Structures of Human Thymidylate Kinase in Complex with Prodrugs: Implications for the Structure-Based Design of Novel Compounds†

Nils Ostermann,‡,§ Dario Segura-Peña,|| Chris Meier,⊥ Thomas Veit,‡,# Christian Monnerjahn,▽ Manfred Konrad,▽ and Arnon Lavie*,‡,||

Department of Physical Biochemistry, Max Planck Institute for Molecular Physiology, Otto-Hahn-Strasse 11, 44227 Dortmund, Germany, Department of Biochemistry and Molecular Biology, University of Illinois at Chicago, 1819 West Polk Street, Chicago, Illinois 60612, Institute for Organic Chemistry, University of Hamburg, Martin-Luther-King-Platz 6, 20146 Hamburg, Germany, and Department of Molecular Genetics, Max Planck Institute for Biophysical Chemistry, 37070 Göttingen, Germany

Received December 4, 2002; Revised Manuscript Received January 8, 2003

ABSTRACT: Nucleoside analogue prodrugs are dependent on efficient intracellular stepwise phosphorylation to their triphosphate form to become therapeutically active. In many cases it is this activation pathway that largely determines the efficacy of the drug. To gain further understanding of the determinants for efficient conversion by the enzyme thymidylate kinase (TMPK) of clinically important thymidine monophosphate analogues to the corresponding diphosphates, we solved the crystal structures of the enzyme, with either ADP or the ATP analogue AppNHp at the phosphoryl donor site, in complex with TMP, AZTMP (previous work), NH_2TMP, d4TMP, ddTMP, and FLTMP (this work) at the phosphoryl acceptor site. In conjunction with steady-state kinetic data, our structures shed light on the effect of 3′-substitutions in the nucleoside monophosphate (NMP) sugar moiety on the catalytic rate. We observe a direct correlation between the rate of phosphorylation of an NMP and its ability to induce a closing of the enzyme's phosphate-binding loop (P-loop). Our results show the drastic effects that slight modifications of the substrates exert on the enzyme's conformation and, hence, activity and suggest the type of substitutions that are compatible with efficient phosphorylation by TMPK.

Derivatives of natural nucleosides are widely used for antiviral chemotherapy (for a recent review see ref *1*). These nucleoside analogues are prodrugs that have to be metabolized to their active triphosphorylated form after entering cells. As triphosphates they exert their antiviral effect by targeting viral polymerases in two different ways: by competitively inhibiting the incorporation of natural nucleotides into a growing viral DNA strand or by incorporating into a growing DNA strand and acting as chain terminators at the 3′ end since most of these compounds are modified at the 3′-position of the ribose moiety (*2*). However, for therapeutic efficacy, such nucleoside analogues must first be substrates for the kinases that catalyze their activation to the triphosphate form before they become substrates of their target polymerases. The activation pathway for the thymidine analogues 3′-azido-3′-deoxythymidine (AZT,[1] Zidovudine), 2′,3′-didehydro-2′,3′-dideoxythymidine (d4T; Stavudine), and 3′-fluoro-3′-deoxythymidine (FLT; Alovudine) is identical to the salvage pathway for thymidine triphosphate synthesis. The first step, the phosphorylation of the respective nucleoside to the corresponding nucleoside monophosphate (NMP), is catalyzed by thymidine kinase (TK), the second phosphorylation step to the diphosphate (NDP) is catalyzed by thymidylate kinase (TMPK), and the last step, the phosphorylation to the active triphosphate (NTP), can be accomplished by the base-unspecific nucleoside diphosphate kinase (NDPK). The efficiency of activation of the different agents varies and is dependent upon the nature of the drug. For d4T, the first activation step has been shown to be rate limiting (*3*). As a means to bypass the TK-dependent activation, masked nucleoside monophosphates (pronucleotides) have been developed that after cell entry undergo spontaneous hydrolysis to release nucleoside monophosphates (*4*−*6*). This strategy seems to be useful for a broad spectrum of nucleotide analogues, potentially making the initial step in the activation pathway no longer rate limiting. The last step in the activation, the phosphorylation of the

† Supported by grants from the National Institutes of Health [RO1-AI46943 (D.S.-P. and A.L.)], the Deutsche Forschungsgemeinschaft, and the Max Planck Society (C.M. and M.K.).

* To whom correspondence should be addressed at the Department of Biochemistry and Molecular Biology, University of Illinois at Chicago, 1819 W. Polk St., Chicago, IL 60612. Tel: (312) 355-5029. Fax: (312) 355-4535. E-mail: Lavie@uic.edu.

‡ Max Planck Institute for Molecular Physiology.

§ Current address: Novartis Pharma AG, Klybeckstrasse 141, CH-4057 Basel, Switzerland.

|| University of Illinois at Chicago.

⊥ University of Hamburg.

Current address: Degussa AG Care Specialties, Goldschmidtstrasse 100, 45127 Essen, Germany.

▽ Max Planck Institute for Biophysical Chemistry.

[1] Abbreviations: AZT, 3′-azido-3′-deoxythymidine; AZTMP, 3′-azido-3′-deoxythymidine monophosphate; d4T, 2′,3′-didehydro-2′,3′-dideoxythymidine; d4TMP, 2′,3′-didehydro-2′,3′-dideoxythymidine monophosphate; FLT, 3′-fluoro-3′-deoxythymidine; FLTMP, 3′-fluoro-3′-deoxythymidine monophosphate; NH_2TMP, 3′-amino-3′-deoxythymidine monophosphate; ddTMP, dideoxythymidine monophosphate; TMPK, thymidylate kinase; NMP, nucleoside monophosphate; NDP, nucleoside diphosphate; NTP, nucleoside triphosphate.

10.1021/bi027302t CCC: $25.00
Published on Web 02/11/2003

diphosphate metabolites to the active triphosphates, is usually not problematic since the NDPK is not base specific. In addition, NDPK is a very fast kinase so that even bad substrates are relatively efficiently phosphorylated. In contrast, the second phosphorylation step, catalyzed by TMPK, is crucial for the activation of a series of prodrugs, and it has been shown that this step is rate limiting in the activation pathway of AZT and FLT (*7*−*9*). In the case of AZT, this bottleneck in the activation pathway results in an accumulation of the partially activated and toxic (*10*−*12*) AZT monophosphate (AZTMP) metabolite in millimolar concentrations and in very low concentrations of the active triphosphorylated AZT (AZTTP) (*13*, *14*). To compound the problem caused by the inefficient activation of AZT, this prodrug in its unphosphorylated and partially phosphorylated forms can undergo a side reaction to yield 3′-amino-3′-deoxythymidine (*15*). The triphosphorylated form, 3′-amino-3′-deoxythymidine triphosphate, has been shown to be a substrate for the human DNA polymerases α and ϵ and thus presents a possible cause for some of the toxic side effects seen during AZT therapy (*16*, *17*). Due to the important role of TMPK in the activation of a number of prodrugs, we have initiated a structural examination of this enzyme, with the aim of understanding the determinants for efficient phosphorylation.

