Structural Requirements for Efficient Phosphorylation of Nucleotide Analogs by Human Thymidylate Kinase

Arnon Lavie*[1] and Manfred Konrad[2]

[1]*Department of Biochemistry and Molecular Genetics, University of Illinois at Chicago, 900 S. Ashland Ave., Molecular Biology Research Building, Room 1108, Chicago, IL 60607, USA*

[2]*Max Planck Institute for Biophysical Chemistry, Department of Molecular Genetics, Am Fassberg 11, D-37077 Göttingen, Germany*

Abstract: Successive phosphorylation of nucleoside analog prodrugs to their triphosphate forms is required for the pharmacological activity of these compounds in the chemotherapeutic treatment of viral infections and cancer. Human thymidylate kinase (TMPK), apart from its essential physiological role in the biosynthesis of TTP, is also required for the activation of thymidine analogs, such as the clinically used anti-HIV prodrugs AZT and d4T. This enzyme is rate determining in the three-step cascade of AZT phosphorylation. Our structural work on human, yeast and *E. coli* TMPKs, in conjunction with sequence homology analyses and biochemical data, has demonstrated that three loops are crucial for the function of this enzyme: the first is the highly conserved P-loop motif, which binds and positions the phosphoryl groups of ATP, the second critical loop contains the DR(Y/H) motif that supplies a catalytic arginine and is also important for the binding and positioning of the magnesium ion complexed to ATP, and the third loop is the so-called Lid-region that is a flexible stretch which closes on ATP when it binds. Modifications of the sugar moieties of nucleoside monophosphates are shown to exert drastic effects on the enzyme's conformation and, thus, reduced activity. Our structural work on several TMPKs has formed the basis for generating mutants of human TMPK that are about 100 times more efficient in phosphorylating AZTMP. These enzyme variants could potentially be introduced into HIV-targeted cells in order to significantly improve AZT's antiviral activity.

INTRODUCTION

The therapeutic activity of nucleoside analog (NA) prodrugs is largely determined by the efficiency of their conversion to the active triphosphorylated form. These agents are administered as uncharged nucleosides since the charged triphosphorylated species cannot traverse biological membranes; inside the cell, NAs are converted into the monophosphate form by nucleoside kinases. For most NAs it is this conversion that is rate determining in the overall activation pathway. Examples of clinically used prodrugs for which the conversion to the monophosphate is rate limiting include d4T [1] which is phosphorylated by human thymidine kinase 1 (see review by Eriksoon), and AraC which is phosphorylated by deoxycytidine kinase [2,3]. The second phosphorylation step, catalyzed by nucleoside monophosphate kinases, appears to be critical in some cases, the most prominent example being the activation pathway of 3'-azido-3'-deoxythymidine (AZT), an additional example being the conversion of 6-thioGMP to 6-thioGDP by human guanylate kinase [4].

In contrast, the final step of NA activation, the conversion to the triphosphorylated form is usually quite an efficient process and is not rate limiting (see review by Veron). A number of enzymes can catalyze the phosphorylation of nucleoside diphosphates to triphosphates, and include the base-unspecific nucleoside diphosphate kinase (NDK), pyruvate kinase, phosphoglycerate kinase and creatine kinase [5]. NDK is a very fast enzyme with physiological substrates, but it has been shown to inefficiently phosphorylate AZT diphosphate [6]. The slow *in vitro* rate of AZT diphosphate (AZTDP) to AZT triphosphate (AZTTP) phosphorylation by NDK, but lack of *in vivo* AZTDP accumulation implies that either this step is still significantly faster than a previous step in the AZT activation pathway, or that some other enzyme(s) are responsible for the *in vivo* activation of AZTDP. In fact, it is the addition of a phosphate to AZT monophosphate (AZTMP) that is the bottle-neck of AZT activation, a reaction catalyzed by thymidylate kinase.

Thymidylate kinase (TMPK) is a nucleoside monophosphate kinase whose physiological role is to catalyze the phosphorylation of thymidine monophosphate (TMP) to thymidine diphosphate (TDP) (Fig. **1**). This enzyme is also required for the phosphorylation of several anti-HIV NAs. These include the monophosphate forms of the clinically used prodrugs d4T and AZT. Importantly, in the case of AZT, the conversion of AZTMP to AZTDP is rate limiting in the overall activation pathway of AZT [7,8]. As a result, the AZTMP species accumulates to millimolar concentration in cells exposed to AZT, while the concentration of the active triphosphorylated form remains very low. In fact, 94% of the AZT metabolites are observed as AZTMP, with AZTDP and AZTTP constituting the rest [9,10]. This low conversion of AZTMP to AZTDP limits the therapeutic efficacy of this agent, as evidenced by the modest decrease in viral load in patients undergoing AZT monotherapy. Moreover, increased viral proliferation as a result of

*Address correspondence to this author at the Department of Biochemistry and Molecular Genetics, University of Illinois at Chicago, 900 S. Ashland Ave., Molecular Biology Research Building, Room 1108, Chicago, IL 60607, USA; E-mail: lavie@uic.edu
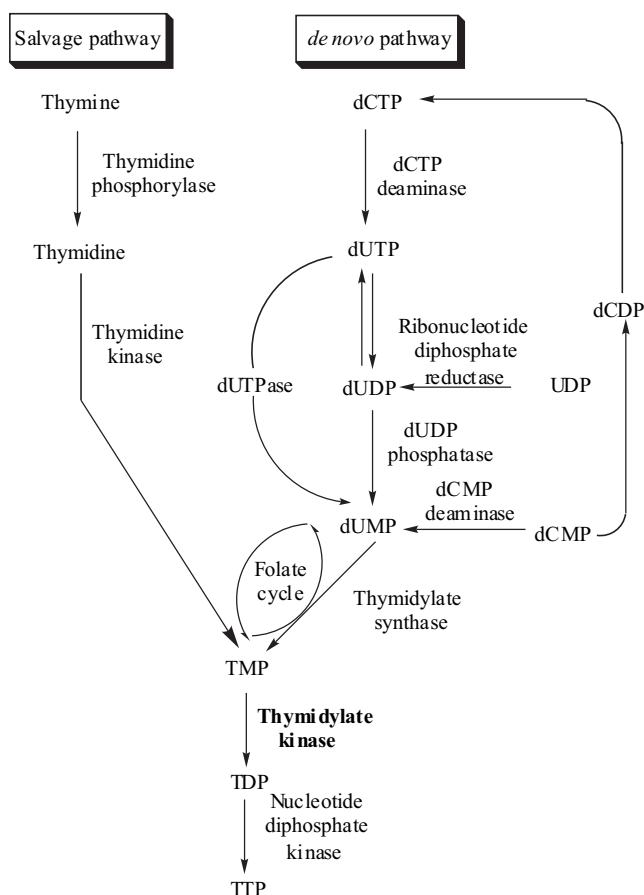
Fig. (1). Thymidylate kinase is at the junction of the *de novo* and salvage pathways of TTP synthesis. The *de novo* pathway is inhibited by the anticancer drugs 5-fluorouracil and 5-fluorodeoxyuridine as well as methotrexate, which target thymidylate synthase and dihydrofolate reductase, respectively. Prodrugs such as d4T and AZT are first phosphorylated by the salvage pathway enzyme thymidine kinase, and subsequently by thymidylate kinase. For AZT, the conversion to the diphosphate form is rate limiting.

suboptimal AZTTP concentrations allows the emergence of AZT resistant HIV strains.

