Role of Histidine-85 in the Catalytic Mechanism of Thymidine Phosphorylase As Assessed by Targeted Molecular Dynamics Simulations and Quantum Mechanical Calculations†

Jesús Mendieta,‡ Sonsoles Martín-Santamaría,‡ Eva-María Priego,§ Jan Balzarini,|| María-José Camarasa,§ María-Jesús Pérez-Pérez,§ and Federico Gago*,‡

Departamento de Farmacología, Universidad de Alcalá, E-28871 Alcalá de Henares, Madrid, Spain, Instituto de Química Médica, CSIC, Juan de la Cierva 3, 28006 Madrid, Spain, and Rega Institute for Medical Research, K. U. Leuven, Minderbroedersstraat 10, B-3000 Leuven, Belgium

Received May 14, 2003; Revised Manuscript Received September 29, 2003

ABSTRACT: The structural changes taking place in the enzyme thymidine phosphorylase (TPase, also known as PD-ECGF) that are required to achieve catalytic competence upon binding thymidine and phosphate have been simulated by means of targeted molecular dynamics (tMD). The hinge regions were characterized by structural homology comparisons with pyrimidine nucleoside phosphorylase, whose X-ray structure has been solved both in a closed and in an open form. The rearrangement of residues around the substrate that was observed during the tMD trajectory suggested that His-85 could be playing an important role in the catalytic mechanism. A quantum mechanical study of the reaction in the presence of the most relevant active site residues was then performed at the semiempirical level. The results revealed that His-85 could be involved in the protonation of the pyrimidine base at the O2 position to yield the enol tautomer of the base. To establish the role of this oxygen atom in the reaction, ground states, transition states, and final products were studied using higher level ab initio methods starting from both thymidine and 2-thiothymidine as alternative substrates. Comparison of both transition states showed that replacing the oxygen at position 2 of the pyrimidine base by sulfur should accelerate the reaction rate. Consistent with this result, 2-thiothymidine was shown to be a better substrate for TPase than the natural substrate, thymidine. For simulating the final step of the reaction, tMD simulations were used to study domain opening upon product formation considering both the enol and keto tautomers of thymine. Product release from the enzyme was easiest in the simulation that incorporated the keto tautomer of thymine, suggesting that the enol intermediate spontaneously tautomerizes back to the more energetically stable keto form. These results highlight a previously unreported role for His-85 in the catalytic mechanism of TPase and can have important implications for the design of novel TPase inhibitors.

Pyrimidine nucleoside phosphorylases are involved in the salvage pathways that replenish pyrimidine nucleotide pools in cells. In lower organisms, there exists only one enzyme (PyNPase) that accepts both thymidine and uridine as substrates. In mammalian cells there are at least two different enzymes: thymidine phosphorylase (TPase, EC 2.4.2.4) and uridine phosphorylase (UPase, EC 2.4.2.3) (*1*). TPase is highly specific for the 2′-deoxyribonucleosides of thymine and related pyrimidine bases whereas UPase does not distinguish between ribose and deoxyribose in pyrimidine nucleosides. TPase catalyzes the reversible phosphorolysis of thymidine (Scheme 1) as follows:

thymidine + phosphate ⇄ 2-deoxyribose 1-phosphate + thymine

Potential inhibitors of this enzyme could be useful since some tumors depend on the nucleotide salvage pathway for their proliferation (*2*). Another aspect that makes this enzyme an interesting therapeutic target is its involvement in the phosphorolysis of different pyrimidine nucleoside analogues that are used as antiviral [e.g., (*E*)-5-(2-bromovinyl)-2′-deoxyuridine and 5-iodo-2′-deoxyuridine] or anticancer (e.g., 5-fluoro-2′-deoxyuridine) agents. As a consequence, TPase inhibitors might enhance the pharmacological effects of these drugs (*3*−*5*). Furthermore, interest in the development of TPase inhibitors has recently grown due to the identification of human TPase activity with the platelet-derived endothelial cell growth factor (PD-ECGF), which is known to stimulate endothelial cell migration in vitro and angiogenesis in vivo (*6*). In fact, the TPase enzymatic activity of PD-ECGF has been shown to be essential for its action as a growth factor (*7*, *8*), and the unusually high levels of PD-ECGF/TPase that

† Financial support from Comunidad de Madrid (Salud 2000 Project 08.1/0039/2000 to M.-J.P.P. and F.G.), the European Commission (QLK2-CT-2000-00291 to J.B., M.-J.C., and F.G. and QLG1-CT-2001-01004 to J.B. and M.-J.P.P.), the National Foundation for Cancer Research (to F.G.), and the Geconcerteerde Onderzoeksacties (Contract 2000/12 to J.B.) is gratefully acknowledged.

* To whom correspondence should be addressed. E-mail: federico.gago@uah.es. Telephone: +34-918 854 514. Fax: +34-918 854 591.

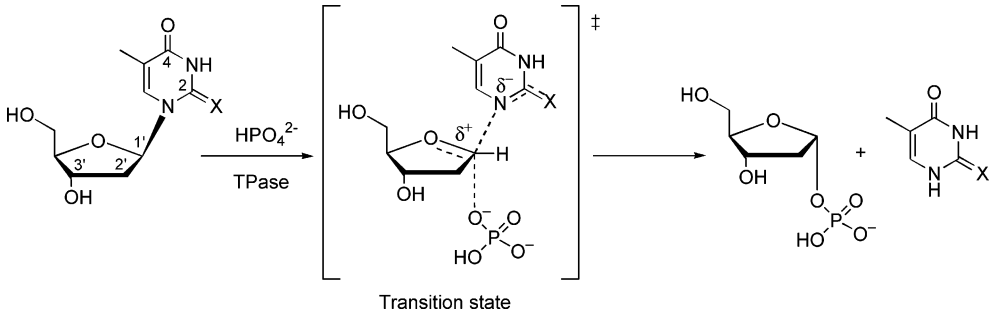
‡ Universidad de Alcalá.

§ Instituto de Química Médica, CSIC.

|| Rega Institute for Medical Research.

10.1021/bi034793o CCC: $27.50
Published on Web 12/20/2003

Scheme 1: Phosphorolysis Reaction of Thymidine (X = O) or 2-Thiothymidine (X = S) Catalyzed by TPase



are found in a variety of human cancers have been related to tumor aggressiveness (*9*−*11*).

Structurally, TPase is a dimer made up of two identical subunits with a dimeric molecular mass ranging from 90 kDa in *Escherichia coli* to 110 kDa in mammals. The fact that human TPase shares 39% of sequence identity with *E. coli* TPase highlights the significant sequence homology that exists among known members of the TPase family. In the X-ray crystal structure of the *E. coli* enzyme, which has been solved at 2.8 Å resolution in a complex with thymine and sulfate (*12*), each subunit appears as a large α/*β* domain and a smaller α domain separated by a cleft. The sulfate ion is bound in the phosphate binding site in the α/*β* domain where it is stabilized by a salt bridge with Lys-84 and several hydrogen bonds with backbone amide nitrogens and hydroxyl groups from residues lining this pocket. The thymine base is bound in the thymidine binding site that is present in the α-domain, and its O4, N3, and O2 atoms are engaged in direct interactions with Arg-171, Ser-186, and Lys-190, respectively. When inorganic phosphate (P_i) and thymidine are modeled into the active site using as templates, respectively, the sulfate ion and the thymine base in this crystal structure to figure out how the forward reaction proceeds, the distance between any anion oxygen and the C1′ of thymidine is higher than 8 Å. This observation has suggested that domain closure is necessary in order to generate a catalytically competent active site in the enzyme (*13*). The lack of experimental information about the structure of the closed form, however, hampers the precise identification of residues that are involved in the enzymatic reaction although some insight has been gained by comparing the structures of *E. coli* TPase in three different crystallographic forms. These structural data, together with molecular modeling results and analogies with members of the purine nucleoside phosphorylase family (*14*), have led to the proposal of a catalytic mechanism (*13*). Nevertheless, the role in this mechanism of His-85, which is absolutely conserved in all TPases and in all PyNPases for which thymidine is a substrate, is presently unknown (*15*).