The physiological role of TMPK, a homodimeric enzyme, is to phosphorylate TMP to TDP using ATP·Mg as its preferred phosphoryl donor. Like other nucleoside monophosphate kinases (NMPKs), TMPK binds both substrates simultaneously. When both substrates are bound, the enzyme is in a closed conformation (*18*). Interestingly, in contrast to other NMPKs, TMPK makes further conformational changes within the closed conformation that are dependent upon the nature of the bound substrates (*19*). The structure of TMPK in complex with TMP and ADP revealed an open P-loop conformation, whereas in the complex of TMP and AppNHp the P-loop adopts a partially closed conformation. The complex of TMPK with TDP and ADP revealed an additional P-loop conformation that we called the fully closed state. Note that all of the above P-loop conformations are within the "globally closed" enzyme conformation. The different states observed for the P-loop correlate with the phosphorylation activity, with "P-loop open" being the inactive state (*18*).

To understand the reasons behind the poor phosphorylation rate of AZTMP by TMPK, we solved analogous structures to those mentioned above but where AZTMP is present instead of TMP. The conclusions from these studies were that the AZT azido group prevents TMPK from adopting the partially closed or closed conformations (*18*, *20*−*22*). This is due to steric hindrance between the 3′-azido group of AZTMP and the P-loop's Asp15 that would ensue if the P-loop undergoes the conformational change from the P-loop open to the "P-loop partially closed" conformation, assuming that the side chain of Asp15 adopts a similar position as seen in the presence of TMP. Thus, in the presence of AZTMP, the P-loop remains in the inactive P-loop open conformation, and this results in the diminished rate of AZTMP phosphorylation by TMPK.

Since the possible sites for modifying nucleosides while maintaining biological function are limited, and since many of the most selective and potent antiviral nucleoside-based prodrugs are modified at the ribose moiety, a detailed understanding of the function of the 3′-substituent for the catalytic function of TMPK is required for the development of novel drugs that are better metabolized to the active species. The steric hindrance caused by the azido group of AZTMP is absent in the isosteric TMP analogues 3′-amino-3′-deoxythymidine monophosphate (NH_2TMP) and 3′-fluoro-3′-deoxythymidine monophosphate (FLTMP) or in those that completely lack a 3′-substituent such as 2′,3′-didehydro-2′,3′-dideoxythymidine monophosphate (d4TMP) and dideoxythymidine monophosphate (ddTMP). Therefore, we determined the phosphorylation rate of these medicinally important nucleotide analogues (NH_2TMP, d4TMP, ddTMP, and FLTMP) by the human TMPK. In addition, we crystallized the ternary complexes of human TMPK with these antiviral metabolites at the phosphoryl acceptor binding site with either ADP or the ATP analogue adenosine β,γ-imido-5′-triphosphate (AppNHp) at the phosphoryl donor binding site (NH_2TMP was crystallized only in complex with AppNHp). These structures serve to explain the different kinetic properties of the respective substrates and, most importantly, provide a rationale for the design of new analogues that can be efficiently phosphorylated by human TMPK.

EXPERIMENTAL PROCEDURES

Purification and crystallization of human TMPK were done as described (*23*). Nucleotides of the highest purity grade were purchased from Pharma-Waldhof (Germany) or Boehringer Mannheim. Since the commercially available ATP analogue AppNHp contains significant amounts of ADP and ADP-NH_2 (∼10%), this nucleotide was further purified on a reversed-phase C18 column. Subsequent HPLC analysis showed >99% of the triphosphorylated form.

d4TMP was prepared from cycloSal-d4TMP pronucleotides (*6*, *24*) by chemical hydrolysis in water containing 2% triethylamine at 40 °C overnight. Purification was done after neutralization (acidic acid) and ethyl acetate extraction followed by lyophilization of the aqueous phase by chromatography on RP-18 columns. This procedure gave pure d4TMP as judged by NMR in 73% yield. FLTMP was synthesized from d4TMP.

NH_2TMP and ddTMP were purchased as the triphosphate from USB and Boehringer Mannheim, respectively. Formation of the monophosphate, achieved by incubating the triphosphate with 5 units of phosphodiesterase for 3−12 h at room temperature, was monitored by HPLC. To separate the products from the phosphodiesterase, a centrifugal filter (Vivaspin) with a 5 kDa cutoff was used.

Steady-state kinetic assays were performed at 25 °C in 50 mM Tris-HCl, pH 7.5, 100 mM KCl, and 5 mM $MgCl_2$ using a colorimetric NADH-dependent spectroscopic assay (*25*) with 100−250 nM TMPK. To determine k_{cat} and K_M values for ATP, TMP, AZTMP, d4TMP, and FLTMP, rates were measured by sets of six substrate concentrations ranging from 2 to about 200 μM. Due to limiting amounts of nucleotide available for NH_2TMP and ddTMP, their K_M values could not be determined, and their rates were measured at only one, apparently saturating, concentration of 200 and 250 μM, respectively.

All of the human TMPK structures reported herein contain the R200A mutation engineered to abolish the interaction

Table 1: Data Collection and Refinement Statistics

	ddTMP ADP	ddTMP AppNHp	d4TMP ADP	d4TDP AppNHp	FLTMP ADP	FLTMP AppNHp	NH_2TMP AppNHp
data collection statistics[a]							
X-ray source	ESRF ID2	rot. anode	DESY X11	DESY X11	DESY X11	DESY X11	rot. anode
wavelength (Å)	0.9887	Cu Kα	0.9076	0.9076	0.9076	0.9076	Cu Kα
unit cell (Å)							
$a = b$	100.535	101.426	101.000	101.100	101.510	101.260	101.301
c	49.963	49.322	49.840	49.800	50.01	49.780	49.513
resolution (Å)	1.55	1.9	1.55	1.5	1.7	1.6	1.75
observed reflections	119378	127182	165795	182998	137414	169968	142253
unique reflections	35723	20663	37349	41213	28602	34069	25995
completeness (%), overall/last shell	94.7/96.4	99.1/99.8	98.4/96.8	98.4 /99.8	97.5/99.3	98.1/98.9	97.1/97.2
R_{sym}[b] (%), overall/last shell	7.3/36.4	4.9/30.5	4.2/32.2	5.0/31.1	8.0/28.3	5.0/26.4	4.5/35.2
refinement statistics							
resolution range (Å)	70.7 −1.6	71.7−1.9	35.7−1.55	35.5−1.5	32.1−1.7	35.8−1.6	71.6−1.75
reflections with $F > 0\sigma$							
working	29245	18558	33602	37062	25759	30665	23213
test	3235	2103	3748	4146	2842	3400	2615
R_{cryst}[c]/R_{free}[d] (%)	21.1/26.5	17.8/24.1	18.5/22.5	19.2/21.5	18.3/22.8	17.5/22.0	18.6/24.3
no. of protein atoms[e]	1642	1708	1771	1611	1651	1642	1645
no. of nucleotide atoms[d]	47	51	47	51	48	100	52
no. of water atoms	305	281	317	297	323	353	282
no. of magnesium ions[f]	2	2	2	2	2	2	2
rms deviations							
bond length (Å)	0.014	0.013	0.010	0.009	0.013	0.010	0.012
bond angles (deg)	1.7	1.9	1.5	1.6	1.4	1.5	1.8
average B (Å^2)							
main chain	24.94	23.97	14.61	16.17	20.25	16.09	22.80
side chain	28.48	28.33	19.71	19.55	23.42	18.71	26.66
water molecules	41.81	39.77	32.96	33.34	36.05	33.90	39.07
nucleotides	24.45	24.84	15.16	16.58	18.74	14.00	21.93
magnesium ions	25.69	26.37	15.38	16.78	19.65	14.86	22.51