Since the activation of nucleosides to their active triphosphorylated form is critical for the agent's success in a clinical setting, others and we have focused attention on the enzymes responsible for nucleoside analog phosphorylation. In this review we will concentrate on the enzyme thymidylate kinase that, in addition to being a part of the activation pathway of AZT, is critical for the activation of d4T, NH_2dT, FLT and ddT (for complete chemical names, see below). We will discuss two aspects of nucleotide analog activation by TMPK. The first will describe our understanding of the factors that determine whether a certain nucleotide analog is a good TMPK substrate. Conclusions from this section can be used for the design of novel nucleoside molecules that are not limited by poor TMPK phosphorylation for gaining their pharmacological activity. The second part of the review will take a different approach to the problem. Instead of making the substrate "fit" the enzyme, we will discuss attempts to make the enzyme "fit" our agent of choice. Specifically, we will discuss results

from mutagenesis experiments geared towards converting TMPK into a more efficient AZTMP phosphorylating enzyme. TMPK variants with enhanced AZTMP activity could be potentially used in therapeutic strategies that involve the delivery of the gene (or gene product) of the catalytically improved enzyme into HIV-infected cells, so that subsequent administration of AZT will result in a potentiated antiviral response.

TMPK SUBSTRATE SPECIFICTY

When it comes to modifying nucleosides to make potential prodrugs, the modifications are essentially limited to two areas – either modifications to the base, or modifications to the sugar moiety. Notable examples of nucleosides analogs with base substitutions are 6-thioguanine and mercaptopurine (for these compounds, guanylate kinase is a critical enzyme for activation). Since human TMPK is highly selective to the base part of the nucleotide, examples of drugs with modifications in the thymine moiety are limited. In fact, some, such as the antiviral sorivudine (1-β-D-arabinofuranosyl-E-5(2-bromovinyl) uracil), a potent inhibitor of varicella-zoster virus (VZV), is not a substrate of the human TMPK and requires activation by the viral kinase to gain pharmacological activity [11]. Likewise, the monophosphate form of (*E*)-5-(2-bromovinyl)-2'-deoxyuridine (BVDU), an inhibitor of herpes simplex virus type 1 (HSV-1) and VZV, is not phosphorylated by the human TMPK and requires conversion to its diphosphate form by the viral-encoded thymidine kinase [12]. Others, such as the base derivative 5-fluorouracil (5-FU) or the corresponding NA 5-fluorodeoxyuridine (5-FdU) are active as the monophosphate form (i.e. 5-fluoro UMP), which targets thymidylate synthase. Therefore, most clinically relevant thymidine analogs are limited to modifications in the sugar moiety. Before addressing thymidine analogs with sugar substitutions we will discuss the determinants of base specificity in TMPK.

(A) Nucleotide Specificity in TMPK

In contrast to human nucleoside kinases that phosphorylate nucleosides with different bases as substrates – e.g. deoxycytidine kinase phosphorylates dC, dG, and dA – TMPK is rather base specific. Apart from thymidine monophosphate, the only physiological nucleotide phosphorylated by TMPK is 2'-deoxyuridine monophosphate. Which structural features of the enzyme are responsible for this high selectivity?

The thymidine monophosphate binding site in TMPK is in a cavity where the base part of the nucleotide binds deeper than the sugar and phosphate groups (Fig. **2**).

The thymine base is sandwiched between Phe72 (numbering of residues correspond to human TMPK) and a face from helix 5. Here especially important is Gly102 whose Cα atom is within 3.0 Å to the carbonyl oxygen atom at C4. Thus, discrimination against purines is due to steric exclusion, as the base-binding pocket becomes more constricted deeper in the binding site. Discrimination against cytosine is achieved by a combination of two factors. One,
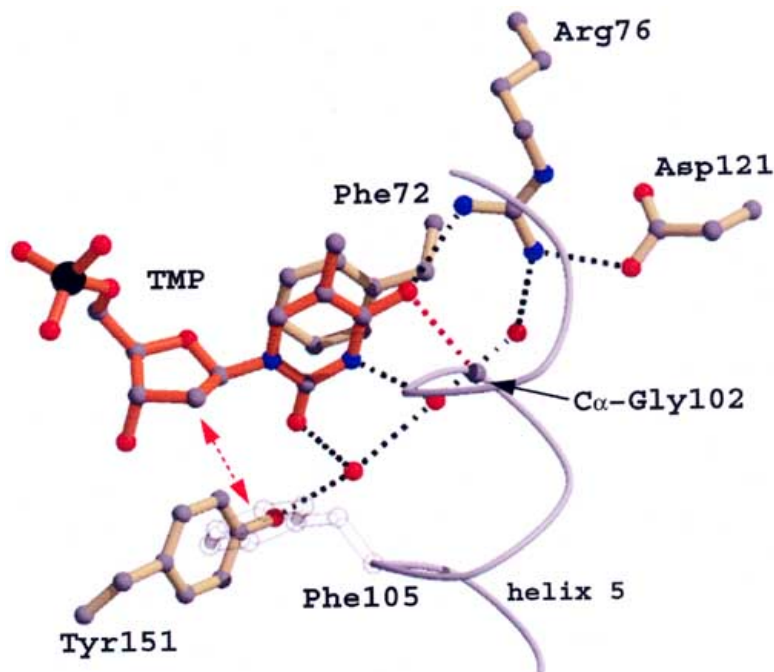
Fig. (2). Base specificity in human TMPK. In this ball-and-stick representation, the thymine base of TMP (orange) is bound between Phe72 and a face made by helix 5. The Cα-atom of Gly102 makes the binding of purine bases sterically disfavored, as for cytosine binding. Arg76 also contributes to the discrimination against cytosine. The methyl group at C5 makes no protein interactions, consistent with dUMP being an efficient TMPK substrate. Structural figures were generated with Bobscript [42] and Raster3D [43].

the above-mentioned short distance between the carbonyl at C4 and Gly102Cα: a NH2 at that position, as present in cytosine, will result in a very close contact. Two, the productive hydrogen bond interaction between Arg76 and the C4 carbonyl will not take place as efficiently with an NH$_2$ at that position. Moreover, this arginine side chain is held rigid in place by a hydrogen-bonding network that includes Asp121, limiting its ability to turn away in the presence of cytosine (Fig. **2**). Since the C5-methyl group is not at close proximity to any protein atom, it is not surprising that TMPK cannot, solely through the base, discriminate between uracil and thymine.

