames and time step as the EVB simulations. As stated above the FEP and EVB calculations were obtained by averaging over the 5 different starting configurations.

We note here that choosing the intermediate state as the model for the TS exploits above mention the fact that the TS for concerted path has similar energetics as the intermediate state (IS), but it provides more stable results for the binding energy calculations. We are well aware of different studies of the TS structure (e.g. [47, 48]) but we believe that our QM/MM analysis [49] is quiet reliable and clearly sufficient for exploring the selectivity towards the hydroxyl of the ribose group which is mainly influenced from the environment of this group. At any rate, in this work we refer to the IS as the TS because of the above considerations.

Our analysis of the reaction mechanism indicated that it is most likely to involve proton transfer (PT) to the bulk and that the PT reaction is not likely to be the rate [50], and thus it is sufficient to calculate ΔG_{PT} instead of $\Delta g^{\ddagger}_{PT}$. This is equivalent to calculating the pK_a shift of the proton donor (the 3′OH group). The calculations of the pKa shift were done by PDLD/S-LRA in the MOLARIS package. [33] The PDLD/S-LRA method is a semimacroscopic method that take the Protein Dipole Langevin Dipoles (PDLD) method and scale it with a physically based dielectric constants. This method evaluates the charging energies with MD relaxation runs for the charged and uncharged stated in the framework of the LRA formulation. The details of this method as been described extensively and can be found for example in ref. [51]

II.2 The Water Flooding approach

The evaluation of binding free energies should involve the use of correct configuration of the water molecules in the active site, in both the bound and unbound states. However, obtaining the corresponding configurations can be very challenging, since overcoming the barriers for water penetration may require very long simulation times. In principle, one may try to randomly insert water molecules and find by Monte Carlo (MC) approach the configuration with the lowest free energy. However, such an approach can take enormous time since most insertion attempt will not be accepted. Our strategy for resolving this challenging problem is the “water flooding “ approach[52], where we first insert clusters of water molecules into the active site, relax and optimize the energy of the selected internal water molecules and use the very efficient LRA approach to estimate the energy of each of the internal water molecule in the given cluster. Next we take the energy of the internal water and evaluate the probability that a water molecule will be inside the protein by running a postprocessing MC simulation with an external program that simply iterate on the energy obtained from the LRA run relative to the energy of water in solution. In this work, we applied the water flooding approach in the TS configuration, and evaluated the TS binding energy in the presence of the most probable water cluster. In the actual simulations we started with standard calculations of the TS in the presence of the original water molecules (obtained by our standard procedure of using an initial grid), but then we removed the water molecules within a radius of 8 Å from the center of our system and reevaluates the population of the internal water by the water flooding approach. In other word, after using MD LRA calculations to estimate the free energy of each water molecule in a relaxed configuration, we then ranked the different configurations by an external MC treatment that obtains the lowest free energy configurations. Next, these processed structures were used in the specific calculation.

In the generation of the relaxed water clusters we solvated the protein with excess water molecules with a minimum distance (r_{cut}) of 1.8Å between the protein atoms and the oxygen of the water molecules (note that the water molecules that are found to be unstable after the LRA relaxation will not be accepted by the MC procedure). Water molecules that are out of the 8 Å radius are inserted normally with r_{cut} = 2.8Å. MD was performed for 100 ps with 0.3 kcal/mol constraints on water. Finally, the solvated structures obtained by the water flooding and the external MC calculation were treated as standard starting structure files and the same protocol and length of the equilibration and relaxation were performed as the simulations without the water flooding.

III. Result

III.1. Analyzing of the selectivity of DNA polymerase

EVB calculations—In first step of our analysis we performed EVB/FEP calculations of the reaction profiles of the different systems. The results of the calculations are given in Fig. 2 As found in our previous studies [4–6, 13], the difference in the activation barrier between the correct and incorrect Watson-Crick base pair substrates was reproduced in this case as well. However, the calculations did not reproduce a significant difference between the calculated barrier of the deoxyribose and ribose sugars systems. Apparently, it is challenging to capture throughout the entire reaction pathway the effect of the small differences in the partial charges of the 2′OH between rNTP and dNTP, because most of the contributions are obtained from the highly charged triphosphate and the base. In fact, Figure 2 shows slightly unfavorable barrier due to the extra hydroxyl group of the ribose sugar. The difficulties of obtaining converging results that reproduce the observed trend reflects in part the fact that the contribution to selectivity from k_{pol} is significantly smaller than that from k_{pol}/K_D. Thus we should try to use the thermodynamic cycle needed to the evaluation of k_{pol}/K_D. With this on mind we considered the cycle in Fig. 3 (see below)

FEP calculation of alchemistic transformations—In view of the discussion above we turn to the use of alchemical transformations in evaluating the energies of the cycle of Fig 3. That is, we evaluated the catalytic efficiency (k_{pol}/K_m), for the wt and the Y416 mutants, by FEP simulations of the relevant alchemical transformations based on the thermodynamic cycles of Fig. 3 (where the energetics of the relevant states is described in Fig.4). We note here that this type of alchemical cycle was used in previous related works ([53, 54]). Finally, it also useful to point out that our alchemical study served also a useful purpose in allowing us to explore the performance of our water flooding approach in an attempt to improve the agreement between the calculated and observed free energies.