[a] All data were collected at 100 K using a MAR345 Image Plate detector. In all cases the space group is $P4_32_12$ with one molecule in the asymmetric unit. [b] $R_{sym} = \sum|I_o - \langle I\rangle|/\sum\langle I\rangle$. [c] $R_{cryst} = \sum|F_o - F_c|/\sum F_o$. [d] $R_{free} = R_{cryst}$, calculated for 10% randomly selected reflections not included in the refinement. [e] Number includes residues modeled in double conformations and residues modeled as alanine if no electron density for the side chain is observed; for details, see PDB header. [f] Only one of the metal ions is bound to ADP/ATP.

of this surface arginine with the active site of a neighboring molecule in the crystal. This mutation does not affect the structure or the kinetics of the enzyme (*19*). Cocrystallization of TMPK with the different nucleotides involved adding to the purified protein the phosphoryl acceptor to a final concentration of 2 mM (NH_2TMP, FLTMP, d4TMP, ddTMP) in combination with 2 mM ADP or 20 mM AppNHp. Using the hanging-drop method, crystals grew in a few days at room temperature to a typical size of 300 × 300 × 100 μm^3 in space group $P4_32_12$. Diffraction data collected at cryogenic temperatures were processed with XDS and XSCALE (*26*). To eliminate model bias, rigid body refinement using a model without nucleotides and waters was subjected to a simulated annealing step as implemented in X-PLOR (*27*). Inspection of the resulting electron density and manual rebuilding was accomplished using the graphics program O (*28*). Subsequent refinement was done using REFMAC (*29*), and water molecules were added automatically using ARP (*29*). Models and structure factors are deposited in the Protein Data Bank (accession numbers 1NMX, 1NMY, 1NMZ, 1NN0, 1NN1, 1NN3, and 1NN5).

RESULTS

All previous structures and those we report here of the human TMPK are characterized by both nucleotide-binding sites being occupied (donor and acceptor sites) and an enzyme in the globally closed conformation (*18*, *19*). The differences we observe in the conformation of active site residues (e.g., P-loop open, partially closed, and fully closed) can be directly attributed to the nature of the nucleotide at the donor site (either ADP or AppNHp) and to the different substituents at the ribose of the acceptor nucleotide. Data collection and refinement statistics for these seven structures are presented in Table 1. In four of these seven structures we interpret the electron density to represent a mixture of two different states that differ in P-loop conformation and/or phosphate conformation of the nucleotide acceptor. We have observed this phenomenon of multiple states during our previous work on TMPK. This fact, the high resolution of our data sets, and the internal correlation between the occupancies of the different states within each structure allow us to make such an interpretation with confidence.

In the following sections the enzyme and nucleotide conformations in the different complexes are presented. While we focus on the different P-loop conformations (open, partially closed, and closed), we note that these conformations are accompanied by concomitant conformational changes in the LID region (residues 133−157) and the adenine-binding loop (residues 180−184). In addition, as seen in previously solved structures of TMPK, the phosphate group of the acceptor nucleotide (i.e., TMP and its analogues), located between Arg45 and Arg97, can be observed in three different positions (Figure 1).

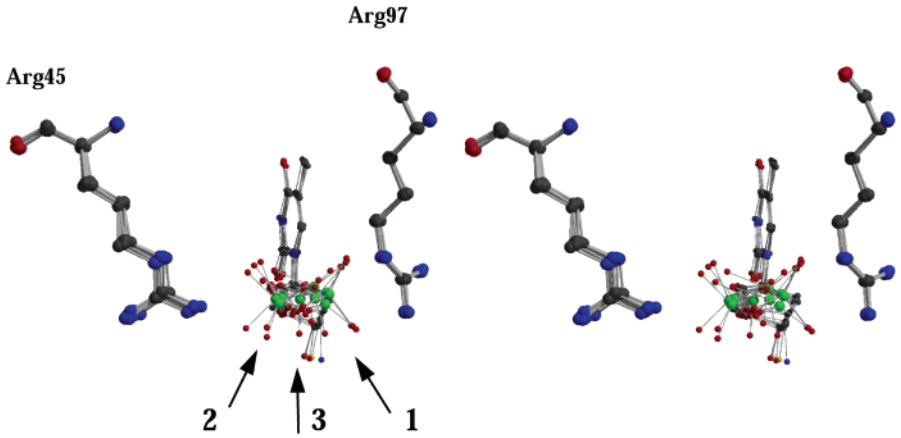

FIGURE 1: The NMP phosphate can occupy three different positions in human TMPK. Stereoview of the overlay of nine TMPK complex structures [the seven reported here plus the complexes of TMP with ADP or AppNHp (*19*)] showing the nucleoside acceptor phosphate group sandwiched between Arg45 and Arg97. The three arrows point to the three different positions we observe for the acceptor nucleotide phosphate group: one, close to Arg97, enabling a bidentate interaction involving Nϵ and NH2 with two of the phosphate's oxygens [this conformation was first observed in the TMP−ADP complex (i.e., in the absence of a γ-phosphate of the phosphoryl donor] (*19*); two, rotated away from Arg97 to allow only an interaction with Arg45, a conformation seen in the presence of TMP and AppNHp; three, a position that occurs midway between the two arginines, maintaining an interaction with each but having the third phosphate oxygen free. This was observed in the transition state mimic structure that included TMP, ADP, and aluminum fluoride. The distance between the phosphate positions is 1.6, 1.1, and 1.9 Å for the d4TMP, ddTMP, and FLTMP structures, respectively. The different NMP phosphate position correlates with the propensity of each TMP analogue to induce P-loop closure. It is in turn the resulting P-loop conformation that to a large extent determines the phosphorylation efficiency of the NMP. All structural figures were generated using Bobscript (*31*) and raster3D (*32*).

Table 2: Apparent Kinetic Constants for TMPK-Catalyzed Phosphorylation

nucleotide	k_{cat} (s^{-1})	K_M for NMP (μM)	K_M for ATP (μM)
TMP	0.70	5	5
AZTMP	0.01	12	69
NH_2TMP	0.14	ND[a]	ND
d4TMP	0.09	12	33
FLTMP	0.03	8	27
ddTMP	0.03	ND	ND

[a] ND, not determined.

TMPK in Complex with NH_2TMP. Steady-state kinetic analysis revealed that TMPK phosphorylates NH_2TMP 14-fold faster than AZT-MP (Table 2). To understand the reasons that confer this considerably higher activity, we solved the crystal structure of TMPK in complex with NH_2-TMP and the ATP analogue AppNHp. Since NH_2TMP shows higher turnover than AZTMP for TMPK, we anticipated the P-loop not to be open in the NH_2TMP−AppNHp complex. This prediction was correct with a slight caveat since the P-loop was observed in the fully closed conformation (and not the partially closed as expected). Both nucleotides are observed with 100% occupancy. In this conformation, the main chain amide of Arg16 of the P-loop makes an ∼3.0 Å (∼3.4 Å in the P-loop open conformation) interaction to the NH group between the β- and γ-phosphates of AppNHp. This interaction is very important for terming this P-loop conformation as an active conformation since the negative charge that arises at this ATP oxygen during the phosphoryl transfer reaction has to be stabilized (Figure 2).