Discrimination against ribonucleotides is due to a close contact between the side chain of Phe105 and the sugar ring. The aromatic ring will sterically exclude a 2'-hydroxyl. As a result, UMP is not accepted as a substrate by TMPK, but dUMP is phosphorylated with good efficiency [13].

(B) Sugar-modified Thymidine Analogs

Several clinically used NAs are the sugar-modified thymidine analogs. The majority of sugar-modifications center on the 3'-position as the hydroxyl at this position is vital for nucleotide polymerization in the process of DNA chain synthesis. In the case of the chain terminator AZT, in which the 3'-hydroxyl is substituted by the bulky, nonreactive azido group, the rate of AZTMP phosphorylation by human TMPK is decreased ~60-fold in comparison to the rate of TMP phosphorylation [14]. However, smaller substituents at the 3'-position do not necessarily make for better TMPK substrates: a fluorine atom at the 3'-position still reduces the rate by 12-fold; yet

an amino group substitution reduces the rate by less than a factor of three. So, which factors determine the activity for 3'-subsitituted thymidine analogs? Since there is no direct correlation between size of a substituent at the 3'-position and the phosphorylation rate by TMPK, and the K_M values for many of the analogs are very similar, a simple binding effect cannot explain the distinct kinetic behavior of such analogs. To gain understanding of the effect of 3'-substitutions on the TMPK phosphorylation rate, we commenced a structural study of the enzyme with various 3'-modified thymidine analogs. These encompassed 3'-azido-3'-deoxythymidine monophosphate (AZTMP), 3'-amino-3'-deoxythymidine monophosphate (NH$_2$TMP), 3'-fluoro-3'-deoxythymidine monophosphate (FLTMP), 2',3'-didehydro-2',3'-dideoxythymidine monophosphate (d4TMP), and dideoxythymidine monophosphate (ddTMP), (Fig. **3**).

(1) AZTMP

Early studies on yeast TMPK that compared a complex of the enzyme with TMP to that with AZTMP suggested that the azido group affects the positioning of the critical phosphate binding loop (P-loop) [15,16]. This loop functions to bind the phosphoryl donor (i.e. ATP); but unique to all TMPKs is a carboxylic acid residue in this loop that interacts with the phosphoryl acceptor *via* the 3'-hydroxyl group. Furthermore, to make the interaction between the 3'-hydroxyl and the side chain of the carboxylic acid residue, the P-loop needs to change its conformation and shift towards the phosphoryl acceptor. The crystal structure of the yeast TMPK-AZTMP complex suggested that such a conformational change is hindered by the 3'-azido group of AZTMP.
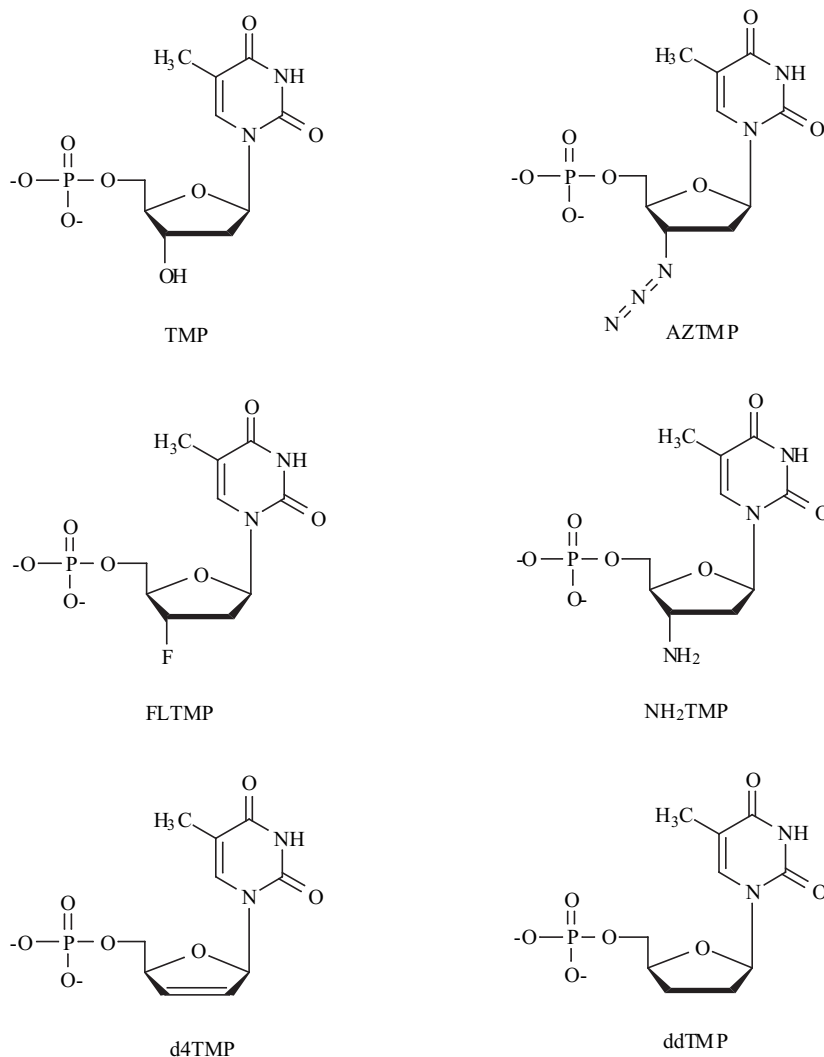
Fig. (3). Structures of ribose-modified TMPK substrates. Kinetic and structural characterization of these molecules as TMPK substrates reveal the requirements for efficient enzymatic phosphorylation (see text for details).

Subsequent work with the human TMPK supported the conclusions derived from the yeast TMPK structures [17]. Importantly, the human TMPK structures are of much higher resolution than the yeast TMPK structures, and include a nucleotide molecule bound to the phosphoryl donor site (either ADP or the ATP analog adenosine β,γ-imido-5'-triphosphate (AppNHp)) and the essential magnesium ion. A surprise was the observation of different conformations for the P-loop when comparing the complex of TMP-ADP to that of TMP-AppNHp. An open conformation was observed for the P-loop in the enzyme-TMP-ADP complex such that the P-loop carboxylic acid was too distant to interact with the 3'-hydroxyl. In contrast, in the TMP-AppNHp complex, the P-loop changed its position moving closer to the TMP molecule. As a result of this conformational change, the carboxylic acid could interact with the TMP 3'-hydroxyl group. This position of the P-loop was termed the partially-closed conformation, as further closing of the loop was observed in certain complex structures, and this conformation was termed fully closed. Only the partially-closed and the fully-closed P-loop conformations were interpreted to represent an active enzyme state. In other

words, the hypothesis was made that during the reaction, a closing of the P-loop is an essential event for catalysis.

