



Mutagenesis of non-conserved active site residues improves the activity and narrows the specificity of human thymidine kinase 2

Monica L. Gerth, Stefan Lutz *

Chemistry Department, Emory University, Atlanta, GA 30322, USA

Received 29 December 2006

Abstract

Human thymidine kinase 2 (TK2) is critical for the nucleotide salvage pathway and phosphorylation of nucleoside analog prodrugs *in vivo*; however, it remains poorly studied because of difficulties in expressing it heterologously. TK2 is strictly pyrimidine-specific, whereas its phylogenetic relative, the *Drosophila melanogaster* deoxyribonucleoside kinase (*DmdNK*), shows higher activity and broader specificity towards both pyrimidines and purines. These differences are counterintuitive, as only two of 29 active site residues differ in the two enzymes: F80 and M118 in *DmdNK* are L78 and L116 in TK2. In addition to reporting an optimized protocol for the expression and purification of TK2, we have used site-directed mutagenesis to introduce the *DmdNK*-like amino acids into TK2, and characterized the three resulting enzymes (L78F-TK2, L116M-TK2, and L78F/L116M-TK2). These mutations improve the K_M for thymidine, increasing the catalytic activity of L78F/L116M-TK2 4.4-fold, yet leaving the activity for deoxycytidine or the purine nucleosides unchanged. © 2007 Elsevier Inc. All rights reserved.

Keywords: Human thymidine kinase 2; Site-directed mutagenesis; Deoxyribonucleoside kinase; Nucleoside analog

The salvage of 2'-deoxyribonucleosides from intracellular or extracellular degradation of DNA is essential for providing deoxyribonucleoside triphosphates for nuclear DNA repair and mitochondrial DNA synthesis. Deoxyribonucleoside kinases (dNKs) are key enzymes of the salvage pathway, catalyzing the initial phosphoryl transfer from a suitable donor (typically ATP) to a 2'-deoxyribonucleoside, forming a 2'-deoxyribonucleoside monophosphate [1]. In the treatment of several cancers and viral diseases, this reaction is the rate-limiting step in activating nucleoside analog prodrugs to their cytotoxic, phosphorylated forms [2]. This central role of dNKs has led to interest in investigating the determinants of their substrate specificity, to facilitate engineering novel dNKs that can phosphorylate nucleoside analogs more effectively.

Human mitochondrial thymidine kinase 2 (TK2) is a pyrimidine-specific dNK, phosphorylating thymidine

(dThd) and deoxycytidine (dCyd) [3,4]. It is constitutively expressed throughout the cell cycle [2], making it critical for the phosphorylation of pyrimidine analog prodrugs, such as AZT (3'-azido-2',3'-dideoxythymidine), FIAU (1-(2'-deoxy-2'-fluoro-β-D-arabinofuranosyl)-5-iodouracil), and AraT (1-β-D-arabinofuranosylthymine), that are routinely used in antiviral and cancer chemotherapy [3,5]. Mutations in the TK2 gene have also been linked to mitochondrial DNA depletion syndrome [6]. Despite the clinical importance of TK2, there is little structural or functional information on this enzyme; unlike other members of the dNK family, TK2 has not been engineered for improved nucleoside analog activation or altered substrate specificity. This may reflect difficulties in expressing TK2 heterologously, as the protein tends to aggregate under standard overexpression conditions [7]. These problems have also hampered efforts to determine the structure of the enzyme.

In this study, we have developed an improved protocol for heterologous expression and purification of recombinant TK2, and used site-directed mutagenesis to explore

* Corresponding author. Fax: +1 404 727 6586.

E-mail address: sal2@emory.edu (S. Lutz).