The domain closing motion that is necessary to generate the active site of TPase is common to many other enzymes and some other proteins presenting two or more domains connected by a few strands of polypeptide chain that can be considered as hinges (*16*). Conformational changes for interconversion between open and closed forms are usually restricted to the hinge region insofar as the domains behave as rigid bodies. The final structure depends on a global displacement of the domains with respect to each other in terms of both translation and rotation.

Standard molecular dynamics (MD) simulations could be useful for modeling the "closed" form of the enzyme starting from an "open" conformation, but the required motions take place at time scales that are longer than those currently achievable by these methods. This is so because the system must surmount the energy barrier that separates both open and closed states, and although the barrier crossing process itself is normally quite fast, the time needed for random thermal fluctuations within the system to produce the local atomic momenta that are necessary for overcoming the barrier may be of the order of milliseconds or longer. In these cases, "targeted" (*17*), "steered" (*18*), or "activated" MD simulations (*19*, *20*) provide an extra energy term that is used to accelerate the crossing of local barriers. This approach has been previously used in conformational studies of several proteins such as the chaperone GroEL (*21*), the glutamate receptor ligand-binding core (*22*), and the tyrosine kinase Src (*23*).

Here, we use a variation of our previously reported MD protocol (*22*) to model the rearrangement of amino acid side chains in the active site of *E. coli* TPase during specified stages of the reaction. To this end, the enzyme active site has been simulated in the presence of both the substrates and the products, as well as the transition states of the reaction, which we have characterized using quantum mechanical methods. In light of the structural information obtained, we propose a catalytic mechanism for this enzyme in which His-85 plays an important role. The feasibility of this proposal has been validated experimentally by comparing the enzymatic reaction rates when thymidine and 2-thiothymidine are used as substrates.

METHODS

Molecular Modeling and Energy Refinement of the Complexes. The atomic coordinates of *E. coli* TPase (*13*) and *Bacilllus stearothermophilus* PyNPase (*24*) were obtained from the Protein Data Bank (*25*) at the Research Collaboratory for Structural Bioinformatics (PDB codes 2TPT and 1BRW, respectively). The structural alignment between the two proteins was performed with DALI (*26*). The natural substrates (thymidine and HPO_4^{2-} or $H_2PO_4^-$) were placed at the same positions that thymine and sulfate occupy in the X-ray crystal structure of TPase (PDB code 1TPT) (*12*). Both tautomers, protonated on either Nδ or Nϵ, as well as the doubly protonated form, were independently considered for His-85. To take into account solvent effects, a sphere of explicit water molecules centered at the Cα of Ser-86 was included in the three simulations. The radius of this water cap was 25 Å to ensure the solvation of the interdomain cleft

containing the substrate binding sites in both the open and the closed conformations.

In order for the crystal structures to adapt to the AMBER force field (*27*), the proteins were gradually energy refined in AMBER using the SANDER module (*28*). Initially, only the water molecules were allowed to move so they could reorientate in the electric field of the complex; then the substrates were also included in the energy minimization, followed by the amino acid side chains. Finally, the whole system was subjected to energy refinement even though protein backbone atoms were restrained to their initial positions with a force constant of 2 kcal mol^{-1} $Å^{-2}$. In all cases, a cutoff of 10.0 Å was used in the evaluation of nonbonded interactions.

Molecular Dynamics Simulations. The refined structures were used as input for the subsequent molecular dynamics (MD) simulations under the same conditions as above. In a 6 ps heating phase, the temperature was raised from 0 to 298 K, and velocities were reassigned at each new temperature according to a Maxwell−Boltzmann distribution. The temperature was then kept constant at 298 K by coupling to an external heat bath. The substrate molecules were completely free to move. The dihedral angles of the Cα trace of both domains were restrained to those of the initial structure during the heating phase by means of a harmonic potential with a force constant of 300 kcal mol^{-1} rad^{-2}. For the rest of the simulation at 298 K, the dihedral restraints involving the hinge regions were removed, and a "template forcing" approach was used to study the motion of the domains. To this end, the Cα atoms of the hinge region in the open structure were forced to adopt the crystallographic positions of equivalent atoms in the closed structure by means of a harmonic potential with a force constant that was progressively increased (0.25 kcal mol^{-1} $Å^{-2}$ per run) during 10 consecutive runs of 60 ps each. The SHAKE algorithm was used to constrain all bonds to their equilibrium values, which allowed us to use an integration time step of 2 fs. The list of nonbonded pairs was updated every 25 steps, and coordinates were saved every 2 ps for further analysis.

To study the process of product release, targeted MD (tMD) simulations in the presence of 2-deoxyribose 1-phosphate and either the enol or the keto form of thymine were performed, including a neutral His-85, protonated at Nδ, in the active site. The template in this case was the open conformation of the enzyme.

Quantum Mechanical Study of the Enzymatic Reaction. To analyze the role in the reaction of the residues that shape the active site of the enzyme, a semiempirical quantum mechanical study of the chemical reaction was undertaken. A simplified representation of the enzyme active site was used, consisting of selected amino acid residues (Lys^+-84, His^+-85, Ser-86, Ser-95, Ser-113, Thr-123, Arg^+-171, Ser-186, and Lys^+-190) and the bound substrate molecules (thymidine and HPO_4^{2-}), with atom positions specified in internal coordinates. The total charge of the system was thus set to +2. The AM1 Hamiltonian, as implemented in the semiempirical quantum chemical program MOPAC 2000 (*29*), with subroutine DRIVER 1.0 (*30*) interfaced to program TRITON 2.0 (*31*), was used for mapping the reaction pathway of the first phosphorolysis step. Attack by the phosphate was modeled by a stepwise reduction (0.05 Å decrements) of the distance between the nucleophilic oxygen from the phosphate anion and the C1′ carbon of thymidine using the following keywords to control the MOPAC calculation: AM1, DEBUG, BFGS, GEO-OK, MMOK, MOZYME, NODIIS, NOINTER, NOXYZ, and CHARGE. After each step, the structure was fully optimized except for the driven coordinate and the heavy atoms of the protein backbone. Since residues from both domains contribute to the binding pocket in TPase and the mobility of the domains appears to be necessary for the catalysis, the backbone atoms of residues belonging to each domain were constrained to move jointly whereas the second domain was allowed to reorientate relative to the first.

Ab Initio Quantum Mechanical Calculations. As a means to assess the importance of position 2 of the pyrimidine ring in the catalytic mechanism (Scheme 1), the reactions of phosphate anion with thymidine and 2-thiothymidine were studied. The relative reactivity of both substrates was compared using the ab initio methods implemented in the Gaussian 98 program suite (*32*). Since the active site environment is the same in both cases, we have focused on the isolated substrates, intermediates, and products in order to make the ab initio calculations more tractable. Ground states, intermediates, and final products (i.e., thymine and 2-thiothymine, as well as their respective enol tautomers) were optimized at the B3LYP/6-31G(d,p) and RHF/6-31G(d,p) levels of theory, both in vacuo and in the presence of a polarizable continuum model (PCM) implicitly representing water. Transition states (TS) for the nucleophilic attack of phosphate anion were located at the RHF/3-21G(d) level of theory using the synchronous transit-guided quasi-Newton (STQN) method. The previously obtained AM1-optimized TS structures were used as initial geometries for the ab initio calculations. Intrinsic reaction coordinate (IRC) calculations were then performed at the RHF/3-21G(d) level of theory from the TS toward both the reactants and the products. Vibrational frequencies were calculated at stationary points to identify them as transition states (one imaginary frequency) or the two neighboring minima (all positive frequencies). Natural bond orbital (NBO) analyses were performed to quantify and compare the role of electron delocalization in each species. The energy of each interaction was calculated as the difference between the total energy and the energy calculated upon removal of the off-diagonal element in the Fock matrix corresponding to the interaction of interest.