The analogous structures in which the enzyme was complexed with AZTMP instead of TMP reveled a P-loop exclusively in the open conformation. Modeling of AZTMP in the partially-closed TMP-AppNHp complex structure predicted a steric clash between the 3'-azido group and the P-loop carboxylic acid. Therefore, as concluded from analysis of the yeast TMPK, AZTMP is a poor substrate for the human TMPK due to steric repulsion between the azido group and the essential P-loop carboxylic acid.

(2) NH₂TMP

The prodrug AZT can undergo a side-reaction to yield 3'-aminothymidine. The triphosphorylated form, NH₂TTP, is a substrate of human DNA polymerases and a possible source of toxicity observed with AZT therapy [18,19]. Kinetic examination of NH₂TMP showed it to be 14-fold a better substrate than AZTMP for the human TMPK. Structural studies of the NH₂TMP-AppNHp complex of TMPK revealed a P-loop in the fully closed conformation [20]. The 3'-amino group interacts with the P-loop carboxylic acid, as

expected. However, the observed fully closed P-loop conformation, in contrast to the partially-closed conformation as was observed in the TMP-AppNHp complex, was a surprise. This could be explained by a stronger interaction between a positively charged amino-group and the negatively charged side-chain of the carboxylic acid. Since P-loop dynamics seems to play a pivotal role in TMPK activity, the strong 3'-amino – carboxylic acid interaction is speculated to ultimately reduce the NH_2TMP phosphorylation rate in comparison to TMP.

(3) d4TMP

Like AZT, d4T is clinically used to treat HIV infections. The rate limiting step of d4T activation is its conversion to d4TMP by cellular thymidine kinase [1]. Consistent with this is a relatively efficient conversion of d4TMP to d4TDP by thymidylate kinase, with a rate one fourth of that with the physiological substrate TMP [20]. Structures of TMPK in complex with d4TMP reveal that this nucleotide analog promotes P-loop closure to the partially-closed conformation. How is this accomplished? The answer seems to lie in the position of the nucleotide analog's phosphate group. Due to the double bond between C2' and C3', the phosphate group of d4TMP is not at the same location as observed in structures with nucleotides having a saturated ribose. As a result, the phosphate group has a weakened interaction with the conserved Arg97 of the DRY/H motif. This arginine is observed to interact with the P-loop's carboxylic acid in the partially-closed and closed conformations (but not in the open state). Thus, due to the weakened Arg97-phosphate interaction in the case of d4TMP, the arginine side-chain is more accessible to the P-loop carboxylic acid residue in the closed states, thereby promoting P-loop closure.

(4) ddTMP

Testing of the above hypothesis, which correlates between the position of the nucleotide analog's phosphate group and the P-loop conformation, was accomplished by solving the structures of TMPK in complex with ddTMP. Like d4TMP, ddTMP has no 2'-and 3'-substituents, and the only difference between these TMP analogs lies in the sugar conformation. Steady-state kinetic experiments showed ddTMP to be a very poor substrate of TMPK, only 3-fold faster than AZTMP [20]. The structures of TMPK in complex with ddTMP revealed that the partially-closed P-loop conformation is disadvantaged in comparison to analogous structures with TMP. Opposite to the situation with d4TMP, the position of the ddTMP phosphate group seems to hinder P-loop closure. While for d4TMP the breaking of the Arg97-phosphate group promoted P-loop closure, in the case of ddTMP, the Arg97-phosphate interaction is strengthened. As a result, Arg97 is less available to interact with the P-loop's carboxylic acid, making the partially-closed P-loop conformation less stable. Hence, the low rate of ddTMP phosphorylation by TMPK is attributed to the destabilization of the partially-closed P-loop conformation as a result of a strengthened interaction between Arg97 and the phosphate group of ddTMP.

(5) FLTMP

For this compound, in which a fluorine atom replaces the 3'-hydroxyl group, the TMPK-catalyzed conversion to the diphosphate form is rate limiting. *In vitro* kinetic data show FLTMP to be a poor TMPK substrate [20]. Structural studies of TMPK in complex with FLTMP were performed to understand the root for the slow phosphorylation of this nucleotide analog. However, in this case, no obvious differences between analogous structures of TMP and FLTMP were observed.

(6) Summary

Consideration of all the available TMPK complex structures allows for the deduction of general principles that determine phosphorylation activity. The conformation of the P-loop is critical for activity. Importantly, the nature of the nucleotide analog directly affects the ability of the P-loop to change its conformation from the inactive open state to the partially-closed and fully-closed active states. A large substituent at the 3'-position, such as an azido group, is sterically not compatible with the closed states. Lack of any substituent at the 3'-position does not necessarily translate into a good substrate: the sugar conformation affects the position of the analog's phosphate group, which in turn can either promote (e.g. d4TMP) or inhibit (e.g. ddTMP) P-loop closure. A small and positively charged substituent is conducive to activity by strengthening the direct interaction between the 3'-substituent and the essential P-loop carboxylic acid residue. On the other hand, an electronegative substituent, particularly fluorine, reduces the phosphorylation rate. Thus, it is a combination of the 3'-substituent's size and hydrogen bonding abilities, in addition to its effect on the sugar ring conformation, that determines the efficiency of TMP analogs phosphorylation by TMPK.

(C) Acyclic-sugar and Conformationally-locked Thymidine Analogs

The acyclic-sugar guanosine analogs acyclovir and ganciclovir have proven very effective antiviral agents [21]. It was speculated that acyclic-sugar thymidine monophosphate analogs might be efficiently phosphorylated by TMPK since the flexible sugar moiety would not hinder P-loop closure. Three such analogs (Fig. **4**) were prepared by the Meier group [22] and given to the Reinstein group (Max-Planck Institute for Molecular Physiology, Dortmund, Germany) for testing [23].