	10	20	30	40	50	60	70	80	
hTK2	M--GAFQCRPSSDKEQEKEKKSVCVEGNIASGKTTCLFEFFSNA--TDVEVLTEPVSKWRNVGRHNPGLMYHDASRWGLTLOTYVQLT								
DmdNK	MAEAASCARKGT--KYAEGTQPFTVFLIEGNIGSGKTTYLNHFHEKYKNDICLLTEPVEKWRNVNGVNLLELMYKDKPKWAMPFQSYVTLT								
AgdNK	MPPIAS----EK--LGASGKPPFTVFVEGNIGSGKTTFLNHFQKF--NDICLLTEPVEKWRNCGGVNLLDLMYKESHWRWAMPFQTYVTLT								
BmdNK	-----MSANNVKPFTVFVEGNIGSGKTTFLFHFQFE--DITLLTEPVEMWRDLKGCNLLLELMYKDEPKWAMTFQSYVSLT								
	90	100	110	120	130	140	150	160	170
hTK2	MLDRHTRPQVSSVRLMERSIHSARYIFVENLYRSRGMPEVDYVVLSEWFDWILRNMDVSVDLIVYLRTNPETCYQLKKRCREEEKVI								
DmdNK	MLQSHAPTNNKLLKIMERSIFSARYCFVENMRRNGSLQGMYNLTLEEWYKFIIEESIHVQADLI IYLRTSPEVAYERIRQRARSEESCVC								
AgdNK	MLDMHTCQTDKSVKLMERSLFSARNCFVESMLASGSLHQGMYNVLQEWYDFICCNIIHQADLIVYLQTSPEVVYERMQRARSEESCVC								
BmdNK	MLDMHRRPAPTVPKLMERSLFSARYCFVEHIMRNNTLHPAQFAVLDEWFRFIQHNIPIDADLIVYLKTSPPSIVYQRIKKRRARSEEQVC								
	180	190	200	210	220	230			
hTK2	PLEYLEAIHHLHEEWLIKGSIFPM--AAPVLVIEADHMMERMLELFEQNRDRILTPEN--RKHCP-----								
DmdNK	PLKYLQELHELHEDWLIHQRR--PQ--SCKVLVLDADLNLENIGTEYQRSESSIFDAIS--SNQPPSPVLVSPSKKQR-----VA--R								
AgdNK	PLEYLKELHELHENWLIHGAS--PR--PAPVLVLDADLDLNTIGAEYERSETSIKLPILIENTNQHALITSPAKRAK-----TD--F								
BmdNK	PLSYIEELHRLHEDWLINRIH--AECAPVLVLDADLDLSQITDEYKRSEHQILRKAV--NVVMSSPNKHSPKPISTTPIKITPHMRIL								

Fig. 1. Structure-based multiple sequence alignment of TK2, *DmdNK*, *AgdNK*, and *BmdNK* computed by 3DCoffee [24]. Numbering is according to TK2. Active-site residues, as determined from the structure of *DmdNK* [13], are highlighted in grey. Residues targeted for mutagenesis (L78, L116) are underlined and bold.

the roles of non-conserved active site residues on the structure/function relationship of the enzyme. Aligning the sequence of TK2 with its closest phylogenetic relatives, the dNKs from *Drosophila melanogaster* (*DmdNK*), *Anopheles gambiae* (*AgdNK*) and *Bombyx mori* (*BmdNK*), reveals remarkable conservation of the active site residues, with the four enzymes only differing in two positions (Fig. 1). Unlike TK2, however, the insect enzymes are each able to phosphorylate all four native 2'-deoxyribonucleosides [8–10]. We hypothesized that the different substrate specificities of these enzymes may be due to the differences in the non-conserved amino acid positions. We attempted to engineer TK2 into an enzyme with *DmdNK*-like broad substrate specificity and high catalytic efficiency, by using site-directed mutagenesis to replace L78 and L116 of TK2 with the corresponding residues in *DmdNK* (F80 and M118). We have characterized the intermediate enzymes L78F-TK2 and L116M-TK2, as well as the double mutant, L78F/L116M-TK2. Our results show that the *DmdNK*-like mutations do not broaden the specificity of TK2; however, they do improve activity towards dThd, and affect the quaternary structure of the enzyme.

Materials and methods

Materials. Enzymes were from New England Biolabs (Beverly, MA) unless otherwise indicated. *Pfu* Turbo DNA polymerase (Stratagene, La Jolla, CA) was used for all cloning. Pyruvate kinase and lactate dehydrogenase were from Roche Biochemicals (Indianapolis, IN). Gel filtration standards were from Bio-Rad (Hercules, CA). All other reagents were from Sigma-Aldrich (St. Louis, MO) unless noted.