As a result of the presence of a γ-phosphate in the donor site and the fully closed P-loop conformation, the phosphate of NH_2TMP rotates away from Arg97, maintaining the interaction via Nϵ to Arg97 but breaking the interaction via NH2. The phosphate group of NH_2TMP is now at an H-bond distance to Arg45 (3.3 Å). The third oxygen atom of the

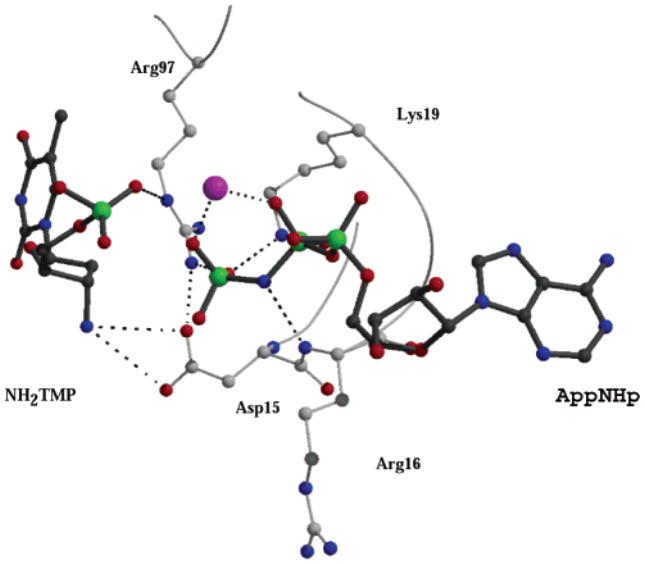

FIGURE 2: AppNHp−TMPK interaction. Shown is the NH_2TMP + AppNHp complex. The phosphates of AppNHp are positioned through many interactions between the oxygen atoms and the P-loop main chain amides and the catalytically essential magnesium ion as described previously (*19*). The magnesium ion interacts with oxygen atoms of both the β-phosphate and the γ-phosphate of AppNHp and thereby contributes to charge distribution during the phosphoryl transfer reaction. The side chain of the highly conserved P-loop Lys19 interacts with oxygen atoms of the β- (3.0 Å) and γ- (2.6 Å) phosphates of AppNHp and the carbonyl group of Val14 (2.8 Å; not shown for clarity). The same oxygen atom of the γ-phosphate that interacts with the P-loop lysine is in addition hydrogen bonded to NH2 of the highly conserved Arg97 (2.8 Å). At the same time, Arg97 interacts with an oxygen atom of the NH_2-TMP phosphate via Nϵ (2.8 Å). Thus, Arg97 stabilizes negative charges from both the phosphoryl donor and acceptor, and it acts as a clamp to bring both nucleotides close to each other. Similar interactions occur with the other NMP's examined.

NH_2TMP phosphate group is not hydrogen bonded to any protein residues and is at a 3.9 Å distance to the phosphorus

atom of the γ-phosphoryl group of AppNHp. The P-loop carboxylic acid is in a position to make a bidentate interaction to the 3′-amino group of NH_2TMP (3.4 Å distance of both oxygen atoms). In addition, atom OD1 of Asp15 is 2.3 Å apart from NH2 of Arg97 (Figure 3a).

TMPK in Complex with d4TMP. d4T is an anti-AIDS prodrug licensed for clinical use. It lacks a 3′-substituent, and in contrast to AZT and FLT the partially activated monophosphate metabolite of d4T does not accumulate in cells treated with this prodrug. A steady-state kinetic analysis of the phosphorylation rate of d4TMP in comparison to FLTMP by human TMPK shows that d4TMP is a 3-fold faster substrate than FLTMP (Table 2). Remarkably, 3′-substitutions appear to affect primarily the rate (k_{cat}) but not the apparent affinity (K_M). The structures of human TMPK in complex with d4TMP at the monophosphate-binding site and ADP or AppNHp at the phosphoryl donor binding site give an explanation for these kinetic data. The ternary structure with ADP-Mg at the phosphoryl donor site shows the P-loop in two conformations, both modeled with 50% occupancy (Figure 3b). The first conformation is the P-loop open (and inactive) conformation that we usually observe with any monophosphate at the phosphoryl acceptor binding site and ADP at the phosphoryl donor binding site. The second conformation is the P-loop partially closed conformation with the Asp15 rotated toward the monophosphate, making interactions to both Arg97 and Gln157 (2.9 Å to NH2 and 3.1 Å to NE2, respectively). Observing the P-loop in the partially closed conformation in the presence of ADP, albeit in 50% occupancy, is unique for the substrate d4TMP. For all other ternary complexes of TMPK with TMP, AZTMP, FLTMP, or ddTMP at the acceptor binding site and ADP at the donor site, we observe the P-loop 100% in the open conformation. Modeling and estimation of relative occupancy of the different P-loop (and nucleotide) conformations are made possible by our high-resolution data. The methodology we used to determine the occupancy of atoms in multiple conformations is shown in Figure 4, with the d4TMP−ADP complex as an illustrative example.

The thymine moiety of d4TMP, just as in all other TMP analogues tested, is bound at the same position with a nearly perfect overlay of the thymine part among the different complexes. However, differences in position of the sugar and phosphate moieties are observed among these structures. In the case of d4TMP, the ribose atoms C1′, C2′, C3′, and C4′ are in plane due to the double bond between atoms C2′ and C3′. This is in contrast to the 3′-endo conformation of the ribose moiety we observed with all other thymidine analogues. As a result of this conformational restraint of the d4TMP ribose, the phosphate group of d4TMP is not in the regular position observed in complexes with TMP, AZTMP, FLTMP, and ddTMP. In the latter, the phosphate group makes two interactions with Arg97 (to Nϵ and NH2) via two of its phosphate oxygens. But in the case of d4TMP, the phosphate group shifts away from Arg97, breaking the interaction with the NH2 atom, and moves within 3.1 Å to Arg45 (Figures 1 and 3b). It maintains a single interaction to Arg97 via Nϵ (2.6 Å). The Arg97 NH2 atom that previously made an interaction with the phosphate group is within H-bond distance (2.9 Å) to the P-loop carboxylic acid Asp15, thus stabilizing the partially closed conformation. In addition, the lack of a 3′-hydroxyl group in d4TMP could also serve to destabilize the open P-loop conformation since this hydroxyl group in TMP interacts with a water molecule that plays a role in the stabilization of the water structure of the open conformation. Together, these factors seem to be sufficient to make the P-loop open and the P-loop partially closed conformations energetically similar, resulting in our observation of the P-loop in both conformations with 50% occupancy.

The ternary complex of human TMPK with d4TMP and AppNHp shows the P-loop in a single, partially closed conformation and with 100% occupancy for both nucleotides (Figure 3c). The triphosphate analogue is at the same position to that seen in the analogous complex with TMP. The γ-phosphate of AppNHp repels the phosphate group of d4TMP toward Arg45 so that the interaction between the d4TMP phosphate group and NH2 of Arg97 is broken.