The thymidine monophosphate analogs of ganciclovir and penciclovir, 9-[1,3-dihydroxy-2-propoxymethyl] thymine monophosphate (T-'GCV'-MP) and 9-[4-hydroxy-3-hydroxy-methylbut-1-yl] thymine monophosphate (T-'PCV'-MP), respectively, are converted by TMPK to their diphosphate form, but at a lower rate than AZTMP. The difference in nature of the atom, an oxygen or a carbon, in the position that would correspond to the ribose ether oxygen O4', between T-'GCV'-MP and T-'PCV'-MP, seems not to affect the rate to a great extent, whereas the K_M of T-'GCV'-MP is more than 50-fold higher. With the acyclic analog containing an azido group, AZT-'PCV'-MP, no appreciable turnover was detected. Structural data on these compounds is still not available, but based on experience with cyclic-sugar TMP analogs (see above), we can speculate that the low rate of phosphorylation of the acyclic-sugar TMP analogs is due to mispositioning of the analog's phosphate group (while the
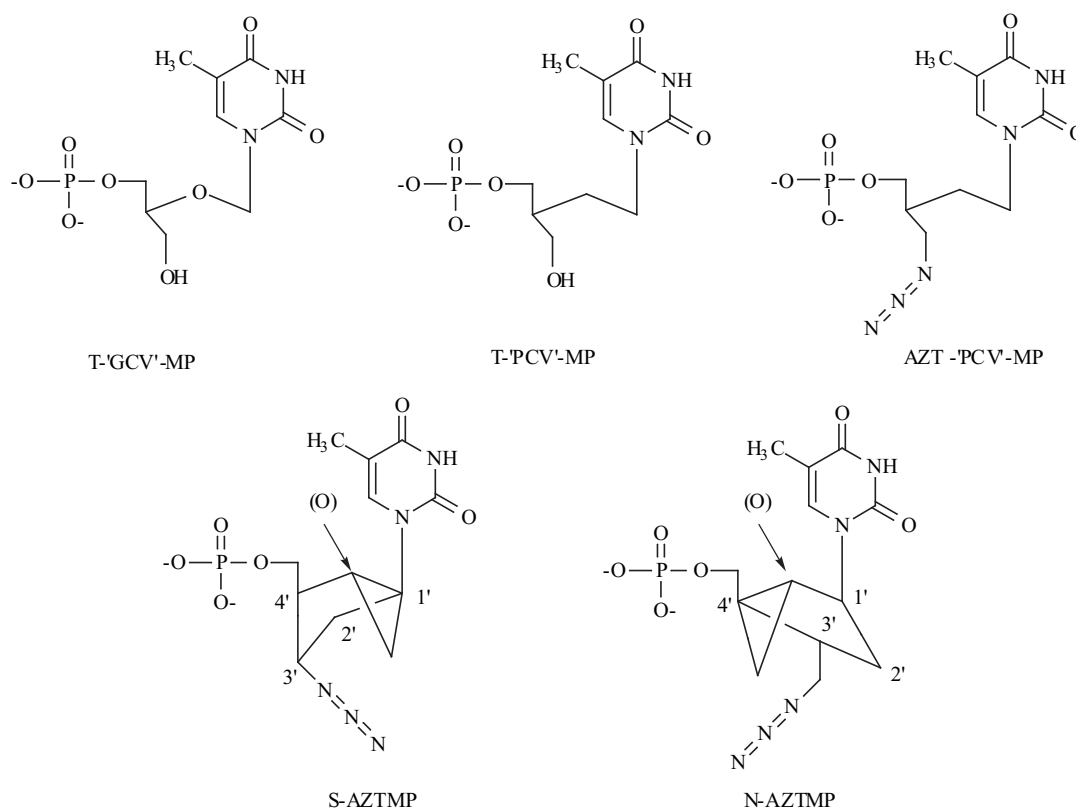
Fig. (4). Structures of acyclic-sugar (analogous to the GMP analogs ganciclovir, 'GCV', and penciclovir, 'PCV') and conformationally-locked (designated Northern, N, and Southern, S) TMP analogs tested as potential substrates for human TMPK. For the latter, the position in the ribose ring normally occupied by O4' is indicated by an arrow.

thymine moiety is expected to bind in an identical fashion as seen with TMP). Correct positioning of this phosphate is critical for promoting closure of the P-loop, a prerequisite for catalytic activity, as well as being essential for a productive geometry for accepting the phosphoryl group from ATP.

With the goal of discovering AZT analogs that are potent HIV inhibitors, analogs with a locked sugar pucker were synthesized [24]. One such compound, N-AZTMP (2'-*exo*/3'-*endo*, also referred to as a Northern conformation), in its triphosphorylated form is a strong HIV reverse transcriptase inhibitor [25]. Unfortunately, both Northern and Southern (2'-*endo*/3'-*exo*) locked AZTMP analogs [22] (Fig. **4**) are phosphorylated at a very low rate by human TMPK [23]. This is a good example for a compound that in principle is a promising drug candidate, but that is limited in its efficacy by inefficient prodrug-to-drug conversion. The importance of the sugar pucker for TMPK-dependent phosphorylation is evident from the results obtained with d4TMP and ddTMP. Since the analog's phosphate group must be able to change its conformation to interact Arg97 and with Arg45 [17], it is possible that by locking the sugar pucker into a rigid conformation, and hence the position of the phosphate group, this required mobility was lost.

TMPK ENZYME ENGINEERING

The rationale for studying the structure-activity relationships of NAs and the enzyme TMPK at the molecular level are two-fold: One, to attain guiding principles for the design of new agents that are well activated by this enzyme (see above),. Two, and the focus of this section, to use this information to develop TMPK variants that possess enhanced activity towards a particular, medicinally used prodrug. Improved enzymes could be delivered to cells *via* a number of strategies; some deliver the gene itself, as in the so-called gene-directed, enzyme-prodrug therapy (GDEPT; for a review see [26]) or antibody-directed, enzyme-prodrug therapy (ADEPT; for a review see [27]), while some aim at delivering the gene product itself. Previous enzyme engineering work with the aim of improving AZT activation focused on herpes simplex type thymidine kinases and utilized a genetic approach including DNA family shuffling and selection in a thymidine kinase-defective bacterial strain [28]. In that work, the phosphorylation rate of AZT was not increased, but its K_M was decreased, enabling the variants to sensitize *E. coli* to 32-fold less AZT in comparison to cells expressing HSV1-TK. We chose a different approach to the problem of improving AZT activation and making the rate determining enzyme more selective for the nucleoside analog. We focused on the human enzyme TMPK, and utilized our structural knowledge to guide the enzyme engineering work.