Molecular modeling. The entire TK2 sequence was submitted for First Approach Mode to the SwissModel server [11]. Residues 9–207 of TK2 were able to be modeled from the structures of *DmdNK* complexed with the substrates dThd (PDB code 1OT3 [12]) and dCyd (PDB code 1J90 [13]).

Cloning and mutagenesis. The genes encoding TK2 and *DmdNK* were isolated previously [14]. The first 19 codons of TK2 were removed by PCR, with a new initiator methionine codon introduced in place of K19 (numbering according to [4]) using primers TK2-F and TK2-R (Table 1). This truncated form of TK2 has identical kinetic properties to the full-length enzyme [3]. Primers DM-F and DM-R were used for PCR ampli-

fication of the full-length *DmdNK* gene. The TK2 gene was corrected for codon usage in *Escherichia coli* using overlap extension PCR with primers CO1-CO10 (Table 1); the rare arginine codons AGA and AGG were mutated to CGC at amino acid positions 89, 103, 109, 118, 139, 161, 165, and 167. This codon-optimized variant was used as the “wild-type”. The L78F and L116M mutations were introduced by overlap extension PCR, using primers L78F-F and L78F-R, and L116M-F and L116M-R, respectively (Table 1). All genes were subcloned via *NdeI* and *SpeI* restriction sites (introduced by the primers) into pET-14b (Novagen, Madison, WI) for overexpression. Constructs were confirmed by DNA sequencing.

Expression and purification. *Escherichia coli* BL21(DE3) pLysS cells (Novagen) were transformed with plasmids containing the wild-type and mutated genes. Cells were grown at 37 °C in LB broth containing carbenicillin (100 µg/ml), chloramphenicol (34 µg/ml) and ethanol (3%, v/v), to $A_{600} \sim 0.6$. Cultures were shifted to 22 °C, isopropyl- β -D-thiogalactopyranoside was added (0.4 mM, final concentration) and cells were grown for 12 h. *DmdNK* was overexpressed as described previously [14]. Cells were harvested by centrifugation and the pellets stored at –20 °C. For protein purification, cell pellets were resuspended in buffer A (50 mM potassium phosphate, 300 mM NaCl, 1 mM β -mercaptoethanol, 10% glycerol, pH 8.0) with protease inhibitor cocktail, rLysozyme (Novagen) and Benzonase nuclease (Novagen) added according to manufacturers' protocols. Bacteria were lysed by sonication on ice, and the lysates clarified by centrifugation (4500g, 4 °C, 30 min) before being loaded onto Talon metal affinity resin (Clontech, Mountain View, CA). After two wash steps with five column volumes of buffer A containing 5 mM and 15 mM imidazole, the target proteins were eluted in buffer A containing 150 mM imidazole. Amicon-Ultra centrifugal filter units (Amicon Bioseparations, Billerica, MA) were used to exchange the purified proteins into either buffer B (50 mM Tris-HCl, 150 mM NaCl, pH 8.0) for biophysical characterization, or storage buffer (50 mM Tris-HCl, 150 mM NaCl, 1 mM β -mercaptoethanol, 10% glycerol, pH 8.0). Protein concentrations were quantified by measuring A_{280} (TK2, $\epsilon = 38,690 \text{ M}^{-1} \text{ cm}^{-1}$; *DmdNK* $\epsilon = 39,880 \text{ M}^{-1} \text{ cm}^{-1}$, calculated according to Gill and von Hippel [15]). Aliquots were stored at –80 °C.

Enzyme assays. Nucleoside phosphorylation activity was determined using a spectrophotometric coupled-enzyme assay [10,16]. Activity was measured at 37 °C, in a 500 µl reaction mixture containing 50 mM Tris-HCl (pH 7.5), 100 mM KCl, 2.5 mM MgCl_2 , 0.18 mM NADH, 0.21 mM phosphoenolpyruvate, 1 mM ATP, 1 mM 1,4-dithio-DL-threitol, 30 U/ml pyruvate kinase, and 33 U/ml lactate dehydrogenase. Substrate concentrations ranged from 1 to 3000 µM. Measurements were made in triplicate and corrected for background. Apparent k_{cat} (calculated assuming one active site per enzyme monomer) and K_M values were determined by fitting data to the Michaelis-Menten equation, using non-linear regression analysis in Origin (OriginLab, Northampton, MA).