An important difference between the d4TMP complex structure with AppNHp to that with ADP is that the d4TMP phosphate group, in the presence of ADP, is in a position midway between Arg97 and Arg45, while with AppNHp it completely breaks its interactions with Arg97 and is only interacting with Arg45. In the case of the physiological substrate TMP, the presence of ADP still maintains the bidentate interaction between the phosphate group of TMP and Arg97. As the phosphate group breaks the interaction with the NH2 group of Arg97, it frees this atom to interact with the carboxylic moiety of Asp15. Since this interaction (Asp15-Arg97) cannot take place in the open P-loop conformation, relative stabilization of the closed P-loop conformation over the open state results. This propensity of d4T to induce closure of the P-loop is interpreted as a main reason for the rather good phosphorylation kinetics observed with this TMP analogue.

TMPK in Complex with ddTMP. To further explore the hypothesis that the positioning of the phosphate group, due to the restraint of the sugar pucker in d4TMP, contributes to the P-loop closure and is not a result of the lack of a 3′-hydroxyl group, we determined the structures of TMPK in complex with ddTMP, both with ADP and with AppNHp. Steady-state kinetics reveals a low rate of ddTMP phosphorylation by the human TMPK, similar to that observed with FLTMP and 3-fold slower than d4TMP (Table 2). The structure of human TMPK in complex with ddTMP and ADP shows that the P-loop is in the open and inactive conformation. The ribose of ddTMP is in the 3′-endo conformation, and the phosphate group makes two hydrogen bonds to Arg97, very similar to the structure with TMP (Figure 3d).

In the structure of TMPK in complex with ddTMP and AppNHp, both nucleotides are bound with 100% occupancy and are observed in a single conformation. In comparison to the complex with ADP, the presence of AppNHp results in a displacement of the phosphate group of ddTMP, away from Arg97 toward Arg45. However, while in the case of TMP AppNHp causes both bonds between Arg97 and the phosphate group to break, in the presence of ddTMP the phosphate group maintains its interaction to Arg97 via the Nϵ atom. As an additional difference to the case when TMP is bound, in the presence of ddTMP, despite having 100% occupancy of AppNHp the P-loop is observed in two conformations, open and partially closed, both with 50% occupancy (Figure 3e). In other words, the presence of AppNHp was not sufficient to totally shift the P-loop

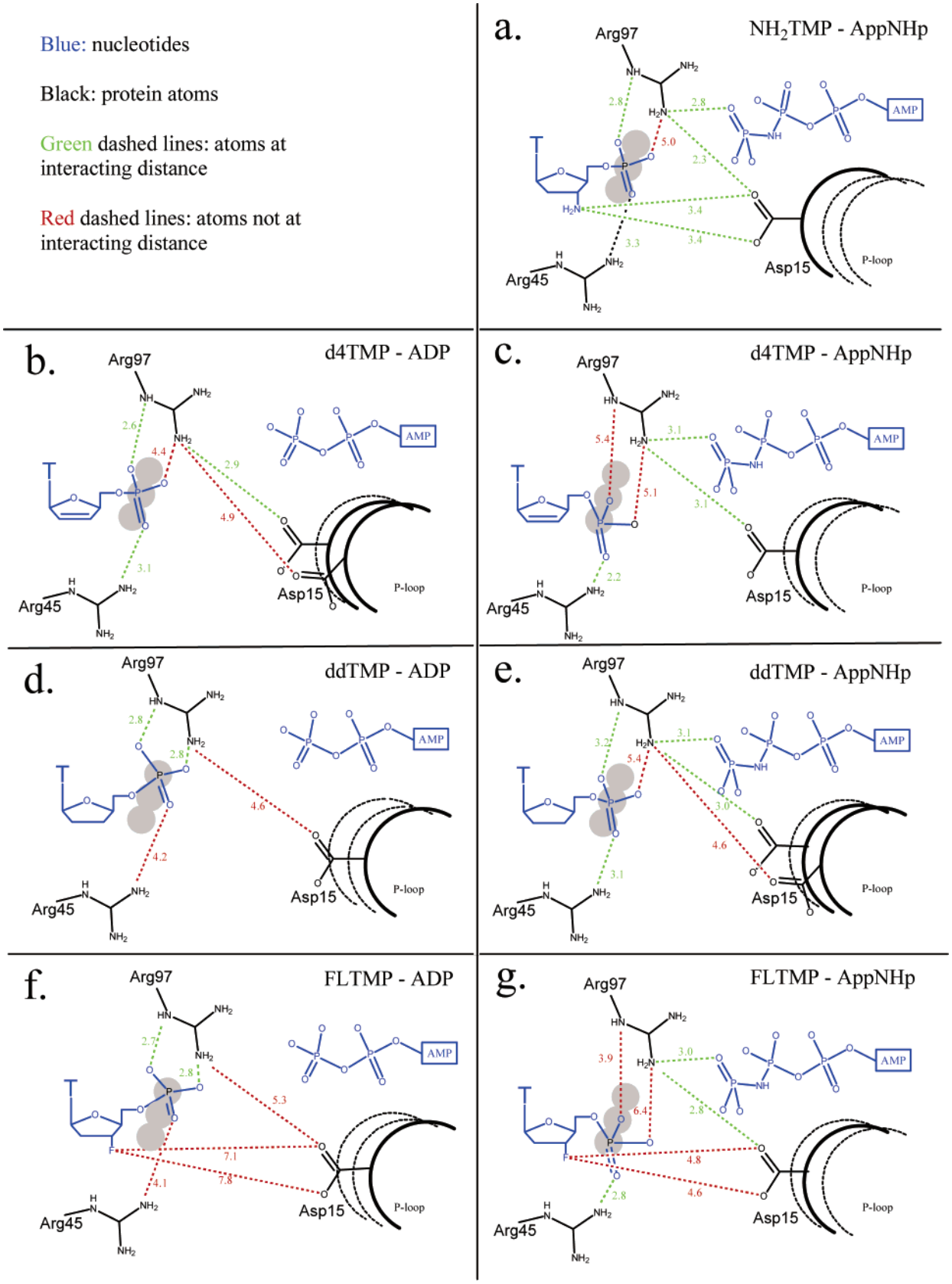

FIGURE 3: Schematic drawing of the interactions between human TMPK and (a) NH_2TMP + AppNHp, (b) d4TMP + ADP, (c) d4TMP + AppNHp, (d) ddTMP + ADP, (e) ddTMP + AppNHp, (f) FLTMP + ADP, and (g) FLTMP + AppNHp. The three possible positions for the NMP phosphate are depicted as gray circles. The different P-loop conformations (open, partially closed, and fully closed) are shown as black arcs, where the conformation(s) observed for the particular complex is (are) in bold; the others shown for reference are depicted as dashed lines. In complexes with AppNHp that showed the presence of ADP (i.e., a mixture of AppNHp and ADP was observed), the ADP state is not shown. The open P-loop conformation is that furthest away from the NMP, the closed conformation the closest, and the partially closed in between.