Though AZT holds great potential as an antiviral agent, its therapeutic efficacy is limited due to the inefficient, TMPK-mediated conversion of AZTMP to AZTDP. As discussed before, the P-loop in the enzyme is prevented by the 3'-azido group of AZTMP from adopting the necessary closed and active conformation [15,29]. With the aim of developing TMPK variants that are not hindered by the 3'-
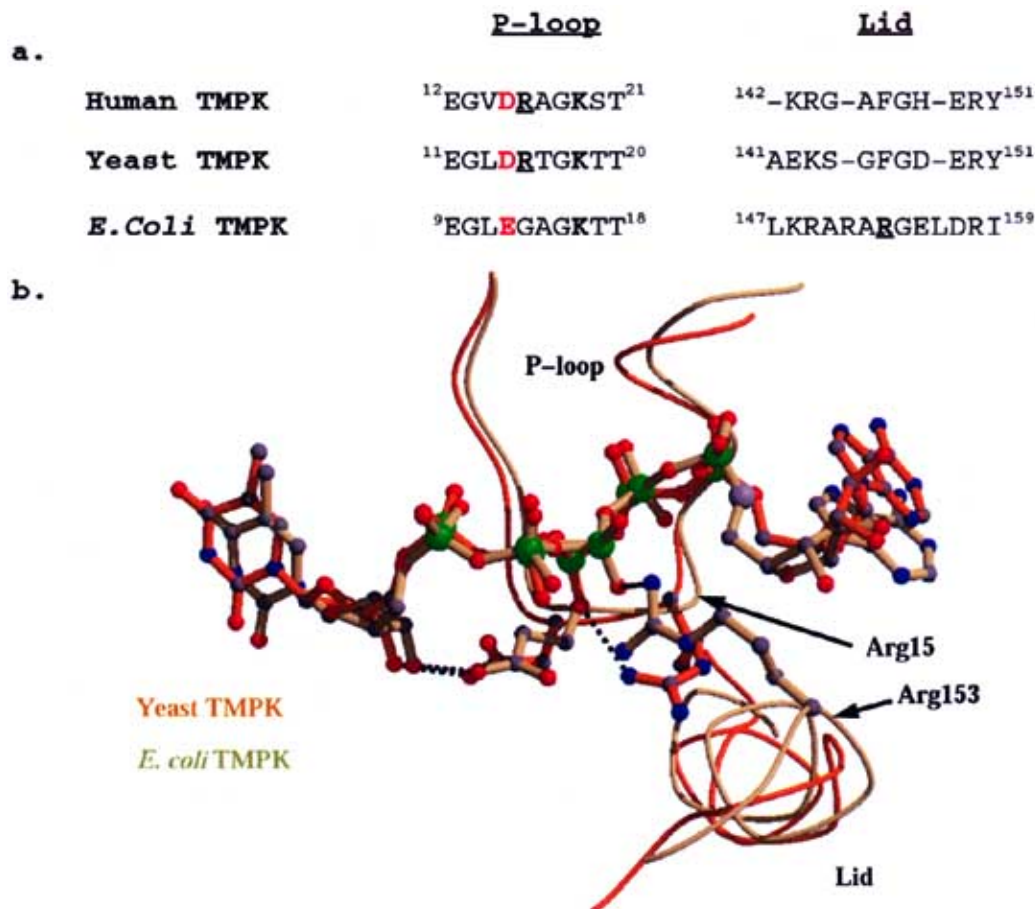
Fig. (5). Differences between eukaryotic and bacterial TMPKs. (a) Only eukaryotic TMPKs possess an arginine residue in their P-loop sequence, whereas bacterial TMPKs code for the functionally equivalent residue in the Lid sequence (*underlined* R). In common for all TMPKs is the conserved and essential P-loop carboxylic acid residue (in *red*). (b) Overlay of yeast TMPK (*orange*) and *E. coli* TMPK (*beige*) TP$_5$A complex structures. Spatially, the side chain of the P-loop Arg15 of the yeast enzyme is close to the *E. coli* Lid arginine Arg153, suggesting an equivalent function for these residues. Note the interactions made by the P-loop carboxylic acid residues with the atom equivalent to the TMP 3'-hydroxyl group.

azido group, analysis of the catalytic mechanism was undertaken to try to understand the essential role of the P-loop in its closed conformation. Several mechanisms for phosphoryl transfer can be envisioned, and at two extremes include the dissociative mechanism and the associative mechanism [30]. For uridylate kinase, an enzyme structurally similar to TMPK and catalyzing a similar reaction (phosphorylation of UMP to UDP), it has been proposed that phosphoryl transfer is associative [31]. Assuming such a mechanism for TMPK implies a transition state that is characterized by a higher negative charge in comparison to the ground state. In this case, the question arises as to how the enzyme would stabilize this additional charge so as to lower the energy barrier for the reaction. Likely candidates are positively charged arginine residues that are invariantly observed in kinase active sites. For UMPK, arginine residues from the Lid region were proposed to stabilize the formed negative charge during the reaction [31]. The Lid region, as the P-loop, is a conserved structural element among nucleoside/nucleotide kinases, despite overall the little sequence identity between these enzymes. Interestingly, the Lid region of eukaryotic TMPKs does not

include these arginines (Fig. **5a**). At the same time, a unique feature of eukaryotic TMPKs is an arginine in the P-loop sequence at a position just following the conserved P-loop carboxylic acid that interacts with the TMP 3'-hydroxyl as described above.

We made the hypothesis that the reason for poor AZTMP phosphorylation by eukaryotic TMPKs is due to a mispositioning of the P-loop arginine residue, as a result of the P-loop being prevented from adopting the partially-closed and closed conformations. In support of this was the observation that in *E. coli* TMPK, which does not have this arginine in the P-loop but rather possesses a number of arginine residues in the Lid region, AZTMP phosphorylation was only 2-fold slower than the TMP phosphorylation rate [32] (this factor is 200 for the yeast TMPK, 60 for the human enzyme). Structural studies on the bacterial TMPK buttress this model. An overlay of the relevant arginine-containing regions of yeast TMPK (arginine in P-loop) and of *E. coli* TMPK (arginine in Lid) reveals that the spatial position of the arginine side chains that originate either from the P-loop or the Lid is nearly identical, and the chains are

similarly interacting with the phosphate groups (Fig. **5b**). Therefore, we made the prediction that the P-loop arginine in eukaryotic TMPKs plays a very similar role to that of the Lid arginine in bacterial TMPKs. In addition, we proposed that the efficient phosphorylation of AZTMP by the *E. coli* TMPK is due to this factor; any inhibition of P-loop closure in the bacterial enzyme due to the 3'-azido group will not perturb the positioning of the catalytic Lid arginine. This model suggested a solution to overcome the poor AZTMP phosphorylation by site-specific engineering of human TMPK, namely by transferring the catalytic arginine from the P-loop sequence to the Lid region.

Simple substitution of residues, i.e. the concomitant change of the P-loop arginine to a glycine to make room for an arginine introduced in the Lid region, did not yield the desired improved catalytic rate. Only when the whole *E. coli* Lid sequence was transferred as a complete unit to replace the Lid sequence in the human enzyme did a variant with the desired properties emerge [14]. Amazingly, the variant labeled *E. coli* large-Lid (Tab. 1) is twice as fast with the physiological substrate TMP [14]. It is very rare to produce a mutant that is even faster than wild type with the physiological substrate. Even more amazing was the nearly 200-fold increase in the AZTMP phosphorylation rate, to the extent that this TMPK variant is more efficient at this prodrug phosphorylation than that of TMP.

Other improved TMPK variants were also produced based on structural considerations. The single Phe105Tyr mutation resulted in a slightly slower enzyme that wild type TMPK, but with higher AZTMP phosphorylation rate than that of TMP (Tab. 1). The impetus to test this mutation originated from comparison of the human and yeast TMPKs. The more active yeast TMPK has a tyrosine in this position, equivalent to the phenylalanine in the human enzyme. Structural characterizations of this mutant suggested that the improved AZTMP rate is due to the increased propensity of the mutant to adopt the partially-closed P-loop conformation. We speculated that the introduced hydroxyl group in the tyrosine residue acts to perturb the water structure that serves to stabilize the open P-loop conformation, and as a result, the closed P-loop conformation becomes energetically favored [29]. Both the *E. coli* Large-Lid and the Phe105Tyr mutants are promising enzymes to be used in gene/enzyme therapy applications with the aim of increasing the concentration of the active triphosphorylated form of AZT.