Our study explored the effect of mutating the deoxy-ribose sugar of dCTP to the ribose sugar of rCTP using the alchemical pathway, trying to obtain the contribution from the hydroxyl group while leaving other parts of substrate unchanged. More specifically, the relative catalytic efficiency,

$$\Omega=(k_{pol}/K_D)^{[OH]}/(k_{pol}/K_D)^{[H]} \quad (2)$$

can be expressed in terms of the corresponding free energies using the thermodynamic cycle of Fig. 3 as

$$log\,\Omega = -\beta(\Delta\Delta G_{TS}^{[H]->[OH]} - \Delta G_{sol}^{[H]->[OH]})/2.303 \quad (3)$$

where $\mathbf{\Delta\Delta G_{TS}^{[H]->[OH]}}$ and $\mathbf{\Delta G_{sol}^{[H]->[OH]}}$ are the free energies of substituting the 2′H of the dNTP to an hydroxyl group in the TS of the reaction in the protein and in solution, respectively. The actual simulations were done in the opposite direction, mutating rCTP to dCTP, since it is more practical to mutate the hydroxyl group to a hydrogen atom and a dummy atom than to perform the opposite transformation. It is known that free energy calculations associated with mutations to dummy atoms frequently experience an extremely large vdW energy explosion, when the mapping parameter, λ, approaches the limit when we the run on the state were the forces are given almost entirely from the potential with the dummy atom (in this case we can obtain enormous artificial free energy contributions from the van der Waals (vdW) energy of the state with the real atom). This situation is known as "vdW catastrophe" [55]. To avoid this problem, we used a modified vdW function whose exponential repulsion part is linear below a cutoff set at 1.8 Å for the hydrogen and 3.0 Å for the oxygen.

The results of our calculations of the relative catalytic efficiency (log Ω) are given in Table 2 and the energy contributions from the FEP procedure are listed in Table 3 and Table S2. We also depict in Figure 5 the relative catalytic efficiency.

In the first set of calculations we obtained overestimates for both the wild type and Y416F, while the results for Y416A agreed with the experimental trend. It is hard to tell whether such overestimates reflect the electrostatics or vdW contributions, since both terms showed about 5 kcal/mol differences compared to the calculations in solution.

Next, we used our water flooding approach, and the corresponding results (Table 3 and Table S2) did improve the agreement with the observed values for the wild type and Y416F (the result for Y416A was slightly worse, but the difference is not significant to be meaningful). The detail comparison between the electrostatic and vdW contributions in Table 3 shows that the improvement comes mostly from the increase in the vdW contributions.

As seen from Figure 6, it is hard to see a clear differences between structures generated by the water flooding approach and the structures obtained by the regular protein solvation procedure. However, there are always extra water molecule besides the tyrosine. Thus, it seems that the water flooding treatment has improved the results by better relaxation rather than by direct interactions between the substrate and the optimally selected water molecules.

To further inspect the source of improvements we calculated the distances between the sugar and the tyrosine (Figure 7). The original structural study [36] has suggested that the 3′OH makes an hydrogen bond with the backbone nitrogen of Y416, and thus increase the binding energy of the incoming nucleotide. Indeed, the hydrogen bond during the calculation with the original structure (without the water flooding) was observed for most of times (Figure 7-a). Interestingly, however, after the structure was processed by the water flooding approach

and then simulated by the FEP approach, it almost always showed a longer H-bonds distance. Note that all results reported in this paper are averaged over 5~10 different initial trajectories. Some trajectories did not even make a hydrogen bonds and still reproduced the observed energetics better than without the water flooding. To see this effect more clearly, we also performed FEP mutation, subjecting the hydrogen bond length to a constraint with an equilibrium length of 2 Å and a force constant of 4.0 kcal/mol $Å^2$. As shown in Table 3, regardless of the use of the water flooding approach, the results were worse with the hydrogen bond constraint except for the alanine mutant. As a consequence of this 'loosen' hydrogen bond in the structures generated by the water flooding approach, the distance between the 2′OH and the Cγ of Y416 was also slightly further as well. This explains the alleviated vdW contributions in the wild type and the Y416F calculations.

III.2 Analyzing the selectivity of RNA polymerase

Building on the success in the evaluation the selectivity in DNA polymerase we used the same alchemical procedure for T7 RNA polymerase. In this case we used the α β methylene ATP elongation complex structure (PDB:1S76 [56]), where we modified the substrate to ATP and mutated Y416 to a phenylalanine, in order to simulate the corresponding mutant. The results are summarized in Tables 5 and 6. Apparently, when we just ran the simulations with our ENZYMIX force field, the proposed ([13, 15]) hydrogen bond between the O2′ and the hydroxyl group of Y416 was not obtained and the observed effect of the mutation was not reproduced. Thus, we imposed a 15 kcal/mol $Å^2$ distance constraint with a 1.9 Å equilibrium value, to force the hydrogen bond in the wild type. To position the mutated phenylalanine in a similar position to that of the tyrosine in the constrained wild type system, a similar distance constraint of 15 kcal/mol $Å^2$ and equilibrium distance of 2.8 Å were imposed on the distance between O2′ and the substituted hydrogen of the side chain of the phenylalanine. With these modifications, we could reproduce the experimental data (a reasonable trend was also obtained when the force constant was reduced to 6 kcal/mol $Å^2$). As seen from Table 6 the sugar discrimination in T7 RNA polymerase is most probably due to electrostatic effects since the only difference between the wild type and the mutant is obtained from the electrostatic contribution. One may assume that imposing constraint on the H bond leads to our conclusions. However, it is quite likely that the hydrogen bond classical function should be refined by comparing to ab initio calculations and also should be represented by an EVB hydrogen bond potential. Exploring this issue is left for future studies.