TPase Enzyme Assays. Thymidine and thymine were purchased from Sigma (St. Louis, MO). 2-Thiothymidine and 2-thiothymine were obtained, respectively, from Berry and Associates (Dexter, MI) and Aldrich (Germany). *E. coli* TPase was recombinantly expressed and purified in our laboratory (by J. Balzarini). The phosphorolysis of thymidine and 2-thiothymidine by *E. coli* TPase was measured by high-performance liquid chromatography (HPLC) analysis. The incubation mixture (500 μL) contained 10 mM Tris·HCl (pH 7.6), 1 mM EDTA, 2 mM potassium phosphate, 150 mM NaCl, a range of concentrations of thymidine and 2-thiothymidine (250, 375, 500, 750, 1000, 1250, and 1500 μM), and an appropriate amount of TPase. Incubations were performed at room temperature. At 10 and 20 min (linear conversion of reaction products), 100 μL fractions were withdrawn, transferred to an Eppendorf tube thermo block, and heated at 95 °C for 5 min. Thereafter, the samples were rapidly cooled on ice, and thymidine and 2-thiothymidine

```
1brwA    .MRMVDLIAKKRDGKALTKEEIEWIVRGYTNGDIPDYQMSALAMAIYFRGMTEEETAALT
1tpt     lFLAQEIIRKKRDGHALSDEEIRFFINGIRDNTISEGQIAALAMTIFFHDMTMPERVSLT

1brwA    MAMVQSGEMLDLS..SIRGVKVDKHSTGGVGDTTTLVLGPLVASVGVPVAKMSGRGLGHT
1tpt     MAMRDSGTVLDWKslHLNGPIVDKHSTGGVGDVTSLMLGPMVAACGGYIPMISGRGLGHT

1brwA    GGTIDKLESVPGFHVEISKDEFIRLVNENGIAIIGQTGDLTPADKKLYALRDVTATVNSI
1tpt     GGTLDKLESIPGFDIFPDDNRFREIIKDVGVAIIGQTSSLAPADKRFYATRDITATVDSI

1brwA    PLIASSIMSKKIAAGADAIVLDVKTGAGAFMKKLDEARRLARVMVDIGKRVGRRTMAVIS
1tpt     PLITASILAKKLAEGLDALVMDVKVGSGAFMPTYELSEALAEAIVGVANGAGVRTTALLT

1brwA    DMSQPLGYAVGNALEVKEAIETLKG.NGPHDLTELCLTLGSHMVYLAEKAPSLDEARRLL
1tpt     DMNQVLASSAGNAVEVREAVQFLTGeYRNPRLFDVTMALCVEMLISGKLAKDDAEARAKL

1brwA    EEAIRSGAAIAAFKTFLAAQGGDASVVDDLD.KLPKAAYTSTVTAAADGYVAEMAADDIG
1tpt     QAVLDNGKAAEVFGRMVAAQKGPTDFVENYAkYLPTAMLTKAVYADTEGFVSEMDTRALG

1brwA    TAAMWLGAGRAKKEDVIDLAVGIVLHKKIGDRVQKGEALATIHSNRPD.VLDVKEKIEAA
1tpt     MAVVAMGGGRRQASDTIDYSVGFTDMARLGDQVDGQRPLAVIHAKDENnWQEAAKAVKAA

1brwA    IRLSPQPVARPPLIYETIV.
1tpt     IKLADKAPESTPTVYRRISe
```

FIGURE 1: Structural alignment of PyNPase (1brwA) and TPase (1tpt) obtained by DALI (*26*). Structurally equivalent residues are in upper case letters; nonequivalent residues are in lower case letters. Homologous residues are shown in color: green, nonpolar residues; red, positively charged residues; blue, negatively charged residues; magenta, polar (noncharged) residues.

were separated from thymine and 2-thiothymine, respectively, and quantified in the samples on a reverse-phase RP-8 column (Merck, Darmstadt, Germany) by HPLC analysis. The separation of 2-thiothymine, thymine, thymidine, and 2-thiothymidine was performed by a linear gradient from 2% buffer B (10 mM potassium phosphate buffer, pH 5.5, + 80% methanol) in 10 mM potassium phosphate buffer, pH 5.5, to 80% buffer B in 10 mM potassium phosphate buffer, pH 5.5. After injection of the samples, 2% buffer B was given for 10 min before the start of the gradient (5 min from 2% to 80% buffer B). The retention times of thymine and thymidine were 4.2 and 8.5 min, and those of 2-thiothymine and 2-thiothymidine were 4.9 and 9.5 min, respectively. UV-based detection of thymine and thymidine was performed at 254 nm and for 2-thiothymine and 2-thiothymidine at 270 nm. The Lineweaver−Burk plots for determining the K_m and V_{max} values used the data from the 20 min incubation experiment.

RESULTS AND DISCUSSION

Determination of the Hinge Regions of TPase. Structural evidence of the hinge-like interdomain bending motion that is necessary to produce the active conformation of TPase can be gathered from the several degrees of domain closure observed in three different crystal forms of the enzyme (*13*). However, an indispensable requisite for modeling the closed form of *E. coli* TPase, using tMD simulations, is the availability of a structure that can be used as a template for inducing the closing motion. A search for proteins structurally homologous to TPase using the structural alignment program DALI (*26*) pointed to the PyNPase from *B. stearothermophilus*: both proteins were shown to share more than 99% structural homology in their open forms [root-mean-square deviation (rmsd) of Cα atoms 2.0 Å] despite having a sequence homology of only 42% (Figure 1). In fact, only 5 out of 442 Cα atoms were shown to be in nonequivalent positions, and these two enzymes have been recently proposed to belong to the same nucleoside phosphorylase II (NP-II) family (*15*). Fortunately, the crystal structure of PyNPase contains both an open apo form and a complex of the enzyme with both phosphate and the substrate analogue inhibitor pseudouridine in a completely closed conformation (*24*).

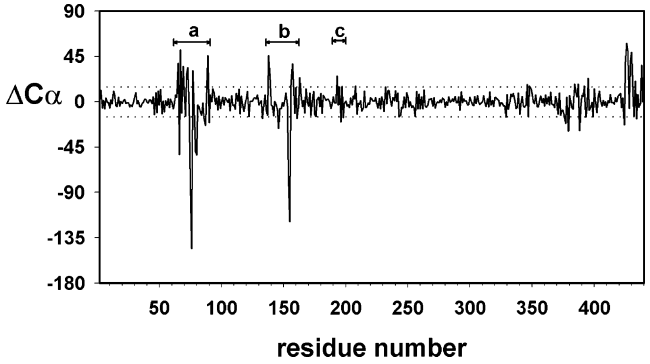


FIGURE 2: Differences in Cα dihedral angles (ΔCα) along the peptide backbone between the open and closed forms of PyNPase. The two enzyme domains are connected by segments a (residues 60−80), b (residues 140−160), and c (residues 193−197), of which only the former two appear to undergo significant conformational changes. The horizontal dashed lines represent the mean of the difference for all residues ± 2σ.