Table 1
Primers used for plasmid construction and sequencing

Name	Sequence
TK2-F	5'-CAGCCATATGTCAGTGATCTGTGTGTCG-3'
TK2-R	5'-GCGACTAGTTCATGGGCAATGCTT-3'
DM-F	5'-CTATCATATGGCGGAGGCAGCATCTGTG-3'
DM-R	5'-CTCACTAGTTCATCTGGCGACCCTCTG-3'
CO1	5'-CATGCTGGACCGCCATACTCGTCCTCAGGTGTC-3'
CO2	5'-GACGAGTATGGCGTCCAGCATGGTGAGCTGCA-3'
CO3	5'-ATGGAGCGCTCGATTACAGCGCACCGCTACATTTTTGTAGAAAAC-3'
CO4	5'-GCGCTGTGAATCGAGCGCTCCATCAACCGTACAGATG-3'
CO5	5'-AAACCTGTATCGCAGTGGGAAGATGCCAGAAGT-3'
CO6	5'-TCTTCCCCTGCGATACAGTTTTCTACAAAAA-3'
CO7	5'-CTGGATCTTGCGCAACATGGACGTGTCTGTTGA-3'
CO8	5'-CGTCCATGTTGCGCAAGATCCAGTCAAACCATT-3'
CO9	5'-CCAGCGCTTAAAGAAGCGCTGCCGGAAGAGGAGAAGGTCATTCC-3'
CO10	5'-GCAGCGCTTCTTTAAGCGCTGGTAACAAGTCTCAGGAT-3'
L116M-F	5'-GTAGAAAACATGTATCGCAGTGGGAAGAT-3'
L116M-R	5'-CACTGCGATACATGTTTTCTACAA-3'
L78F-F	5'-TCTTACGTTCCAGACTTATGTGCAGCTC-3'
L78F-R	5'-CTGCACATAAGTCTGGAACGTAAGAC-3'

Restriction sites are italicized. Mutations are underlined.

Biophysical characterization. Far-UV circular dichroism (CD) spectra were obtained using a J-810 spectropolarimeter (Jasco, Inc., Easton, MD). Spectra were recorded at 22 °C from 260 to 190 nm (0.5 nm increments) using a 0.1 mm pathlength cell, 20 nm/min scan rate, 4 s response time, and 2 nm bandwidth. Proteins were analyzed at concentrations of 0.9–1.2 mg/ml. Recorded data represent the means of eight scans. Spectra were corrected for buffer absorbance and converted to mean residue ellipticity ($[\theta]_{\text{mrw}}$).

Intrinsic tryptophan fluorescence emission spectra were measured using a FluoroMax-3 spectrophotometer (Horiba Jobin Yvon, Edison, NJ). Protein samples were excited at 295 nm in a 1 cm pathlength cuvette. Emission data were accumulated at 20 °C, from 305 to 405 nm, in 1 nm increments, using 5 nm excitation and emission bandwidths. Protein concentrations were 3.2–3.6 μM . Each spectrum represents the mean of three scans, normalized for protein concentration and corrected for contribution of buffer to the signal.

Oligomeric states of the purified enzymes were determined by size-exclusion chromatography using a Superdex 200 10/300 GL column and AKTApurifier (both Amersham Biosciences). The column was equilibrated in buffer B and calibrated with the standard proteins thyroglobulin, bovine γ -globulin, chicken ovalbumin and equine myoglobin. Blue dextran and acetone were used to determine the void volume ($V_0 = 8.00$ ml) and total exclusion volume ($V_1 = 20.75$ ml) of the column, respectively. The partition coefficients of the standards, $K_{\text{av}} = (V_e - V_0)/(V_1 - V_0)$, where V_e is the elution volume, were linearly related to log molecular mass. Approximately 100 μl of each purified protein (1 mg/ml) was loaded onto the column and eluted at a flow rate of 0.5 ml/min.