III. Concluding remarks

This work explored the origin of discrimination between rNTP and to dNTP by DNA/RNA polymerases, using systematic calculations of the change in the transition state (TS) binding free energy in RB69 DNA polymerase and T7 RNA polymerase. The calculations reproduced the observed trend and supported the idea that the discrimination is due to the steric interaction between of the 2′OH and Tyr416 in DNA polymerase, while the electrostatics is the source of the discrimination in RNA polymerase.

The calculations demonstrate the importance of the water contribution by the improvement of the results by using the water flooding approach, and indicate that the determination of the correct water configuration can be crucial for reliable binding calculations.

Overall, we believe that the elucidation of the source of the discrimination can help significantly in developing selective drugs that act against viruses and cancers.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health Grants 5U19CA105010. We acknowledge the University of Southern California's High Performance Computing and Communications Center for computer time.

References

1. Beard WA, Wilson SH. Structural insights into the origins of DNA polymerase fidelity. Structure. 2003; 11(5):489–496. [PubMed: 12737815]
2. Doublie S, Sawaya MR, Ellenberger T. An open and closed case for all polymerases. Structure. 1999; 7(2):R31–5. [PubMed: 10368292]
3. Echols H, Goodman MF. Fidelity Mechanisms in DNA-Replication. Annual Review of Biochemistry. 1991; 60:477–511.
4. Florian J, Goodman MF, Warshel A. Computer simulation of the chemical catalysis of DNA polymerases: Discriminating between alternative nucleotide insertion mechanisms for T7 DNA polymerase. Journal of the American Chemical Society. 2003; 125(27):8163–8177. [PubMed: 12837086]
5. Florian J, Goodman MF, Warshel A. Computer simulation studies of the fidelity of DNA polymerases. Biopolymers. 2003; 68(3):286–299. [PubMed: 12601790]
6. Florián J, Goodman MF, Warshel A. Computer simulations of protein functions: Searching for the molecular origin of the replication fidelity of DNA polymerases. Proceedings of the National Academy of Sciences of the United States of America. 2005; 102(19):6819–6824. [PubMed: 15863620]
7. Florian J, Warshel A, Goodman MF. Molecular dynamics free-energy simulations of the binding contribution to the fidelity of T7 DNA polymerase. Journal of Physical Chemistry B. 2002; 106(22): 5754–5760.
8. Goodman MF. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. Annual Review of Biochemistry. 2002; 71:17–50.
9. Goodman MF, et al. Biochemical Basis of DNA-Replication Fidelity. Critical Reviews in Biochemistry and Molecular Biology. 1993; 28(2):83–126. [PubMed: 8485987]
10. Joyce CM. Choosing the right sugar: how polymerases select a nucleotide substrate. Proc Natl Acad Sci U S A. 1997; 94(5):1619–22. [PubMed: 9050827]
11. Joyce CM, Benkovic SJ. DNA polymerase fidelity: Kinetics, structure, and checkpoints. Biochemistry. 2004; 43(45):14317–14324. [PubMed: 15533035]
12. Kool ET. Active site tightness and substrate fit in DNA replication. Annual Review of Biochemistry. 2002; 71:191–219.
13. Prasad BR, Warshel A. Prechemistry versus preorganization in DNA replication fidelity. Proteins-Structure Function and Bioinformatics. 2011; 79(10):2900–2919.
14. Showalter AK, et al. Mechanistic comparison of high-fidelity and error-prone DNA polymerases and ligases involved in DNA repair. Chemical Reviews. 2006; 106(2):340–360. [PubMed: 16464009]