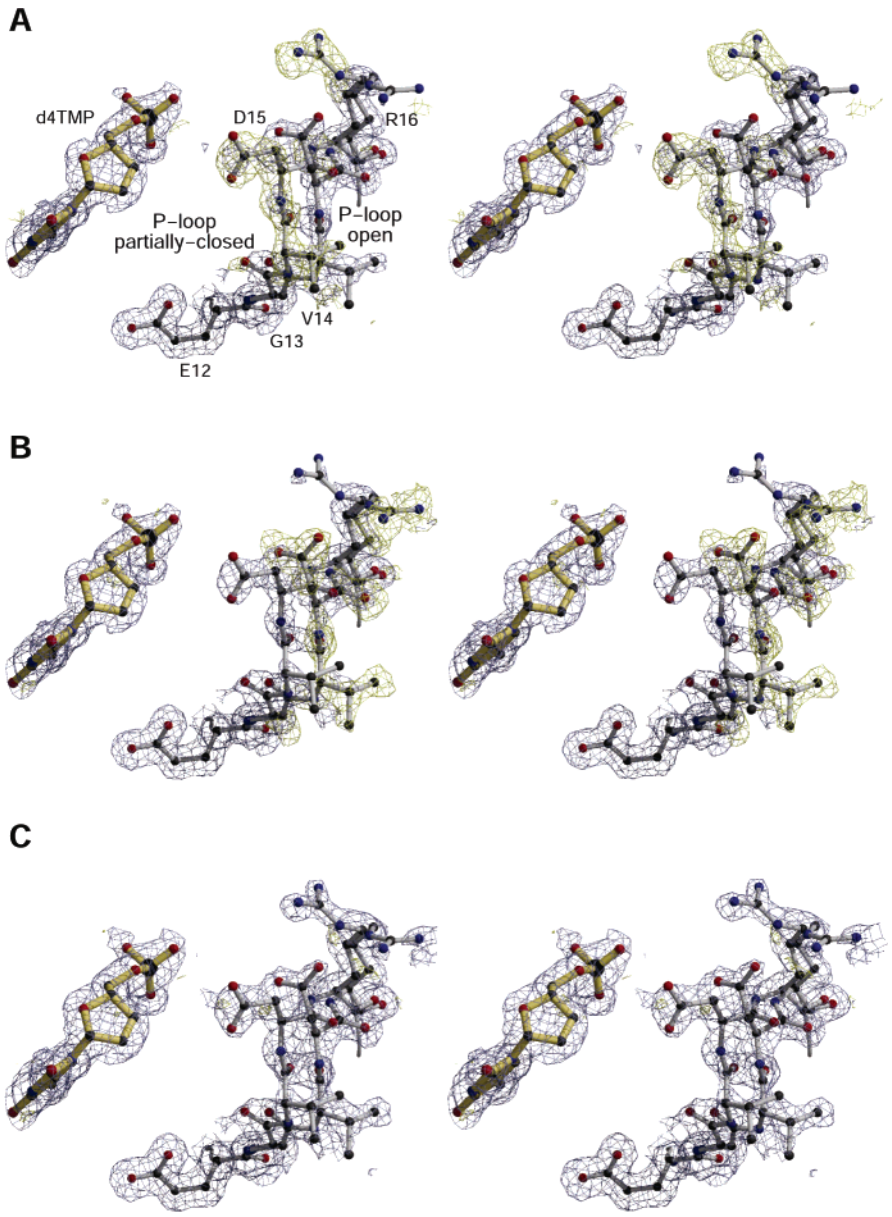

FIGURE 4: Use of difference maps guided the determination of the different conformations and occupancies of enzyme loops and nucleotides. For example, in the case presented for the complex of TMPK with d4TMP and ADP, which was solved at 1.55 Å resolution, P-loop residues 12−16 and d4TMP are depicted in a ball-and-stick representation. The final refined model is shown overlaid with the sigma A weighted electron density maps: $2mF_o - DF_c$ [gray, contoured at 1.2σ for (A) and (B) and 0.8σ for (C)] and a $mF_o - DF_c$ (yellow, contoured at 2.5σ) for differing occupancies of the open and partially closed conformation. In (A), 100% occupancy of the open conformation was modeled. In the $mF_o - DF_c$ difference map, the density for the not modeled partially closed conformation is clearly visible. In (B), 100% occupancy of the P-loop partially closed conformation was modeled, resulting in difference density for the open conformation. In (C), 50% occupancy for both the P-loop open and P-loop partially closed is refined, resulting in no noticeable electron density in the difference map.

conformation from the open, and inactive, to the partially closed, and active, state. We take this fact to rationalize the very low turnover rate with ddTMP in comparison to the rate with the physiological substrate TMP.

TMPK in Complex with FLTMP. FLTMP, in which the 3′-hydroxyl group of TMP is replaced by a fluoro atom, is the first phosphorylated intermediate of the anti-AIDS prodrug FLT [development of FLT was suspended after phase II trials due to hemotoxicity (*30*)]. The activation pathway of FLT is identical to that described for AZT. For both drugs, the monophosphate form accumulates in cells. For AZT we attribute the reasons for the poor activation to the bulkier azido group. For steric reasons, the 3′-substituent of AZTMP prevents human TMPK from occupying the P-loop closed and active conformation. The poor activation of FLTMP (Table 2) by human TMPK cannot be caused by steric hindrance since the 3′-fluoro atom is even smaller than the hydroxyl group in TMP. What then makes FLTMP a much poorer substrate than TMP?

A comparison of the structure of human TMPK in complex with FLTMP and ADP with the previously reported structure of human TMPK in complex with TMP and ADP shows that both structures (protein residues as well as the nucleotides) are nearly identical (Figure 3f). In both structures we observe a very similar water structure between the P-loop and the respective monophosphate, including the water molecule 510 being in a hydrogen-bonding distance to the 3′-substituent, contributing to the stabilization of the open

Table 3: Conformations and Occupancy of Observed Structures

cocrystallized with	nucleotides observed	P-loop conformation	TMP phosphate interacts with	occupancy[c] (%)
TMP and ADP[a]	TMP and ADP	open	Arg97	100
TMP and AppNHp[a]	TMP and ADP	open	Arg97	33
	TMP and AppNHp	partially closed	Arg45	67
TMP + ADP + AlF_3[a]	TMP + ADP + AlF_3	partially closed	Arg97-Arg45	100
TMP + ATP[a]	TDP + ADP	fully closed	product[d]	100
AZTMP and ADP[a]	AZTMP and ADP	open	Arg97	100
AZTMP and AppNHp[a]	AZTMP and ADP	open	Arg97	33
	AZTMP and AppNHp	open	Arg97-Arg45	67
NH_2TMP and AppNHp[b]	NH_2TMP and AppNHp	fully closed	Arg97-Arg45	100
d4TMP and ADP[b]	d4TMP and ADP	open	Arg97-Arg45	50
	d4TMP and ADP	partially closed	Arg45	50
d4TMP and AppNHp[b]	d4TMP and AppNHp	partially closed	Arg45	100
ddTMP and ADP[b]	ddTMP and ADP	open	Arg97	100
ddTMP and AppNHp[b]	ddTMP and AppNHp	open	Arg97-Arg45	50
	ddTMP and AppNHp	partially closed	Arg97-Arg45	50
FLTMP and ADP[b]	FLTMP and ADP	open	Arg97	100
FLTMP and AppNHp[b]	FLTMP and ADP	open	Arg97	50
	FLTMP and AppNHp	partially closed	Arg45	50

[a] These complex structures are presented in ref *18*. [b] This work. [c] We observe in many of the structures two conformations of both protein and nucleotide atoms that are interpreted as representing two distinct states. The partial occupancy of AppNHp can be attributed to contaminating amounts of ADP present despite HPLC purification of the ATP analogue, to the enzyme preparation not being nucleotide free (∼20% of TMP and ADP), or to a breakdown of AppNHp during the time required for crystallization. The high resolution of our data sets allows us to approximate the relative occupancy with good confidence. [d] The phosphates of the product TDP are largely displaced in comparison to the substrate TMP due to a conformational change in the position of Arg97.

conformation. The observed slight displacement of LID region residues 144−149 can be an artifact caused by poor electron density for these residues especially in the FLTMP structure, as indicated by the ∼3-fold higher *B*-factors in comparison to the average values for the complete structures for this region.