Results and discussion

Rational redesign of the TK2 active site

Our sequence alignment (Fig. 1) suggested that L78 and L116 were the only two active site residues that are not identical in TK2 and the multi-substrate insect kinases. This was confirmed when we constructed a homology model of TK2 using the structure of *DmdNK* [13] as a template (Fig. 2A). Overall, we identified 29 amino acid residues within the active site; 27 of these are identical in TK2

and *DmdNK*. Consistent with previous models [7,13], L78 and L116 in TK2 are in structurally equivalent positions to F80 and M118 of *DmdNK*. It has been proposed that the broad specificity of *DmdNK* is due to its large, hydrophobic active site [13]. Interestingly, our model predicts that the active site of TK2 (with the two leucines replacing a phenylalanine and a methionine) is even larger and more hydrophobic. Based on these findings, we used site-directed mutagenesis to produce the L78F and L116M single-site mutants, as well as the L78F/L116M double mutant.

Expression and purification

Expression and purification of TK2 in sufficient quantity and quality is critical for enzyme engineering, yet has been problematic. Previously reported protocols have suffered from low yields [3] or required complex, multi-step expression and purification [7]. We therefore sought a robust, simple and high yield protocol to prepare TK2 for *in vitro* characterization.

First, we optimized codon usage in the human gene by replacing eight AGG and AGA arginine codons with the preferred *E. coli* codon CGC. This codon-optimized variant of TK2 served as the wild-type control and as the template for site-directed mutagenesis. We next attempted to follow the protocol of Wang et al. [3], as the authors have shown that an N-terminal (His)₆-tag does not affect activity and therefore would not need to be proteolytically cleaved (in turn, obviating the need for further purification). However, this approach yielded negligible amounts of soluble enzyme. Instead, we found that using the pET system with a reduced induction temperature and ethanol added to the medium (to induce a heat shock-like response in the bacterial host [7,17–19]) increased the yield of soluble, purified TK2 to an average of 11 mg/L culture. This yield

is similar to that reported by Barroso et al. (10 mg/L; [7]), but was obtained with a simpler protocol, requiring fewer purification steps. Our improved protocol was used to purify L78F-TK2, L116M-TK2, and L78F/L116M-TK2 to >95% purity.

Kinetics

To examine the effects of the mutations on the activity and substrate specificity of TK2, each enzyme was evaluated for its ability to phosphorylate the four natural 2'-deoxyribonucleosides (Table 2).

The catalytic efficiency (k_{cat}/K_M) of L78F-TK2 was improved 2.5-fold for dThd, and 1.8-fold for dCyd (Table 2). The L78F mutation introduces the favorable base stacking interaction observed in *DmdNK* [13] and consequently improves the Michaelis constants for dThd and dCyd. We expected that this mutation would also have equal effects on the turnover numbers for each pyrimidine, since the new aromatic side chain is angled towards the edge of the substrate that is identical in dThd and dCyd (i.e., C2 and N3). However, the mutation improves the k_{cat} for dThd and worsens it slightly for dCyd, suggesting that this single-site mutation causes

structural rearrangements that are not strictly paralleled in *DmdNK*.

Replacement of L116 with methionine had little effect on the kinetics of dThd phosphorylation. In contrast, L116M-TK2 retained only 23% of dCyd phosphorylation activity relative to wild-type TK2; the k_{cat} and K_M each decreased approximately 2-fold. Examination of the *DmdNK* structures with dThd and dCyd reveals that upon binding dCyd, the side chain of Q81 rotates, forming hydrogen bonds with the nucleobase, while the side chain of M118 must move 2.8 Å to avoid a steric clash with the amino group of Q81 (Q79 in TK2; Figs. 2B and C). In *DmdNK*, this conformational change is accommodated by a small hydrophilic residue (S123) located behind the methionine. TK2 has a lysine residue at the equivalent position (K121), which may interfere with the analogous conformational rearrangement in L116M-TK2, due to its size and charge.