15. Steitz TA. DNA polymerases: Structural diversity and common mechanisms. Journal of Biological Chemistry. 1999; 274(25):17395–17398. [PubMed: 10364165]
16. Forterre P. The origin of DNA genomes and DNA replication proteins. Current Opinion in Microbiology. 2002; 5(5):525–532. [PubMed: 12354562]
17. Olsen GJ, Woese CR. Archaeal genomics: An overview. Cell. 1997; 89(7):991–994. [PubMed: 9215619]
18. Lazcano A, et al. On the Early Evolution of Rna-Polymerase. Journal of Molecular Evolution. 1988; 27(4):365–376. [PubMed: 3146647]
19. Lazcano A, et al. The Evolutionary Transition from Rna to DNA in Early Cells. Journal of Molecular Evolution. 1988; 27(4):283–290. [PubMed: 2464698]
20. Jordheim LP, et al. Advances in the development of nucleoside and nucleotide analogues for cancer and viral diseases. Nature Reviews Drug Discovery. 2013; 12(6):447–464. [PubMed: 23722347]
21. Maga G, Hubscher U. Repair and translesion DNA polymerases as anticancer drug targets. Anti-Cancer Agents in Medicinal Chemistry. 2008; 8(4):431–447. [PubMed: 18473728]
22. Huang Y, et al. Mechanism of ribose 2′-group discrimination by an RNA polymerase. Biochemistry. 1997; 36(27):8231–8242. [PubMed: 9204868]
23. Brown JA, Suo ZC. Unlocking the Sugar “Steric Gate” of DNA Polymerases. Biochemistry. 2011; 50(7):1135–1142. [PubMed: 21226515]
24. Yang GW, et al. A conserved Tyr residue is required for sugar selectivity in a pol alpha DNA polymerase. Biochemistry. 2002; 41(32):10256–10261. [PubMed: 12162740]
25. Sousa R, Padilla R. Mutant T7 Rna-Polymerase as a DNA-Polymerase. Embo Journal. 1995; 14(18):4609–4621. [PubMed: 7556104]
26. Georgiadis MM, et al. Mechanistic Implications from the Structure of a Catalytic Fragment of Moloney Murine Leukemia-Virus Reverse-Transcriptase. Structure. 1995; 3(9):879–892. [PubMed: 8535782]
27. Makarova OV, et al. Transcribing of Escherichia coli genes with mutant T7 RNA polymerases: Stability of lacZ mRNA inversely correlates with polymerase speed. Proceedings of the National Academy of Sciences of the United States of America. 1995; 92(26):12250–12254. [PubMed: 8618879]
28. Warshel A, Weiss RM. An Empirical Valence Bond Approach for Comparing Reactions in Solutions and in Enzymes. Journal of the American Chemical Society. 1980; 102(20):6218–6226.
29. Warshel, A. Computer modeling of chemical reactions in enzymes and solutions. New York: Wiley; 1991. p. xiv, 236
30. Aqvist J, Warshel A. Computer-Simulation of the Initial Proton-Transfer Step in Human Carbonic Anhydrase-I. Journal of Molecular Biology. 1992; 224(1):7–14. [PubMed: 1312606]
31. Aqvist J, Fothergill M. Computer simulation of the triosephosphate isomerase catalyzed reaction. Journal of Biological Chemistry. 1996; 271(17):10010–10016. [PubMed: 8626554]
32. Lee FS, et al. Calculations of Antibody Antigen Interactions – Microscopic and Semimicroscopic Evaluation of the Free-Energies of Binding of Phosphorylcholine Analogs to Mcpc603. Protein Engineering. 1992; 5(3):215–228. [PubMed: 1409541]
33. Lee FS, Chu ZT, Warshel A. Microscopic and Semimicroscopic Calculations of Electrostatic Energies in Proteins by the Polaris and Enzymix Programs. Journal of Computational Chemistry. 1993; 14(2):161–185.
34. Warshel A, et al. Examining methods for calculations of binding free energies: LRA, LIE and PDLD/S-LRA calculations of ligands binding to proteins. Biophysical Journal. 2000; 78(1):332a–332a.
35. Singh N, Frushicheva MP, Warshel A. Validating the vitality strategy for fighting drug resistance. Proteins-Structure Function and Bioinformatics. 2012; 80(4):1110–1122.
36. Xia SL, Wang JM, Konigsberg WH. DNA Mismatch Synthesis Complexes Provide Insights into Base Selectivity of a B Family DNA Polymerase. Journal of the American Chemical Society. 2013; 135(1):193–202. [PubMed: 23214497]

37. Kamerlin SCL, et al. A Computational Study of the Hydrolysis of dGTP Analogues with Halomethylene-Modified Leaving Groups in Solution: Implications for the Mechanism of DNA Polymerases (vol 48, pg 5963, 2009). Biochemistry. 2009; 48(32):7776–7776.
38. Rosta E, Kamerlin SCL, Warshel A. On the interpretation of the observed linear free energy relationship in phosphate hydrolysis: A thorough computational study of phosphate diester hydrolysis in solution. Biochemistry. 2008; 47(12):3725–3735. [PubMed: 18307312]
39. Kamerlin SCL, Florian J, Warshel A. Associative versus dissociative mechanisms of phosphate monoester hydrolysis: On the interpretation of activation entropies. Chemphyschem. 2008; 9(12): 1767–1773. [PubMed: 18666265]
40. Klahn M, Rosta E, Warshel A. On the mechanism of hydrolysis of phosphate monoesters dianions in solutions and proteins. Journal of the American Chemical Society. 2006; 128(47):15310–15323. [PubMed: 17117884]
41. Vicatos S, et al. An effective Coarse-grained model for biological simulations: Recent refinements and validations. Proteins-Structure Function and Bioinformatics. 2014; 82(7):1168–1185.
42. Schutz CN, Warshel A. What are the dielectric “constants” of proteins and how to validate electrostatic models? Proteins-Structure Function and Bioinformatics. 2001; 44(4):400–417.
43. Fothergill M, et al. Structure-Energy Analysis of the Role of Metal-Ions in Phosphodiester Bond Hydrolysis by DNA-Polymerase-I. Journal of the American Chemical Society. 1995; 117(47): 11619–11627.
44. Oelschlaeger P, et al. Magnesium-cationic dummy atom molecules enhance representation of DNA polymerase beta in molecular dynamics simulations: Improved accuracy in studies of structural features and mutational effects. Journal of Molecular Biology. 2007; 366(2):687–701. [PubMed: 17174326]
45. Rosta E, Yang W, Hummer G. Calcium Inhibition of Ribonuclease H1 Two-Metal Ion Catalysis. Journal of the American Chemical Society. 2014; 136(8):3137–3144. [PubMed: 24499076]
46. Lopata A, et al. Mutations Decouple Proton Transfer from Phosphate Cleavage in the dUTPase Catalytic Reaction. Acs Catalysis. 2015; 5(6):3225–3237.
47. Klvana M, et al. An abridged transition state model to derive structure, dynamics, and energy components of DNA polymerase beta fidelity. Biochemistry. 2011; 50(32):7023–32. [PubMed: 21739967]
48. Zhang Z, Eloge J, Florian J. Quantum mechanical analysis of nonenzymatic nucleotidyl transfer reactions: kinetic and thermodynamic effects of beta-gamma bridging groups of dNTP substrates. Biochemistry. 2014; 53(25):4180–91. [PubMed: 24901652]
49. Ram Prasad B, Warshel A. Prechemistry versus preorganization in DNA replication fidelity. Proteins. 2011; 79(10):2900–19. [PubMed: 21905114]
50. Matute RA, Yoon H, Warshel A. Exploring the mechanism of DNA polymerases by analyzing the effect of mutations of active site acidic groups in Pol β (submitted). Proteins: Structure, Function, and Bioinformatics. 2016
51. Singh N, Warshel A. Absolute binding free energy calculations: On the accuracy of computational scoring of protein-ligand interactions. Proteins-Structure Function and Bioinformatics. 2010; 78(7):1705–1723.
52. Chakrabarty S, Warshel A. Capturing the energetics of water insertion in biological systems: the water flooding approach. Proteins: Structure, Function, and Bioinformatics. 2013; 81(1):93–106.
53. Klvana M, et al. Catalytic effects of mutations of distant protein residues in human DNA polymerase beta: theory and experiment. Biochemistry. 2012; 51(44):8829–43. [PubMed: 23013478]
54. Frushicheva MP, Warshel A. Towards quantitative computer-aided studies of enzymatic enantioselectivity: the case of Candida antarctica lipase A. Chembiochem. 2012; 13(2):215–23. [PubMed: 22190449]
55. Beutler TC, et al. Avoiding Singularities and Numerical Instabilities in Free-Energy Calculations Based on Molecular Simulations. Chemical Physics Letters. 1994; 222(6):529–539.
56. Yin YW, Steitz TA. The structural mechanism of translocation and helicase activity in T7 RNA polymerase. Cell. 2004; 116(3):393–404. [PubMed: 15016374]