To characterize precisely the regions in PyNPase that are involved in the hinge-like interdomain motion, the differences in dihedral angles between the Cα traces of the polypeptide chains of both open and closed forms were measured. From the results shown in Figure 2 it is clear that the most important conformational changes occur in two regions (residues 60−80 and 140−160, denoted as a and b in Figure 2) corresponding to two of the three peptide stretches that connect the two domains of both PyNPase and TPase. The conformational changes affecting the third domain-connecting segment (denoted c in Figure 2) are almost negligible. Changes detected close to the C-terminus are most likely due to the different packing arrangements of the proteins in the crystal lattice ($P4_32_12$ vs $P12_11$). No significant variation is detected in any other region belonging to either the α or

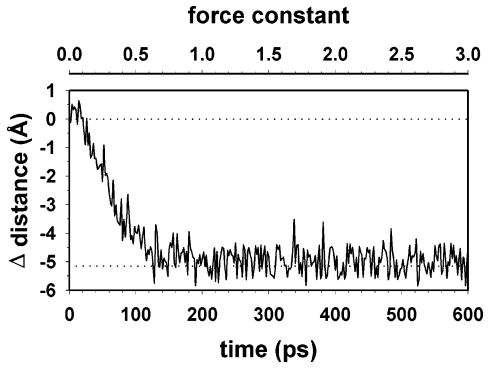


FIGURE 3: Degree of interdomain closure in TPase expressed as a function of the distance between Cα atoms of residues Leu-117 and Arg-171 (each located in a different domain) over the tMD simulation time (lower axis) and applied incremental template force constant (upper axis).

the α/*β* domain. This finding supports earlier interpretations of the crystallographic data for different TPase forms (*13*) and further confirms that the conformational change associated with the catalytic activity of this enzyme can be described as a hinge-like bending motion affecting two rigid-body domains.

Modeling of the Closed Form of TPase by Targeted Molecular Dynamics. To preserve the overall shape of the α and α/*β* domains, the dihedrals of their Cα traces were restrained during the tMD simulation to the values found in the closed form. The hinge-like bending motion was then induced by forcing the Cα trace of the hinge region amino acids (40 atoms) to adopt the conformation it shows in the closed form of PyNPase. The rmsd between open and closed forms for these atoms is only 2.5 Å. We note that the choice of force constant can dramatically affect the results of the simulation: a very large force constant could arbitrarily superimpose the moving stretches to the template, overcoming artifactual energy barriers and depriving the simulation of any physical sense; on the contrary, a small force constant may not supply sufficient extra energy to overcome some or all of the local energy barriers. To circumvent this problem, our approach was to progressively increase the force constant along the simulation time, as previously reported for other systems (*22*). Since domain-closing motions are activated by the extra energy supplied to the hinge regions, the time scale is shorter that in the real situation and is amenable to simulation methods. By monitoring the energy, we get an initial overall value of about 40 kcal mol^{-1} being "pumped" into the system, a value that is lower than 2RT per hinge atom.

Different simulations were performed, each with a different protonation state for both His-85 and the phosphate anion. Figure 3 shows how the open form of TPase evolves toward the closed form along the simulation time for a system containing HPO_4^{2-} and a protonated His-85 residue. The steep slope and the short time are indicative of the low-energy barrier that separates open and closed forms. This hinge-based closing motion was monitored as the decrease of the distance between Cα atoms corresponding to residues Leu-117 and Arg-171, which are placed in different domains. The results were very similar in the simulations of the alternatively protonated systems (data not shown). No significant differences were found among the Cα traces of average structures obtained from the last 300 ps of different simulations (rmsd <1.5 Å). Schematic representations of the initial open structure and the final closed form of TPase are shown in Figure 4.

The rearrangement in the active site of amino acid side chains around the substrates was dependent on the protonation state of both His-85 and the phosphate although this did not affect the interactions between P_i and the residues that make up the phosphate-binding pocket. Most of the interactions between these amino acids and P_i in the open form were maintained in the closed form regardless of whether HPO_4^{2-} or $H_2PO_4^-$ was used in the modeling. However, when HPO_4^{2-} and doubly protonated His-85 were included in the system, the Nδ proton of His-85 hydrogen bonded to the phosphate and the Nϵ proton approached the thymidine O2, which lost the interaction with Lys-190 that is found both in the X-ray crystal structure and in the initial open conformation. This Lys-190, in turn, formed a salt bridge with Asp-164. On the contrary, the hydrogen-bonding interactions of thymidine O4 and N3 with Arg-171 and Ser-186, respectively, as seen in the open form, were also maintained in the closed conformation (substrates, Figure 5). A similar arrangement was obtained when $H_2PO_4^-$ and neutral His-85, singly protonated at Nϵ, were included in the system but in this case the biphosphate acted as the proton donor to Nδ. When the opposite case was simulated, i.e., when the neutral His-85 was protonated at Nδ, no direct interaction between this residue and the thymidine O2 was observed (data not shown). Since the protonation state of His residues depends on the neighboring groups, the most probable arrangement of side chain residues lining the active site is that found in the presence of either HPO_4^{2-} and a positively charged His-85 or $H_2PO_4^-$ and a neutral His residue singly protonated at Nϵ.

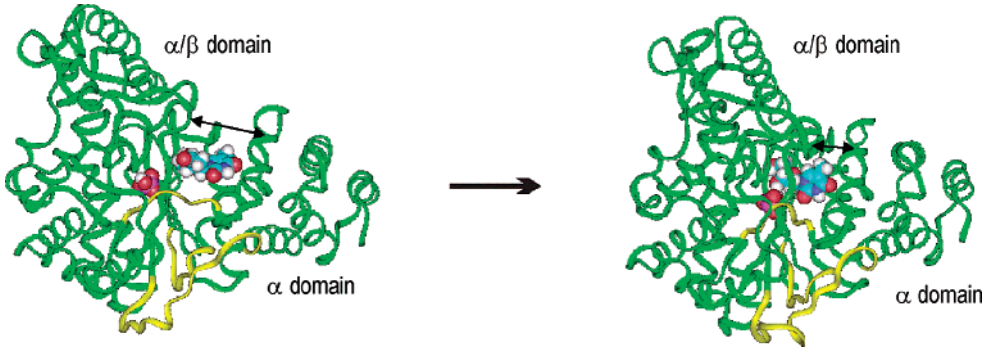


FIGURE 4: Schematic ribbon representations of TPase in the initial open structure (left) and the final closed form (right) obtained after targeted molecular dynamics. The hinge region is colored in yellow, and the Cα atoms of reference residues Leu-117 and Arg-171 are joined by a double-headed arrow.