CONCLUSION AND PERSPECTIVE

The location of TMPK at the junction of the *de novo* and the salvage pathways of TTP biosynthesis renders this enzyme essential for DNA replication. Its fundamental role in nucleotide metabolism is emphasized by the fact that a knock-out of the single TMPK encoding gene in yeast is lethal [33] and TMPK genes have been identified in all recently sequenced genomes, including the smallest microbial genomes, which are thought to consist of a minimal set of genes required for cellular life [34,35]. In mammals, a correlation has been demonstrated between TMPK activity and the rate of DNA synthesis during rapid growth of e.g. regenerating tissue and tumor cells [36], and, therefore, selective inhibition of TMPK could be very effective in blocking DNA synthesis in rapidly dividing cells. Human TMPK, which is cytosolic and encoded by a single gene, may thus be a selective target for cancer chemotherapy. A prerequisite for the rational design of anti-cancer drugs is a detailed understanding of the structural requirements of substrate binding to TMPK and of the catalytic mechanism of phosphoryl transfer at the molecular level. Our structural work on the human enzyme, which we have analyzed in complex with a series of natural nucleotides and medicinally relevant nucleotide analogs, has laid the basis for understanding the reasons of the enzyme's pronounced substrate specificity, and also for explaining the poor metabolic activation of clinically well-approved antiviral drugs such as AZT and d4T. These studies, combining structural and kinetic data, provide a rationale for the design of novel compounds that could be more efficiently phosphorylated by human TMPK and, thus, may act as more potent drugs for combating viral diseases. Moreover, detailed knowledge of the human enzyme allows us to define the unique structural and catalytic features that distinguish this enzyme from its bacterial counterparts. This is particularly relevant in light of recent efforts to evaluate novel strategies of chemotherapeutic intervention of infectious diseases such as tuberculosis. Here, attempts are being made to design nucleoside-based compounds that act as inhibitors of bacterial TMPK [37]. Ideally, these analogs should be inert on human nucleoside and nucleotide kinases in order to show a high therapeutic index by not interfering with normal cellular nucleotide metabolism.

Our effort of generating rationally designed variants of human TMPK that are able to efficiently phosphorylate

Table 1. **Steady-State Phosphorylation Rates for Wild-Type (WT) and Mutants of Human TMPK.** The 60-fold reduced rate of AZTMP in comparison to TMP by WT-TMPK defines the challenge that the TMPK variants were designed to overcome. The *E. coli* small-lid designates variants in which four amino acids from the human TMPK were replaced by six from the *E. coli* sequence. In the large-lid constructs, 13 residues were replaced by 15. Both the small-lid and large-lid constructs are active towards TMP and AZTMP, but only the large-lid displays the dramatic reversal in substrate specificity, where AZTMP is a better substrate than TMP. A change in substrate specificity is also observed with the Phe105Tyr mutant, albeit with a lower absolute rate. No synergy was observed upon combining the large-lid and Phe105Tyr point mutation

	k_{obs} for TMP (sec^{-1})	k_{obs} for AZTMP (sec^{-1})
WT	0.7	0.012
WT+ *E. coli* Small-Lid	0.15	0.003
WT+ *E. coli* Large-Lid	0.17	0.002
Arg16Gly + *E. coli* Small-Lid	0.7	0.22
Arg16Gly + *E. coli* Large-Lid	1.5	2.14
Phe105Tyr	0.17	0.25
Arg16Gly + *E. coli* Large-Lid + Phe105Tyr	0.33	0.60

AZTMP may open a way to significantly improve the intracellular phosphorylation of the anti-HIV prodrug AZT. Since it was shown that the bottle-neck in AZT activation is the addition of the second phosphoryl group catalyzed by cellular TMPK [1,7,16], leading to the intracellular accumulation of cytotoxic AZTMP [38, 39], we envision gene transfer into HIV-targeted cells of a catalytically improved human TMPK variant that has the potential to enhance both the efficacy and selectivity of the prodrug AZT. Indeed, in human lymphoid cell culture assays using radiolabeled AZT it was shown that expression of HSV-1 TK, which also has thymidylate kinase activity and converts AZTMP to AZTDP much more efficiently than human TMPK, leads to significantly higher levels of AZTTP, the antivirally active metabolite. Importantly, this enhanced AZT phosphorylation in genetically modified cells translates into improved efficacy of AZT in blocking HIV replication [40]. In suitable clinical applications, ultimately, this approach, which is also called gene-directed enzyme prodrug therapy, or genetic pharmacomodulation, may help to suppress virus replication drastically, or at least prolong the interval of emergence of drug-resistant viruses. Furthermore, expressing modified human enzymes instead of enzymes of viral or bacterial origin could be advantageous since an immune response [41] to the introduced human protein appears less likely. Also, it might be hoped that this technique will foster the continued development of (pro)drugs that have been abandoned because of poor metabolism to their pharmacologically active derivatives in target cells.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Balzarini, J.; Herdewijn, P.; De Clercq, E. *J. Biol. Chem.* **1989**, *264*, 6127-6133.

[2]   Momparler, R.L.; Fischer, G.A. *J. Biol. Chem.* **1968**, *243*, 4298-4304.

[3]   Chabner, B.A.; Allegra, C.J.; Curt, G.A.; Calabresi, P. Antineoplastic agents. *Goodman & Gilman's The Pharmacological Basis of Therapeutics*; 9th ed.; McGraw-Hill, **1996**; pp 1251-1252.

[4]   Crabtree, G.W.; Nelson, J.A.; Parks, R.E., Jr. *Biochem. Pharmacol.* **1977**, *26*, 1577-1584.

[5]   Miller, W.H.; Daluge, S.M.; Garvey, E.P.; Hopkins, S.; Reardon, J.E.; Boyd, F.L.; Miller, R.L. *J. Biol. Chem.* **1992**, *267*, 21220-21224.

[6]   Bourdais, J.; Biondi, R.; Sarfati, S.; Guerreiro, C.; Lascu, I.; Janin, J.; Veron, M. *J. Biol. Chem.* **1996**, *271*, 7887-7890.

[7]   Furman, P.A.; Fyfe, J.A.; St, C.M.; Weinhold, K.; Rideout, J.L.; Freeman, G.A.; Lehrman, S.N.; Bolognesi, D.P.; Broder, S.; Mitsuya, H.; *et al. Proc. Natl. Acad. Sci. USA* **1986**, *83*, 8333-8337.

[8]   Balzarini, J.; Pauwels, R.; Baba, M.; Herdewijn, P.; De Clercq, E.; Broder, S.; Johns, D.G. *Biochem. Pharmacol.* **1988**, *37*, 897-903.