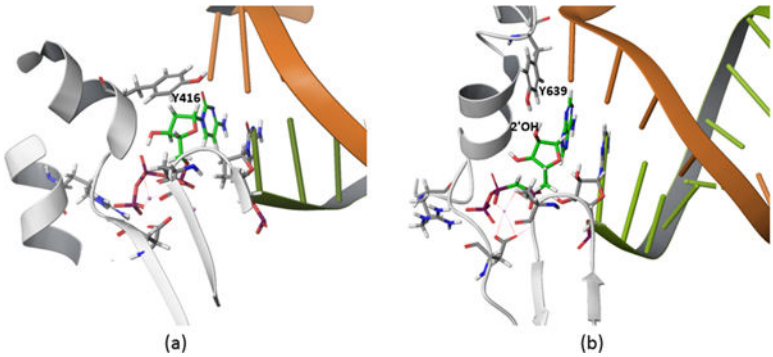

Figure 1.

Analogous active sites of DNA polymerase (a) and RNA polymerase (b). These two polymerase catalyze the same reaction with the same key groups; two aspartate acids and magnesium ions. Furthermore, both have a tyrosine right above the 2′ of the ribose sugar.

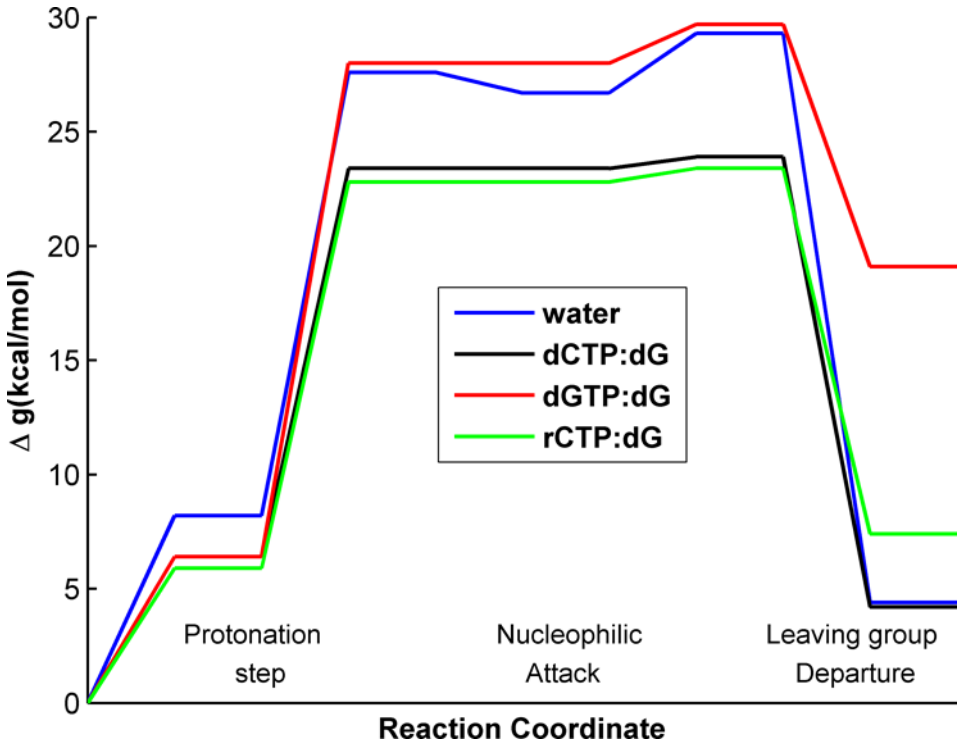

Figure 2.
The EVB free energy profiles for the catalytic reaction with different substrates.

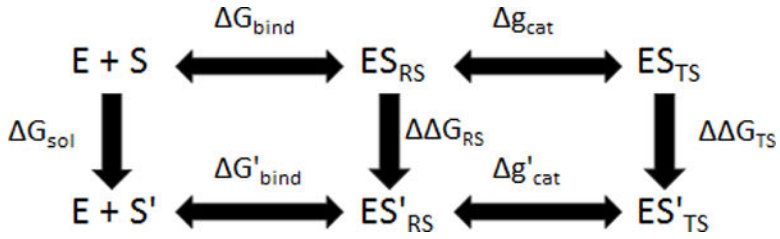

Figure 3.