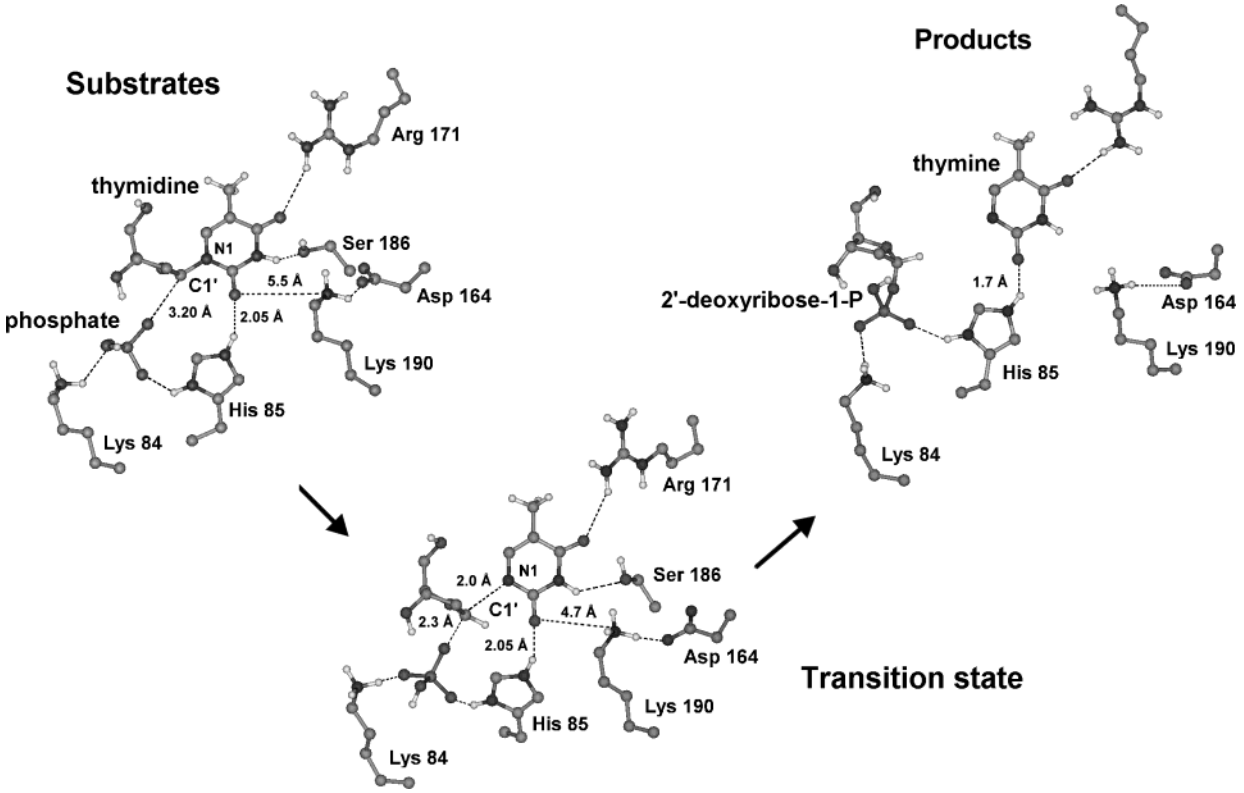


FIGURE 5: Representation of the structures of substrates, transition state, and products within the TPase active site as obtained from semiempirical calculations. The overall active site cavity corresponds in the three cases to the closed form of the enzyme. For clarity, only relevant hydrogens are shown.

Characterization of the Transition State (TS). An accurate characterization of the TS of the phosphorolysis reaction that TPase catalyzes was achieved using a STQN method. Starting with the AM1-optimized structures of reactants, products, and TS as guesses, the STQN calculation at the RHF/3-21G(d) level yielded one TS for thymidine with one imaginary frequency of 430.0*i* cm^{-1}, corresponding to the breaking/formation of the C1′−N1 bond. A further characterization of the TS features will be provided below.

When the TS was docked into the active site of the closed conformation of the enzyme and the complex was simulated without any restraints in the hinge region for 1 ns, the two domains remained close to each other. The arrangement of amino acid side chains in the active site for TS stabilization was similar to that found in the presence of the substrates. Further adaptation of the enzyme to accommodate the TS structure appears to require only minor backbone adjustments in the hinge region.

Theoretical Study of the Enzymatic Reaction within the Protein Active Site. Bond breaking and bond formation cannot be adequately studied with molecular mechanics tools alone. To analyze the role of the residues that make up the active site in catalysis, a semiempirical quantum mechanical study of the chemical reaction was undertaken in the presence of the most important protein atoms. Figure 5 shows the structures of substrates, TS, and products within a reduced representation of the TPase active site as obtained from these calculations in which a positively charged imidazole ring for His-85 and a phosphate dianion were considered. Interestingly, the required protonation on N1 of the leaving thymine was not observed along the reaction coordinate. This is not surprising as no proton-donor residues are placed close

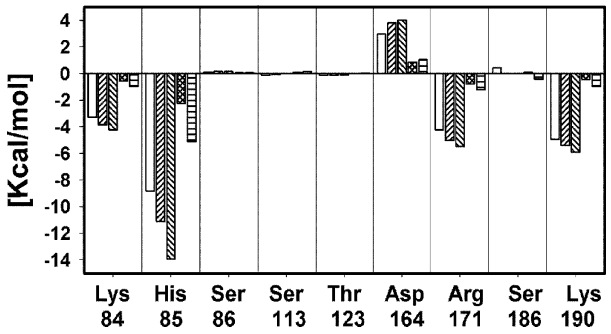


FIGURE 6: Electrostatic interaction energy between the enzyme residues lining the TPase active site and O2 of the pyrimidine base, as calculated for substrates (□), transition state (▨), products (▧), [transition state] − [substrates] (crosshatched bars), and [products] − [substrates] (bars with horizontal lines).

to this atom at the end of the reaction, which suggests an alternative and indirect way for protonation. In fact, the negative charge on the O2 atom of thymine increases during the reaction (data not shown), and the protein residue that contributes most to its stabilization in the three complexes (enzyme−substrates, enzyme−TS, and enzyme−products) is the doubly protonated His-85 (Figure 6). These results suggest that His-85 can be involved in the transfer of the proton from the phosphate to the O2 thymine and also that the enol form of the free base is the most likely intermediate product. This view is in agreement with previous suggestions that His-85 could be potentially involved in stabilizing the TS but appears to rule out a possible role in donation of a proton to the N-1 position of the pyrimidine ring following glycosidic cleavage (*15*). In fact, the MD simulation of the TPase−TS complex reported above consistently showed the Hϵ−O2 distance (1.99 ± 0.05 Å) to be substantially shorter

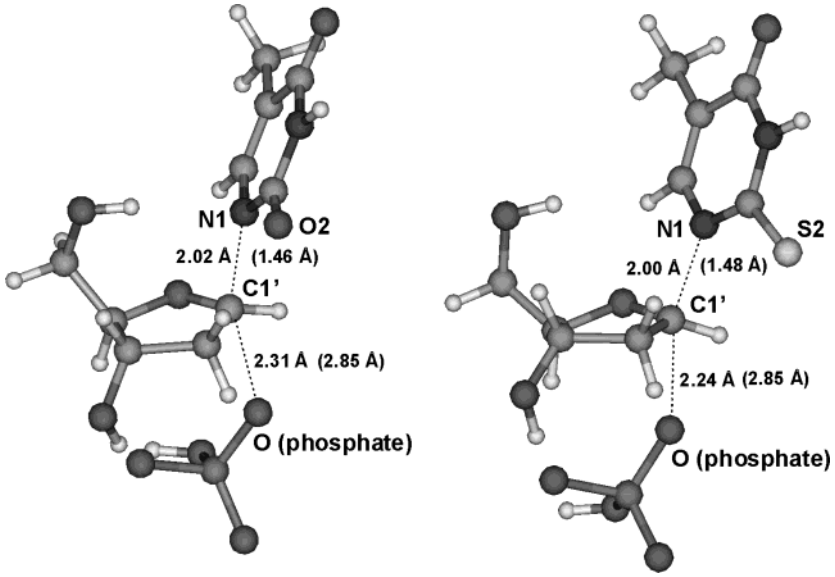


FIGURE 7: Geometry and bond distances for the transition states (left, thymidine; right, 2-thiothymidine) as calculated at the RHF/3-21G(d) level of theory. For comparison, equivalent distances in the ground state are shown in parentheses.