The electron density of the human TMPK structure in complex with FLTMP and AppNHp is interpreted to be composed of a mixture of the bound nucleotides and two P-loop conformations (Figure 3g). By comparing the electron density that we observe for the γ-phosphoryl group for AppNHp in comparison to the α- and β-phosphate, we modeled AppNHp with an occupancy of 50% and ADP with an occupancy of 50%. Consistent with this interpretation are two conformations of the phosphate group of FLTMP and two P-loop conformations: one conformation modeled with 50% occupancy that corresponds to ADP bound at the phosphoryl donor binding site, being identical to that seen in the complex of TMPK with FLTMP and ADP, and the second displaced conformation also modeled with 50% occupancy that corresponds to AppNHp bound at the donor site. The conformation of TMPK with FLTMP and AppNHp is identical to that seen of TMPK with the physiological substrate TMP and AppNHp.

Thus, we interpret the lack of any obvious differences between structures of TMPK in complex with TMP or with FLTMP, with either ADP or AppNHp, to suggest that the reasons that make FLTMP a poorer substrate than TMP reside in a step of the reaction coordinate other to that we have visualized. The most likely one is the P-loop shift from the partially closed to the fully closed conformation, a protein isomerization step that follows the chemical step but is prior to product release (*19*). The partial negative charge of the fluorine atom would destabilize the close contact with the side chain of Asp15, as is observed in the fully closed P-loop conformation. It is this destabilization that could slow the transition of the P-loop from the partially closed to the fully closed conformation and thus result in the reduced rate of FLTMP phosphorylation.

DISCUSSION

Correlation between kinetic and structural data, presented here and in previous work, has established a strong connection between the conformation of the P-loop and the rate of phosphorylation. The open P-loop conformation is interpreted to represent an inactive state, whereas the partially closed and fully closed conformations are interpreted to represent active states. A good example for the importance of the partially closed P-loop conformation for catalysis comes from our work with the F105Y TMPK mutant (*18*). This single substitution, by interfering with the water structure that stabilizes the open conformation, promotes P-loop closure (i.e., the open state is destabilized by the F105Y mutation). Noteworthy, the F105Y mutant is ∼20-fold faster at AZTMP phosphorylation than wild-type TMPK, a fact we attributed to this mutant's preference for the closed (i.e., active) conformation.

Considering the substrates, the nature of both nucleotide donor and acceptor plays a role in determining the P-loop conformation (Table 3). The consequence of a γ-phosphate in the nucleotide bound at the donor site was observed upon comparing structures with AppNHp to those with ADP. For example, in combination with the physiological substrate TMP at the acceptor side, the presence of a γ-phosphate at the donor site causes the P-loop to adopt the partially closed conformation. The shift in P-loop conformation is a result of the displacement of two waters, important for the stabilization of the open state, by the γ-phosphoryl group of AppNHp. This weakens the stabilization of the open state and hence promotes P-loop closure. However, if TMP is replaced by the HIV prodrug AZTMP, the presence of a γ-phosphate is not sufficient to promote P-loop closure. The presence of the bulky azido group at the 3′-position of AZTMP was given as the cause behind this observation. If

the 3′-position is occupied by an amine group, as in NH_2-TMP, the presence of a γ-phosphate is not only sufficient to promote P-loop closure (Figure 3a) but in fact results in the fully closed P-loop conformation and not the partially closed one as seen with TMP and AppNHp. In this case, it seems that the ability of the amine moiety to interact with the P-loop Asp15 plays a crucial role in stabilizing the fully closed P-loop conformation. The close proximity of the side chain of Asp15 to the amine group is stabilized by hydrogen bond interactions or by ionic interactions if assuming the protonated state of NH_2TMP. Since P-loop closure is required for TMPK activity, and NH_2TMP is even better than TMP at promoting P-loop closure, why is the rate of NH_2TMP phosphorylation ~3-fold slower than the TMP phosphorylation rate? Here we can only speculate that, by too strongly stabilizing the closed P-loop conformation, other steps in the reaction that might require an open P-loop, such as product release, are slowed, thus resulting in a slightly reduced steady-state rate. If this hypothesis were correct, it would serve as another example to demonstrate the importance of enzyme dynamics for catalysis.

While an H-bond donor such as an amine group at the 3′-position is favorable for catalysis, a partially negatively charged fluorine atom, an H-bond acceptor, at that position is detrimental for activity (Table 2). We attribute the low activity of FLTMP phosphorylation by TMPK to the destabilizing interaction between the 3′-flourine atom and the side chain of Asp15 that would take place in the fully closed P-loop conformation. In the presence of TMP, the fully closed P-loop conformation [as observed in the complex with TDP and ADP (*19*)] is a part of the reaction coordinate, being adopted after the phosphoryl transfer step. To test our hypothesis concerning FLTMP, we require an additional structure in which products are present (i.e., FLT diphosphate and ADP).

Since a hydrogen bond donor at the 3′-position of the acceptor nucleotide is beneficial for TMPK activity, whereas a hydrogen bond acceptor is detrimental, we asked the question: What is the effect of no 3′-substituent on the phosphorylation rate? This question takes special relevance since the approved HIV drug d4T lacks a 3′-substituent, as does ddT, but d4TMP is phosphorylated three times faster than ddTMP. The structures with d4TMP and ddTMP reveal the effect of the position of the phosphate group of the acceptor nucleotide on the P-loop conformation. This phosphate can occupy three distinct positions: making two hydrogen bonds to Arg97, making only an interaction to Arg45, or having an intermediate position that interacts concomitantly with both Arg97 and Arg45 (see Figure 1). The conformation in which the phosphate makes two hydrogen bonds to Arg97 is not conducive for P-loop closure for the following reasons. One, in that conformation, repulsion between the NMP phosphate and the γ-phosphate of ATP will take place. Two, by the phosphate oxygen atom interacting with the NH2 atom of the arginine side chain, this atom is not free to interact with Asp15. Both of these facts will destabilize the active closed P-loop conformation. In other words, closure of the P-loop requires the rotation of the acceptor nucleotide phosphate group away from Arg97 toward Arg45.