The thermodynamic cycle for the alchemical transformation of the ribose sugar in NTP. E, S, and S′ represent enzyme, cognate substrate, and non-cognate substrate, respectively. The energetics of the states involved is described schematically in Fig. 4

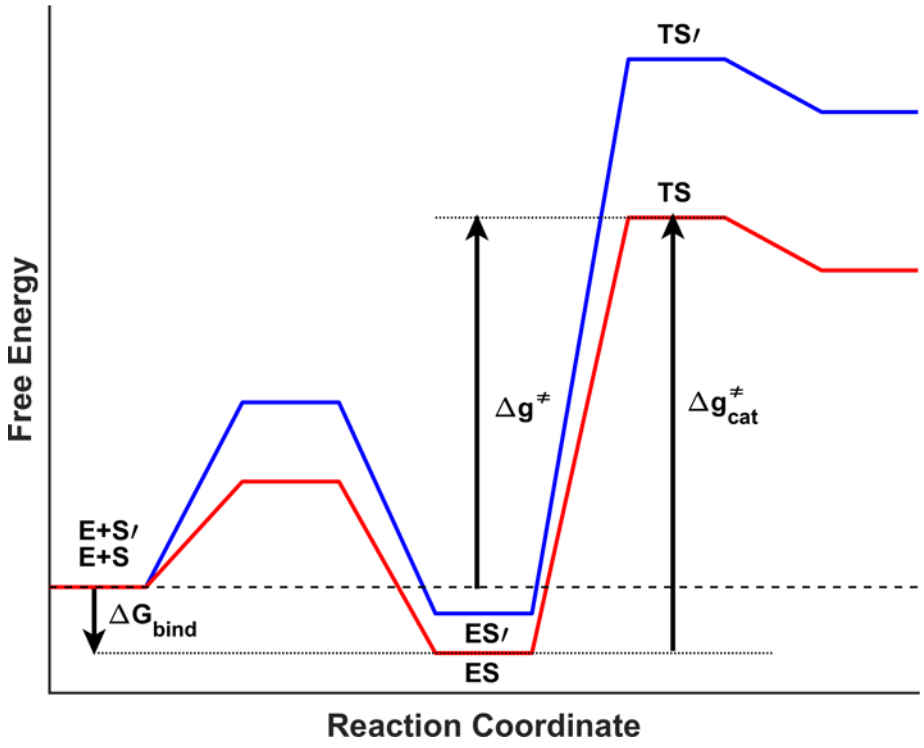

Figure 4.
Schematic description of the energetics of the states considered in the thermodynamic cycle of Fig.3

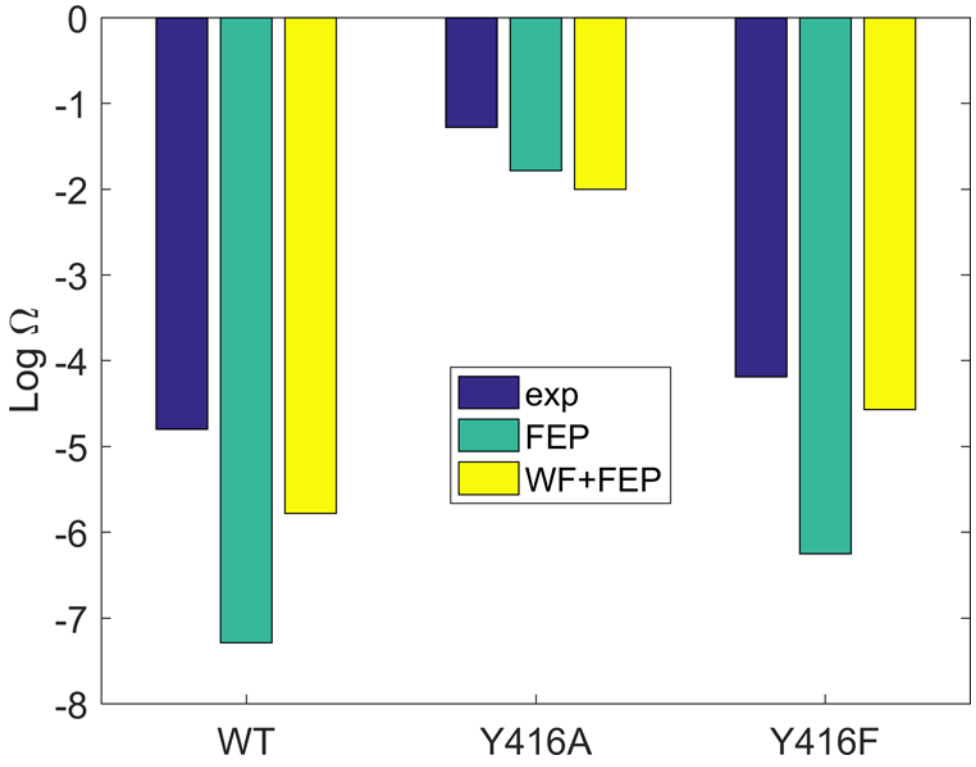

Figure 5.
The relative catalytic efficiency (log Ω) calculated for DNA polymerase