Table 1: Activation Enthalpies (kcal mol^{-1}) and Selected Bond Lengths and Dihedral Angles for the Reactant Nucleosides (Thd, Thymidine; SThd, 2-Thiothymidine), Transition States (TS), and the 2-Deoxyribose 1-Phosphate (2drP) End Product Calculated at the RHF/3-21G(d) Level of Theory

	X = O	X = S
activation enthalpy	22.25	20.20
free energy of reaction	−50.28	−59.18

	X = O			X = S	
	Thd	TS	2drP[a]	SThd	TS
bond length					
C1′−N1	1.46	2.02		1.48	2.00
C1′−O (phosphate)		2.31	1.39		2.24
dihedral angle					
C4′−O−C1′−C2′	−1.79	−2.61	−43.81	−1.99	−19.16
O−C1′-C2′−C3′	−21.13	−1.58	28.23	−21.59	4.88
C1′−C2′−C3′−C4′	34.65	4.76	−4.39	35.54	10.33
C2′−C3′−C4′−O	−36.14	−6.21	−20.85	−37.17	−20.85
C3′−C4′−O−C1′	24.18	5.67	41.30	24.95	25.33

[a] Fully optimized geometry of this molecule at the same level of theory.

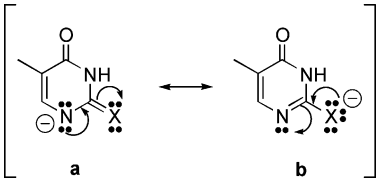
than Hϵ−N1 (4.90 ± 0.10 Å) and the Nϵ−Hϵ−O2 angle (142.1 ± 5.1°) to be more suitable than Nϵ−Hϵ−N1 (98.0 ± 6.8°) for proton abstraction from the histidine.

Role of the O2 Position of Thymine in the Interaction with His-85. These results led us to think that a change in the nature of the atom bonded to C2 of the thymine ring, which according to our calculations is involved in a crucial interaction with His-85, should affect the catalytic rate of TPase. We then compared the transition states upon nucleophile attack of phosphate dianion on either thymidine or 2-thiothymidine (Scheme 1). The calculated TS when X = S also yielded just one imaginary frequency (430.3*i* cm^{-1}) corresponding to the breaking/formation of the C1′−N1 bond. Activation enthalpies for both reactions and selected geometrical parameters of reactants, TS, and the 2-deoxyribose phosphate end product are listed in Table 1. IRC calculations from the TS toward both the products side (resulting in 2-deoxyribose 1-phosphate plus either thymine anion or 2-thiothymine anion) and the reactants side (giving rise to phosphate anion and either thymidine or 2-thiothymidine) showed the products to be lower in energy than the reactants.

Both TS structures are characterized by long attack (phosphate O···C1′) and leaving (C1′···N1) distances (see Figure 7 and Table 1), and positive charge development at C1′ (data not shown), which is compatible with formation of an oxycarbocation and supports the S_N1 mechanism proposed for the phosphorolysis reaction (*13*). Examination of the dihedral angles shows the change in the ribose puckering along the reaction, from thymidine (or 2-thiothymidine) toward free 2-deoxyribose 1-phosphate. For X = O, the ribose ring at the TS is nearly planar, while for X = S, dihedral angle values are much closer to those of 2-deoxyribose 1-phosphate, revealing a geometry more similar to that of this final product (Figure 7). The energetic content of the transition states also varies as a function of X. The calculated activation enthalpy for phosphorolysis of 2-thiothymidine was 2 kcal mol^{-1} lower than that calculated for thymidine, and the exothermicity of the reaction was greater for X = S than for X = O (Table 1). These results are suggestive of a higher reaction rate for TPase when 2-thiothymidine is used as a substrate in place of the natural thymidine.

To determine the factors that would differentiate thymidine and 2-thiothymidine as TPase substrates, we decided to study the putative leaving groups for both S_N reactions by performing an NBO analysis. Solvation models showed greater consistency among the different levels of theory in comparison to the NBO analysis in a vacuum (Table 2). The NBO analysis revealed that the Lewis structure preferred for X = O is that corresponding to resonance structure **a**, in which O2 is double bonded to C2 and N1 is bearing the negative charge of the anion, with strong electron delocalizations due to $n_{N1} \rightarrow \pi^*_{C2-O2}$ interactions. However, for X = S the preferred resonance structure is **b**, in which N1 is double bonded to C2, and the negative charge of the anion is borne by the sulfur which, due to its larger size and its diffuse orbitals, is able to accommodate electrons better than nitrogen and oxygen can. In this case, the negative charge

Table 2: Selected Contributions of Delocalization Interactions to the Stabilization Energies of the Leaving Group Generated upon Phosphate Attack to Thymidine (X = O) and 2-Thiothymidine (X = S)[a]



	X = O			X = S		
method[b]	donor orbital[c]	acceptor orbital[d]	stabilization energy (kcal mol^{-1})	donor orbital[c]	acceptor orbital[d]	stabilization energy (kcal mol^{-1})
1	n_{N1}	π^*_{C2-O2}	128.3	n_{S2}	π^*_{N1-C2}	65.5
	n_{N1}	π^*_{C5-C6}	132.6			
2	n_{N1}	π^*_{C2-O2}	92.4	n_{S2}	π^*_{N1-C2}	46.7
	n_{N1}	π^*_{C5-C6}	81.3			
3	n_{N1}	π^*_{C2-O2}	149.6	n_{S2}	π^*_{N1-C2}	55.0
	n_{N1}	π^*_{C5-C6}	106.1			
4	n_{N1}	π^*_{C2-O2}	103.9	n_{S2}	π^*_{N1-C2}	40.9

[a] As calculated in the NBO analysis on the leaving groups (resonance structures **a** and **b** shown) generated from the transition states depicted in Figure 7. [b] Energy-based geometry optimizations and NBO analyses were carried out at the RHF/6-31G(d,p) level (method 1), the B3LYP/6-31G(d,p) level (method 2), the RHF/6-31G(d,p) level and PCM solvation model (method 3), and the B3LYP/6-31G(d,p) level and PCM solvation model (method 4). [c] n = nonbonding orbital (lone pair). [d] π^* = antibonding orbital.

is mainly stabilized by a $n_{S2} \rightarrow \pi^*_{N1-C2}$ interaction. The better leaving group character of **b** in the case of X = S compared to that of **a** in the case of X = O is in agreement with the lower calculated activation enthalpy for the reaction involving 2-thiothymidine. Besides, the greater accumulation of electron density on S2 compared to O2 should favor protonation of the sulfur by abstraction of Hϵ from His-85, as outlined above.

Enzymatic Studies with Thymidine and 2-Thiothymidine. To get additional, albeit indirect, support for the reaction mechanism outlined above, we decided to study the suitability of 2-thiothymidine as a substrate for *E. coli* TPase and compare the results with those obtained for the natural substrate. Figure 8 shows the double reciprocal plot of the enzymatic reaction using either thymidine or 2-thiothymidine as substrates, under the same experimental conditions. The K_M value obtained using thymidine as the substrate (0.801 mM) was about half the value obtained for 2-thiothymidine (1.587 mM), which suggests a greater affinity of the enzyme for the natural substrate. However, the almost 3-fold difference in V_{max} values (24.2 nmol mL^{-1} min^{-1} for thymidine vs 65.4 nmol mL^{-1} min^{-1} for 2-thiothymidine, obtained from their respective 485 and 1308 μM/20 min values) supports our contention that the glycosidic bond of 2-thiothymidine is more easily cleaved by TPase than that of thymidine.