How does this requirement for the position of the acceptor's phosphate group relate with d4TMP and ddTMP? The double bond between atoms C2′ and C3′ of the d4TMP ribose forces a planar arrangement of the atoms C1′, C2′, C3′, and C4′. As a result, the phosphoryl group cannot occupy the position that makes two hydrogen bonds to Arg97. This leaves the NH2 atom of Arg97 free to interact with the carboxylate group of Asp15 in the P-loop partially closed conformation. Thus, in the case of d4TMP, its phosphate group is hindered from making the bidentate interaction with Arg97. This explains why even in the d4TMP−ADP complex, which should not induce P-loop closure, we observe 50% occupancy for the partially closed P-loop conformation, whereas the TMP−ADP complex is 100% in the open conformation. In the presence of AppNHp, the complex with d4TMP is very similar to the complex with TMP, with the phosphate group of d4TMP only interacting with Arg45 (Figure 3c). In contrast, in the case of ddTMP (an unrestrained ribose) and AppNHp, the phosphate group of ddTMP remains at the intermediate position between Arg97 and Arg45 (Figure 3e). As a result, despite the presence of AppNHp, the P-loop is only 50% in the partially closed conformation, and 50% remains in the open state. Hence, ddTMP is a poorer substrate than d4TMP due to its inferior ability to promote P-loop closure.

CONCLUSION

The activity of nucleoside analogue prodrugs is to a large extent determined by the ability of enzymes to convert the inactive prodrug to the active drug. This conversion, in the case of most nucleoside analogues, means the serial phosphorylation of the prodrug from the 5′-unphosphorylated nucleoside to the 5′-triphosphorylated form. In this work we explored the effects of sugar modifications in four thymidine analogues: NH_2TMP, d4TMP, ddTMP, and FLTMP. The relative phosphorylation rates of these analogues and of TMP can be correlated with the propensity of the P-loop to form a closed conformation in their presence. Factors directly related to the nature of the analogue that play a role in P-loop conformation, and hence the phosphorylation rate, are (a) the hydrogen bond capability of a 3′-substituent, where a donor (or positive charge) promotes P-loop closure and an H-bond acceptor (or negative charge) hinders full P-loop closure, and (b) the position of the analogues' phosphate group, which can be influenced by the sugar moiety and its propensity to interact with Arg97, where the break of interaction between Arg97 and the phosphate group promotes P-loop closure. The structure−activity relationships presented herein, together with the previously reported structures of human TMPK in complex with the natural substrate TMP and the AZT metabolite AZTMP, have important implication for the further development of novel and selective antiviral nucleoside analogues that can be more efficiently phosphorylated by human TMPK and thus may finally show a higher therapeutic index.

ACKNOWLEDGMENT

We thank the staff of EMBL, Hamburg, Germany, and of the ESRF beamline ID12, Grenoble, France, for help in data collection.

REFERENCES

1. De Clercq, E. (2001) *J. Clin. Virol. 22*, 73−89.

2. Arts, E. J., and Wainberg, M. A. (1996) *Antimicrob. Agents Chemother. 40*, 527−540.
3. Balzarini, J., Herdewijn, P., and De Clercq, E. (1989) *J. Biol. Chem. 264*, 6127−6133.
4. Meier, C., Lorey, M., De Clercq, E., and Balzarini, J. (1997) *Nucleosides Nucleotides 16*, 1303−1306.
5. Meier, C., Lorey, M., De Clercq, E., and Balzarini, J. (1997) *Bioorg. Med. Chem. Lett. 7*, 99−104.
6. Meier, C., Lorey, M., De Clercq, E., and Balzarini, J. (1998) *J. Med. Chem. 41*, 1417−1427.
7. Furman, P. A., and Barry, D. W. (1988) *Am. J. Med. 85*, 176−181.
8. Matthes, E., Lehmann, C., Scholz, D., Rosenthal, H. A., and Langen, P. (1988) *Biochem. Biophys. Res. Commun. 153*, 825−831.
9. Qian, M., Bui, T., Ho, R. J., and Unadkat, J. D. (1994) *Biochem. Pharmacol. 48*, 383−389.
10. Bridges, E. G., Faraj, A., and Sommadossi, J. P. (1993) *Biochem. Pharmacol. 45*, 1571−1576.
11. Törnevik, Y., Ullman, B., Balzarini, J., Wahren, B., and Eriksson, S. (1995) *Biochem. Pharmacol. 49*, 829−837.
12. Yan, J. P., Ilsley, D. D., Frohlick, C., Steet, R., Hall, E. T., Kuchta, R. D., and Melancon, P. (1995) *J. Biol. Chem. 270*, 22836−22841.
13. Frick, L. W., Nelson, D. J., St. Clair, M. H., Furman, P. A., and Krenitsky, T. A. (1988) *Biochem. Biophys. Res. Commun. 154*, 124−129.
14. Fridland, A., Connelly, M. C., and Ashmun, R. (1990) *Mol. Pharmacol. 37*, 665−670.
15. Cretton, E. M., Xie, M. Y., Bevan, R. J., Goudgaon, N. M., Schinazi, R. F., and Sommadossi, J. P. (1991) *Mol. Pharmacol. 39*, 258−266.
16. Chidgeavadze, Z. G., Beabealashvilli, R. S., Atrazhev, A. M., Kukhanova, M. K., Azhayev, A. V., and Krayevsky, A. A. (1984) *Nucleic Acids Res. 12*, 1671−1686.
17. Jasko, M. V., Fedorov, I. I., Atrzhev, A. M., Mozzherin, D. Y., Novicov, N. A., Bochkarev, A. V., Gurskaya, G. V., and Krayevskt, A. A. (1995) *Nucleosides Nucleotides 14*, 23−37.
18. Ostermann, N., Lavie, A., Padiyar, S., Brundiers, R., Veit, T., Reinstein, J., Goody, R. S., Konrad, M., and Schlichting, I. (2000) *J. Mol. Biol. 304*, 43−53.
19. Ostermann, N., Schlichting, I., Brundiers, R., Konrad, M., Reinstein, J., Veit, T., Goody, R. S., and Lavie, A. (2000) *Struct. Folding Des. 8*, 629−642.
20. Lavie, A., Vetter, I. R., Konrad, M., Goody, R. S., Reinstein, J., and Schlichting, I. (1997) *Nat. Struct. Biol. 4*, 601−604.
21. Lavie, A., Schlichting, I., Vetter, I. R., Konrad, M., Reinstein, J., and Goody, R. S. (1997) *Nat. Med. 3*, 922−924.
22. Lavie, A., Ostermann, N., Brundiers, R., Goody, R. S., Reinstein, J., Konrad, M., and Schlichting, I. (1998) *Proc. Natl. Acad. Sci. U.S.A. 95*, 14045−14050.
23. Brundiers, R., Lavie, A., Veit, T., Reinstein, J., Schlichting, I., Ostermann, N., Goody, R. S., and Konrad, M. (1999) *J. Biol. Chem. 274*, 35289−35292.
24. Meier, C. (2002) *Mini-Rev. Med. Chem. 2*, 219−234.
25. Agarwal, K. C., Miech, R. P., and Parks, R. E., Jr. (1978) *Methods Enzymol. 51*, 483−490.
26. Kabsch, W. (1993) *J. Appl. Crystallogr. 24*, 795−800.
27. Brünger, A. T. (1993) *X-PLOR: a system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
28. Jones, T. A., Zhou, J.-Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. A47*, 110−119.
29. CCP4 (1994) *Acta Crystallogr. D50*.
30. Kinchington, D., Minshull, C., and Drummond, C. (1996) *Int. Antiviral Res. News 4*.
31. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
32. Merrit, E. A., and Murphy, M. E. P. (1994) *Acta Crystallogr. D50*, 869−873.

BI027302T