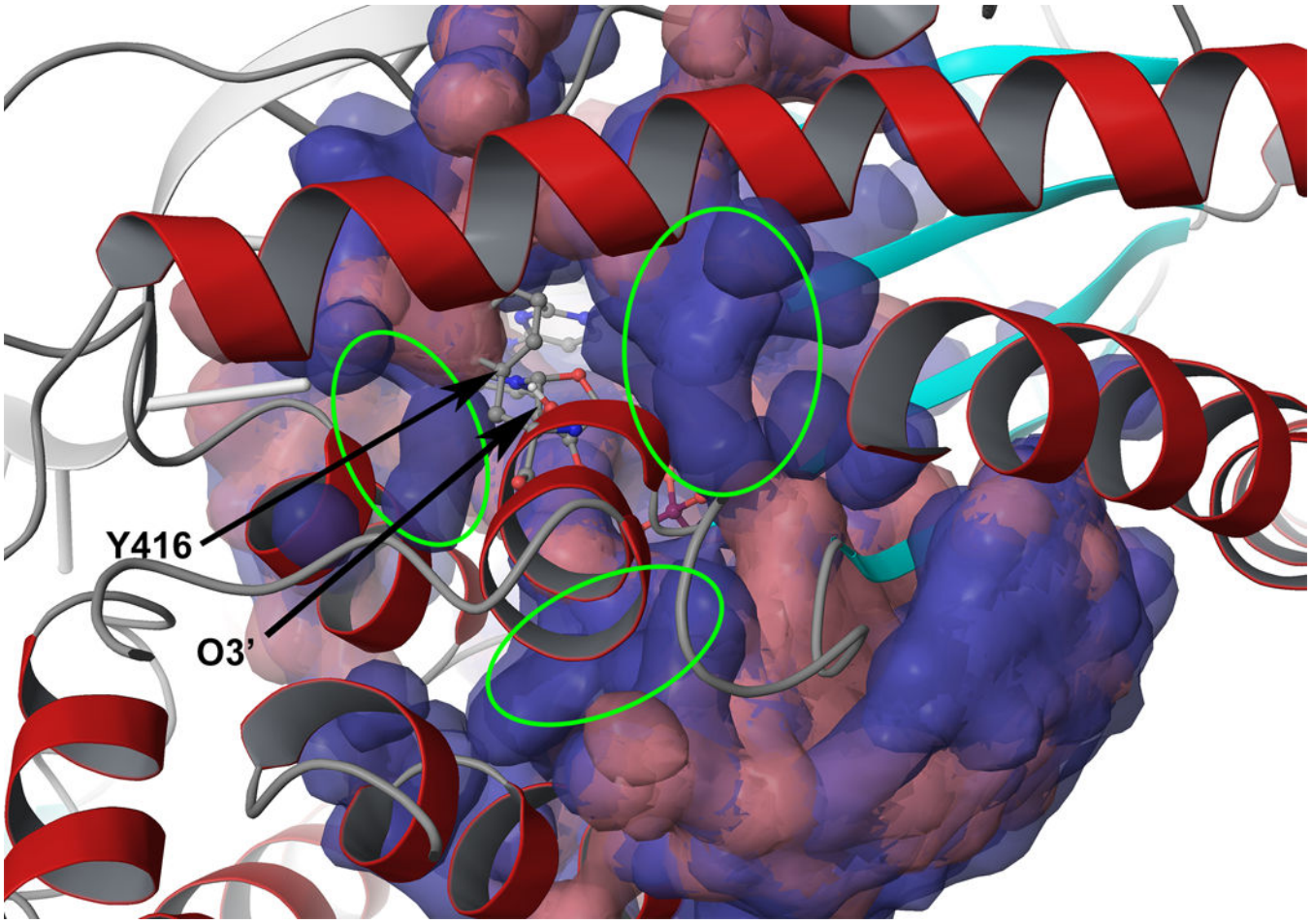

Figure 6.

A comparison between normally solvated structures (red) and structures processed by the water flooding approach (blue). All initial trajectories used were superimposed and the water molecules are presented as molecular surface with transparency. The extra water molecules found around the Y416 after the water flooding are highlighted with circles.