[9]   Frick, L.W.; Nelson, D.J.; St. Clair, M.H.; Furman, P.A.; Krenitsky, T.A. *Biochem. Biophys. Res. Commun.* **1988**, *154*, 124-129.

[10]   Fridland, A.; Connelly, M.C.; Ashmun, R. *Mol. Pharmacol.* **1990**, *37*, 665-670.

[11]   Yokota, T.; Konno, K.; Mori, S.; Shigeta, S.; Kumagai, M.; Watanabe, Y.; Machida, H. *Mol. Pharmacol.* **1989**, *36*, 312-316.

[12]   Balzarini, J.; McGuigan, C. *Biochim. Biophys. Acta* **2002**, *1587*, 287-295.

[13]   Jong, A.Y.; Campbell, J.L. *J. Biol. Chem.* **1984**, *259*, 14394-14398.

[14]   Brundiers, R.; Lavie, A.; Veit, T.; Reinstein, J.; Schlichting, I.; Ostermann, N.; Goody, R.S.; Konrad, M. *J. Biol. Chem.* **1999**, *274*, 35289-35292.

[15]   Lavie, A.; Vetter, I.R.; Konrad, M.; Goody, R.S.; Reinstein, J.; Schlichting, I. *Nature Struct. Biol.* **1997**, *4*, 601-604.

[16]   Lavie, A.; Schlichting, I.; Vetter, I.R.; Konrad, M.; Reinstein, J.; Goody, R.S. *Nature Medicine* **1997**, *3*, 922-924.

[17]   Ostermann, N.; Schlichting, I.; Brundiers, R.; Konrad, M.; Reinstein, J.; Veit, T.; Goody, R.S.; Lavie, A. *Structure Fold Des.* **2000**, *8*, 629-642.

[18]   Chidgeavadze, Z.G.; Beabealashvilli, R.S.; Atrazhev, A.M.; Kukhanova, M.K.; Azhayev, A.V.; Krayevsky, A.A. *Nucleic Acids Res.* **1984**, *12*, 1671-1686.

[19]   Jasko, M.V.; Fedorov, I.I.; Atrzhev, A.M.; Mozzherin, D.Y.; Novicov, N.A.; Bochkarev, A.V.; Gurskaya, G.V.; Krayevskt, A.A. *Nucleosides and Nucleotides* **1995**, *14*, 23-37.

[20]   Ostermann, N.; Segura-Pena, D.; Meier, C.; Veit, T.; Monnerjahn, C.; Konrad, M.; Lavie, A. *Biochemistry* **2003**, *42*, 2568-2577.

[21]   De Clercq, E. *J. Clin. Virol.* **2001**, *22*, 73-89.

[22]   Meier, C.; Mueller, C. *University of Hamburg*, unpublished results.

[23]   Veit, T. Kinetischer Mechanismus der menschlichen Thymidylatekinase - der induced fit als Grund für die unzureichende Aktivierung des Anti-HIV Medikaments 3'-Azido-2'-3'-dideoxy thymidinemonophshat (AZTMP). In *Chemistry Department*; University of Dortmund: Dortmund, **2000**.

[24]   Marquez, V.E.; Siddiqui, M.A.; Ezzitouni, A.; Russ, P.; Wang, J.; Wagner, R.W.; Matteucci, M.D. *J. Med. Chem.* **1996**, *39*, 3739-3747.

[25]   Marquez, V.E.; Ezzitouni, A.; Russ, P.; Siddiqui, M.A.; Ford, H.; Feldman, R.J.; Mitsuya, H.; George, C.; Barchi, J.J. *J. Amer. Chem. Soc.* **1998**, *120*, 2780-2789.

[26]   Springer, C.J.; Niculescu-Duvaz, I. *Adv. Drug Delivery Reviews* **1996**, *22*, 351-364.

[27]   Niculescu-Duvaz, I.; Springer, C.J. *Adv. Drug Delivery Reviews* **1997**, *26*, 151-172.

[28]   Christians, F.C.; Scapozza, L.; Crameri, A.; Folkers, G.; Stemmer, W.P. *Nat. Biotechnol.* **1999**, *17*, 259-264.

[29]   Ostermann, N.; Lavie, A.; Padiyar, S.; Brundiers, R.; Veit, T.; Reinstein, J.; Goody, R.S.; Konrad, M.; Schlichting, I. *J. Mol. Biol.* **2000**, *304*, 43-53.

[30]   Mildvan, A.S. *Proteins* **1997**, *29*, 401-416.

[31]   Schlichting, I.; Reinstein, J. *Biochemistry* **1997**, *36*, 9290-9296.

[32]   Lavie, A.; Ostermann, N.; Brundiers, R.; Goody, R.S.; Reinstein, J.; Konrad, M.; Schlichting, I. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14045-14050.

[33]   Sclafani, R.A.; Fangman, W.L. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 5821-5825.

[34]   Fraser, C.M.; Gocayne, J.D.; White, O.; Adams, M.D.; Clayton, R.A.; Fleischmann, R.D.; Bult, C.J.; Kerlavage, A.R.; Sutton, G.; Kelley, J.M.; *et al. Science* **1995**, *270*, 397-403.

[35]   Tamas, I.; Klasson, L.; Canback, B.; Naslund, A.K.; Eriksson, A.S.; Wernegreen, J.J.; Sandstrom, J.P.; Moran, N.A.; Andersson, S.G. *Science* **2002**, *296*, 2376-2379.

[36]   Maehara, Y.; Moriguchi, S.; Emi, Y.; Watanabe, A.; Kohnoe, S.; Tsujitani, S.; Sugimachi, K. *Cancer* **1990**, *66*, 156-161.

[37]   Haouz, A.; Vanheusden, V.; Munier-Lehmann, H.; Froeyen, M.; Herdewijn, P.; Van Calenbergh, S.; Delarue, M. *J. Biol. Chem.* **2003**, *278*, 4963-4971.

[38]   Törnevik, Y.; Ullman, B.; Balzarini, J.; Wahren, B.; Eriksson, S. *Biochem. Pharmacol.* **1995**, *49*, 829-837.

[39]   Sales, S.D.; Hoggard, P.G.; Sunderland, D.; Khoo, S.; Hart, C.A.; Back, D.J. *Toxicol. Appl. Pharmacol.* **2001**, *177*, 54-58.

[40]   Guettari, N.; Loubiere, L.; Brisson, E.; Klatzmann, D. *Virology* **1997**, *235*, 398-405.

[41]   Dewey, R.A.; Morrissey, G.; Cowsill, C.M.; Stone, D.; Bolognani, F.; Dodd, N.J.; Southgate, T.D.; Klatzmann, D.; Lassmann, H.; Castro, M.G.; Lowenstein, P.R. *Nat. Med.* **1999**, *5*, 1256-1263.

[42]   Esnouf, R. *J. Mol. Graphics* **1997**, *15*, 133-138.

[43]   Merrit, E.A.; Murphy, M.E.P. *Acta. Crystallogr.* **1994**, *D50*, 869-873.