Product Release. In a mechanism such as the one proposed here, the reaction intermediates after phosphorolysis of the nucleoside would be 2-deoxyribose 1-phosphate and the enol form of the free base, which must subsequently tautomerize to the more stable keto form. Unrestrained MD simulations in the presence of either the enol or the keto form of thymine

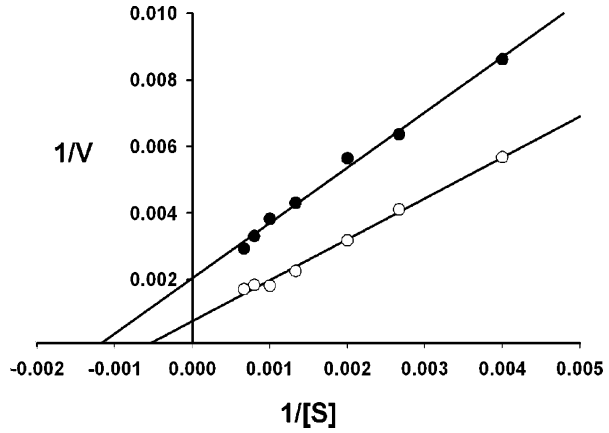


FIGURE 8: Double reciprocal plot from assays of TPase activity using either thymidine (●) or 2-thiothymidine (○) as substrates. Substrate concentrations are in μM, and V are in μM/20 min.

were undertaken to analyze the stability of the different complexes. The need for activation of the opening process became patent after 1 ns of dynamics as the domains remained in a closed conformation (data not shown). When tMD simulations were used to accelerate the process, the energy barrier was surmounted relatively easily (Figure 9). Since the force constant used for the activation was gradually increased, the time needed for the opening to take place was directly proportional to the amount of energy that was pumped into the system and was used as an indirect measure of complex stability. This approach has recently been shown in our hands to be useful for understanding the conformational changes of the glutamate receptor ligand-binding core upon binding of two different agonists (*22*).

Domain separation was monitored in this case by the increase of the distance between the same two Cα atoms from different domains that were previously used to monitor domain closure. Control simulations were also run in the presence of either the substrates or the TS under the same conditions. Opening of the active site in the presence of bound substrates took place at a low force constant value, in fact very similar to that needed for simulating the closing motion starting from the open conformation (Figure 9). This result, which can be expected for an enzyme that catalyzes a reversible reaction (Scheme 1), suggests that no high-energy barrier exists between these two enzyme conformers. The energy required to separate the two domains increased the most when the TS was bound in the active site (Figure 9a), as would be expected for a protein that is optimized for catalyzing a chemical reaction by means of a preorganized polar environment in the active site that provides more electrostatic stabilization to the charges of the TS than water does (*33*). The closed form of the enzyme was also relatively stable in the presence of the enol tautomer of the leaving thymine (Figure 9b). An important contribution to this stabilization stems from the hydrogen bond formed between the OH in position 2 of thymine and the unprotonated Nϵ of His-85. Finally, domain opening was the easiest for the simulation that incorporated the keto tautomer of thymine, in which case the energy-dependent structural profile virtually overlapped that obtained for the substrate−enzyme complex (Figure 9c).

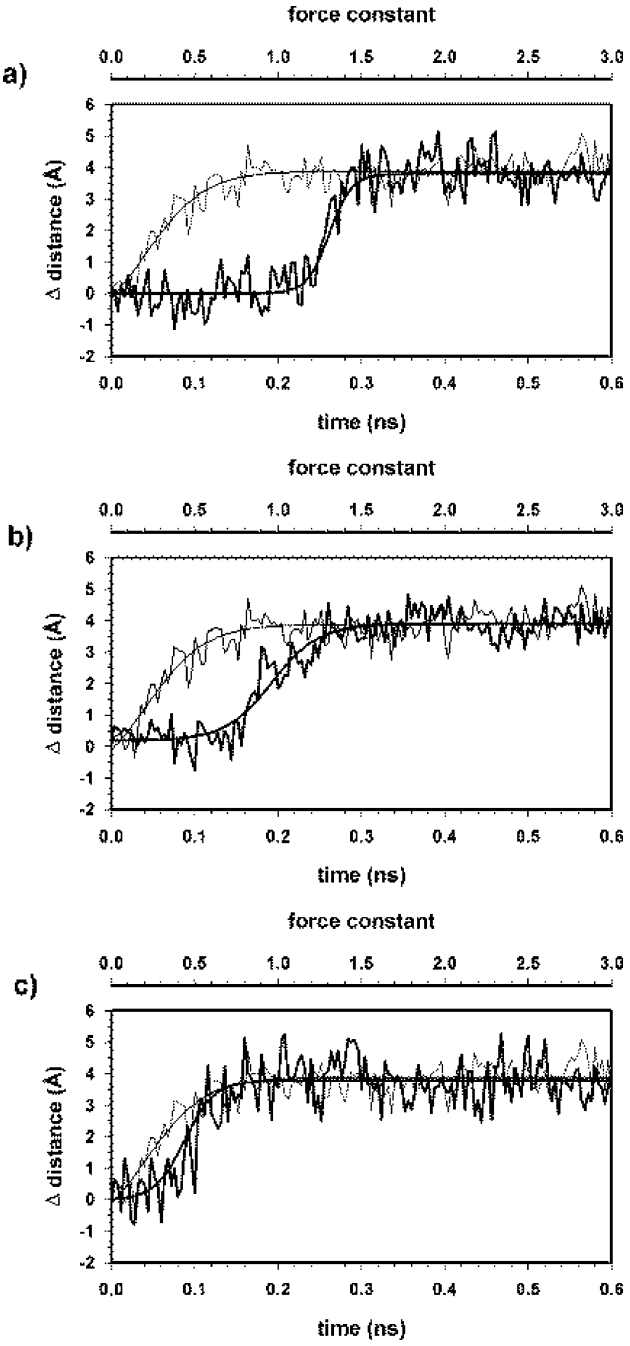


FIGURE 9: Degree of interdomain opening in TPase with bound substrates (thin line) and (a) TS, (b) products with the enol form of thymine, and (c) products with the keto form of thymine (thick lines), expressed as the increment in the distance between Cα atoms of residues Leu-117 and Arg-171 over the tMD simulation time (lower axis). The upper axis shows the value of the applied incremental template force constant.

CONCLUSIONS

TPase must undergo important conformational changes during catalysis. As is observed for other proteins, switching motions for interconversion between open and closed forms are usually restricted to the hinge regions insofar as each domain behaves as a rigid body. Taking the open and closed forms of PyNPase as templates, the hinge regions in TPase were accurately identified and the closed form of this enzyme was modeled using tMD simulations. The resulting rearrangement of the catalytic site showed a relevant interaction between the His-85 imidazole ring and the O2 of thymidine, suggesting that this highly conserved residue could be playing an important role in the catalytic mechanism of TPase. Since the required protonation on N1 of the leaving thymine was not observed along the reaction coordinate within the enzyme active site, an indirect alternative mechanism was proposed and studied.

Our results support an S_N1 mechanism for the phosphorolysis reaction, i.e., a two-step mechanism. In the first step, usually slower and rate limiting, the O2 of thymidine interacts with a proton from the imidazole ring of a protonated histidine residue, which makes the pyrimidine base a better leaving group. The weakening of the glycosidic bond generates an intermediate oxycarbocation in the sugar. The second step is a fast reaction between this oxycarbocation and the phosphate dianion acting as a nucleophile that gives rise to 2-deoxyribose 1-phosphate and the enol form of thymine. In this proposed mechanism, the highly conserved His-85 plays a crucial role: by deprotonating the incoming phosphate, the two catalytically active species (the phosphate dianion and the protonated His) are generated. Besides, it is this positively charged histidine that donates a proton to the O2 of thymidine whereas the phosphate dianion contributes to TS stabilization and acts as the nucleophile in the second step. The enol form of the base would thus be the most likely intermediate product of the reaction, which would then spontaneously tautomerize to the more stable keto form.

Ab initio calculations on the conversion of thymidine and 2-thiothymidine into thymine and 2-thiothymine, respectively, allowed us to predict a higher reaction rate for the sulfur-containing nucleoside. This proposal is in agreement with the higher reaction rate that is observed for TPase when 2-thiothymidine is used as a substrate compared with thymidine, since S2 protonation by His-85 should be facilitated over protonation of O2 in thymidine. The implications of these results for the design of novel TPase inhibitors are currently being evaluated.