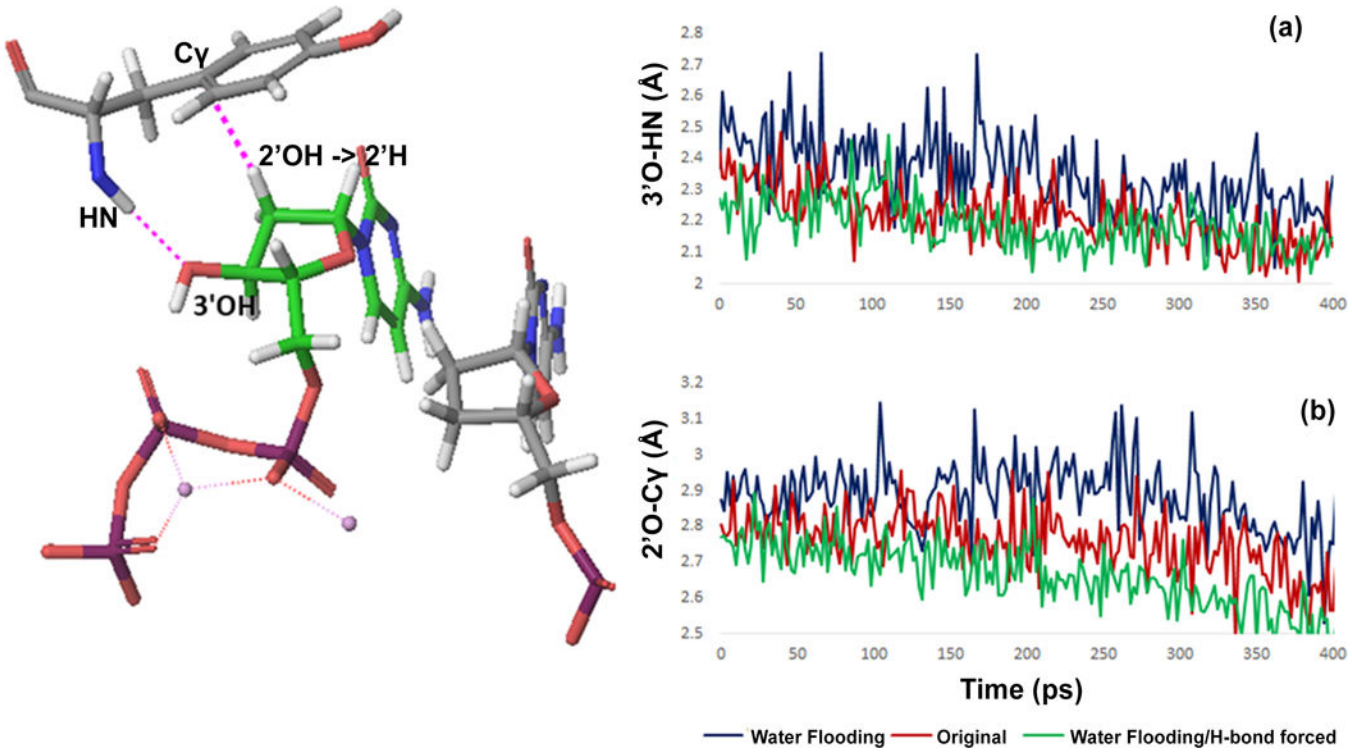

Figure 7.

(a) Hydrogen bond distances between 3′OH and the Y416 backbone during the FEP calculation for structures. **(b)** The distance between 2′O and Cγ of the side chain of Y416 during the FEP calculation. Three cases were tested with the original structure (red), with the structure processed by water flooding (blue), and the same structure by water flooding but the hydrogen bond 3′O and HN is forced by a 4.0 kcal/mol $Å^2$ harmonic constraint at 2 Å (green).

Table 1

Experimental data set from Yang et al. [24] and the catalytic effect of changing from deoxyribose to ribose sugar in RB69 DNA polymerase

	WT				Y416A				Y416F			
	Kd	**kpol**	**kpol/Kd**	**log Ω**	**Kd**	**kpol**	**kpol/Kd**	**log Ω**	**Kd**	**kpol**	**kpol/Kd**	**log Ω**
dCTP	69	2.0×10^{2}	2.9	−4.8	70	22	0.31	−1.3	2.1×10^{2}	3.6×10	0.17	−4.2
rCTP	1.6×10^{4}	0.74	4.6×10^{-5}		5.2×10^{2}	8.6	1.7×10^{-2}		1.8×10^{4}	0.2	1.1×10^{-5}	

Table 2

The relative catalytic efficiency (log Ω) calculated for DNA polymerase with and without the water flooding postprocessing and 4.0 kcal/mol Å^2 hydrogen bond constraint between the backbone nitrogen of the Y416 and 3OH′ in the substrate.

		Without H bond const		With H bond const	
	exp	**Original**	**Water Flooding**	**Original**	**Water Flooding**
WT	−4.80	−7.29	−5.78	−8.48	−7.03
Y416A	−1.28	−1.78	−2.00	−1.24	−1.52
Y416F	−4.19	−6.25	−4.57	−9.32	−6.37

[a]The calculations are based on the alchemical thermodynamic cycle described in Figure 4. The last column includes the effect of using our water flooding approach.

Table 3

Breakdown of the calculated free energy contributions to the TS binding in RB69 DNA polymerase[a]

	Electrostatic	VDW
solvent	9.08	−3.05
WT	4.43	−5.76
Y416A	5.58	−2.40
Y416F	7.29	−6.83

[a]Energies in kcal/mol. The calculations involve the water flooding approach and no H bond constraint.

Table 4

Experimental data set from [25] and the catalytic effect of changing from deoxyribose to ribose sugar in T7 RNA polymerase

	WT				Y639F			
	K_D	**kpol**	**kpol/K_D**	**log Ω**	**KD**	**kpol**	**kpol/K_D**	**log Ω**
rATP	0.1	205.0	4100.0	−3.3	0.1	180.0	1914.9	−0.9
dATP	3.5	7.5	2.1		0.3	60.0	222.2	

Table 5

The relative catalytic efficiency (log Ω) for T7 RNA polymerase structures with and without the constrained hydrogen bond[a].

	exp	normal	H bond forced
Wild Type	−3.3	1.1	−3.2
Y639F	−0.9	−2.3	−1.4

[a] with 15kcal/mol hydrogen bond constraint between O2′ and the hydroxyl group in Y416 at 1.9 Å was imposed in the wild type while a corresponding distance constraint at 2.8 Å between O2′ and Hζ in F416 was imposed in the mutant

Table 6

Free energy contributions of the alchemical mutation in T7 RNA polymerase[a]

	WT/H-bond	Y639F with const	Solution
Electrostatic	10.1	6.8	9.1
VDW	0.3	0.5	−3.0

[a]Energies in kcal/mol. The results were obtained with a hydrogen bond constraint between O2′ and the hydroxyl group. In Y639F, a corresponding constraint was imposed to position the mutated phenylalanine at the constrained wild type.