ACKNOWLEDGMENT

We thank the University of Alcalá Computing Centre and the CIEMAT for generous allowances of computer time on their SGI R8000 Power Challenge and Cray J916 servers, respectively.

REFERENCES

1. Krenitsky, T. A., Mellors, J. W., and Barclay, R. K. (1965) Pyrimidine nucleosidases: their classification and relationship to uric acid ribonucleoside phosphorylase, *J. Biol. Chem. 240*, 1281−1286.
2. Schwartz, P. M., and Milstone, L. M. (1988) Thymidine phosphorylase in human epidermal keratinocytes, *Biochem. Pharmacol. 15*, 353−355.
3. Desgranges, C., Razaka, G., Rabaud, M., Bricaeud, H., Balzarini, J., and de Clercq, E. (1983) Catabolism of thymidine in human blood platelets, *Biochem. Pharmacol. 32*, 3583−3590.
4. Woodman, P. W., Sarrif, A. M., and Heidelberger, C. (1980) Specificity of pyrimidine nucleoside phosphorylase and the phosphorolysis of 5-fluoro-2′-deoxyuridine, *Cancer Res. 40*, 507−511.
5. Schwartz, E. L., Baptiste, N., Megati, S., Wadler, S., and Otter, B. A. (1995) 5-Ethoxy-2′-deoxyuridine, a novel substrate for thymidine phosphorylase, potentiates the antitumor activity of 5-fluorouracil when used in combination with interferon, an inducer of thymidine phosphorylase expression, *Cancer Res. 55*, 3543−3550.

6. Griffiths, L., and Stratford, I. J. (1997) Platelet-derived endothelial cell growth factor thymidine phosphorylase in tumor growth and response to therapy, *Br. J. Cancer 76*, 689−693.
7. Miyadera, K., Sumizawa, T., Haraguchi, M., Yoshida, H., Konstanty, W., Yamada, and Akiyama, S. (1995) Role of thymidine phosphorylase activity in angiogenic effect of platelet derived endothelial cell growth factor/thymidine phosphorylase, *Cancer Res. 55*, 1687−1690.
8. Moghaddam, A., Zhang, H. T., Fan, T. P., Hu, D. E., Lees, V. C., Turley, H., Fox, S. B., Gatter, K. C., Harris, A. L., and Bicknell, R. (1995) Thymidine phosphorylase is angiogenic and promotes tumor growth, *Proc. Natl. Acad. Sci. U.S.A. 92*, 998−1002.
9. O'Brien, T. S., Fox, S. B., Dickinson, A. J., Turley, H., Westwood, M., Moghaddam, A., Gatter, K. C., Bicknell, R., and Harris, A. L. (1996) Expression of angiogenic factor thymidine phosphorylase/platelet-derived endothelial cell growth factor in primary bladder cancers, *Cancer Res. 56*, 4799−4804.
10. Igarashi, M., Dhar, D. K., Kubota, H., Yamamoto, A., El-Assal, O., and Nagasue, N. (1998) The prognostic significance of microvessel density and thymidine phosphorylase expression in squamous cell carcinoma of the esophagus, *Cancer 82*, 1225−1232.
11. Koukourakis, M. I., Giatromanolaki, A., O'Byrne, K. J., Comley, M., Whitehouse, R. M., Talbot, D. C., Gatter, K. C., and Harris, A. L. (1997) Platelet-derived endothelial cell growth factor expression correlates with tumor angiogenesis and prognosis in non-small-cell lung cancer, *Br. J. Cancer 75*, 477−481.
12. Walter, M. R., Cook, W. J., Cole, L. B., Short, S. A., Koszalka, G. W., Krenitsky, T. A. and Ealick, S. E. (1990) Three-dimensinal structure of thymidine phosphorylase from *Escherichia coli* at 2.8 Å resolution, *J. Biol. Chem. 265*, 14016−14022.
13. Pugmire, M. J., Cook, W. J., Jasanoff, A., Walter, M. R., and Ealick, S. E. (1998) Structural and theoretical studies suggest domain movement produces an active conformation of thymidine phosphorylase, *J. Mol. Biol. 281*, 285−299.
14. Erion, M. D., Stoeckler, J. D., Guida, W. C., Walter, R. L., and Ealick, S. E. (1997) Purine nucleoside phosphorylase. 2. Catalytic mechanism, *Biochemistry 36*, 11735−11748.
15. Pugmire, M. J., and Ealick, S. E. (2002) Structural analyses reveal two distinct families of nucleoside phosphorylases, *Biochem. J. 361*, 1−25.
16. Gerstien, M., Lesk, A. M., and Chothia, C. (1994) Structural mechanism for domain movements in proteins, *Biochemistry 33*, 6739−6749.
17. McCammon, J. A., and Karplus, M. (1979) Dynamics of activated processes in globular proteins, *Proc. Natl. Acad. Sci. U.S.A. 76*, 3585−3592.
18. Isralewitz, B., Gao, M., and Schulten, K. (2001) Steered molecular dynamics and mechanical functions of proteins, *Curr. Opin. Struct. Biol. 11*, 224−230.
19. Collins, J. R., Burt, S. K., and Erickson, J. W. (1995) Flap opening in HIV-1 protease simulated by "activated" molecular dynamics, *Nat. Struct. Biol. 2*, 334−338.
20. Ma, J., and Karplus, M. (1997) Molecular switch in signal transduction: reaction paths of the conformational changes in ras p21, *Proc. Natl. Acad. Sci. U.S.A. 94*, 11905−11910.
21. Ma, J. P., Sigler, P. B., Xu, Z. H., and Karplus M. A. (2000) A dynamic model for the allosteric mechanism of GroEl, *J. Mol. Biol. 302*, 303−313.
22. Mendieta, J., Ramírez, G., and Gago, F. (2001) Molecular dynamics simulations of the conformational changes of the glutamate receptor ligand-binding core in the presence of glutamate and kainate, *Proteins 44*, 460−469.
23. Young, M. A., Gonfloni, S., Superti-Furga, G., Roux, B., and Kurigan, J. (2001) Dynamic coupling between SH2 and SH3 domains of c-Src and Hck underlies their inactivation by C-terminal tyrosine phosphorylation, *Cell 105*, 115−126.
24. Pugmire, M. J., and Ealick, S. E. (1998) The crystal structure of pyrimidine nucleoside phosphorylase in a closed conformation, *Structure 6*, 1467−1479.
25. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. (2000) The Protein Data Bank, *Nucleic Acids Res. 28*, 235−242.
26. Holm, L., and Sander, C. (1996) Mapping the protein universe, *Science 273*, 595−560.
27. Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M. Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc. 117*, 5179−5197.
28. AMBER (UCSF) (1995) Assisted Model Building with Energy Refinement, version 4.1, Department of Pharmaceutical Chemistry, University of California, San Francisco.
29. Stewart, J. P. (1990) MOPAC: A semiempirical molecular-orbital program, *J. Comput.-Aided Mol. Des. 4*, 1−45.
30. Cernohorsky, M., Kuty, M., and Koca, J. (1997) A Multidimensional Driver for Quantum Chemistry Program MOPAC, *Comput. Chem. 32*, 35−44.
31. Damborsky, J., Prokop, M., and Koca, J. (2001) TRITON: Graphic software for rational engineering of enzymes, *Trends Biochem. Sci. 26*, 71−73.
32. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A. Jr., Stratmann, R. E., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G. A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A. (1998) *Gaussian 98*, Gaussian, Inc., Pittsburgh, PA.
33. Warshel A. (1998) Electrostatic origin of the catalytic power of enzymes and the role of preorganized active sites, *J. Biol. Chem. 273*, 27035−27038.

BI034793O