The double mutant, L78F/L116M-TK2, retains wild-type activity towards dCyd, suggesting that the favorable stacking interaction introduced by the L78F mutation and the steric constraints of the L116M mutation offset one another. For dThd, the combination of the two mutations is synergistic, raising the catalytic performance almost 5-fold over wild-type TK2. L78F/L116M-TK2

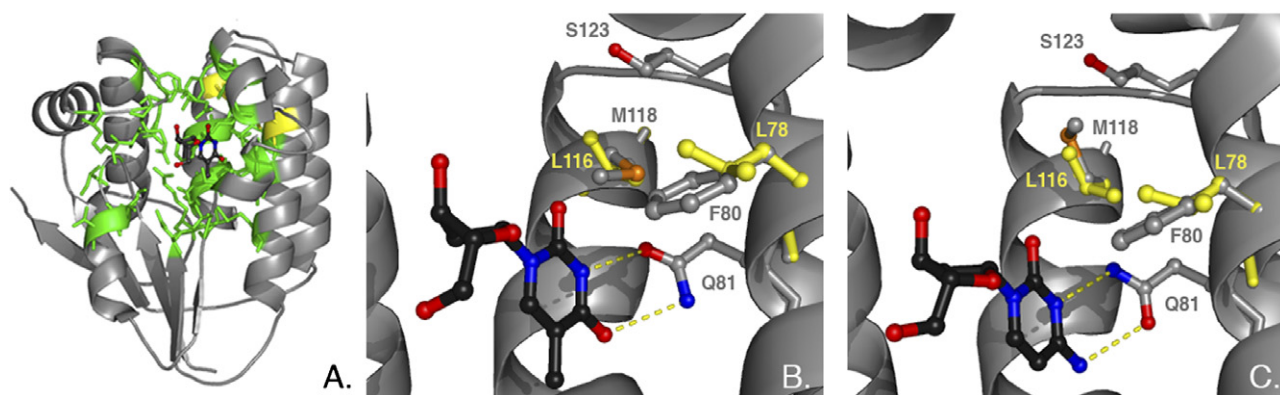


Fig. 2. Comparison of TK2 and *DmdNK* conserved residues and structure. (A) Homology model of TK2 monomeric subunit. Active-site residues are shown as sticks and colored green. L78F and L116M are yellow. (B) Overlay of TK2 homology model and *DmdNK* with dThd (PDB code 1OT3 [12]) or (C) dCyd (PDB code 1J90 [13]) in the binding pocket. *DmdNK* and TK2 residues are colored grey and yellow, respectively.

Table 2

Kinetic parameters of wild-type and mutant deoxyribonucleoside kinases

	dThd			dCyd			dAdo	dGuo
	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($\text{M}^{-1} \text{s}^{-1}$)	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($\text{M}^{-1} \text{s}^{-1}$)	k_{cat}/K_M ($\text{M}^{-1} \text{s}^{-1}$)	k_{cat}/K_M ($\text{M}^{-1} \text{s}^{-1}$)
<i>DmdNK</i> ^a	17.5 ± 0.6	1.9 ± 0.3	9.1 × 10 ⁶	19.3 ± 0.5	2.2 ± 0.3	9.0 × 10 ⁶	8.3 × 10 ⁴	1.8 × 10 ⁴
TK2	1.8 ± 0.2	6.7 ± 1.1	2.7 × 10 ⁵	2.0 ± 0.2	16.2 ± 1.1	1.2 × 10 ⁵	nd	nd
L78F-TK2	2.5 ± 0.2	3.7 ± 0.6	6.7 × 10 ⁵	1.7 ± 0.1	7.8 ± 0.6	2.2 × 10 ⁵	nd	nd
L116M-TK2	1.1 ± 0.1	4.9 ± 1.2	2.3 × 10 ⁵	1.0 ± 0.1	36.1 ± 4.9	2.7 × 10 ⁴	nd	nd
L78F/L116M-TK2	1.3 ± 0.1	1.1 ± 0.3	1.2 × 10 ⁶	1.7 ± 0.1	14.9 ± 1.5	1.1 × 10 ⁵	nd	nd

Values represent the mean of three experiments. nd, not detectable.

^a The kinetic parameters of *DmdNK* have been determined previously using a radiometric assay [10]; our results using the coupled-enzyme spectrophotometric assay are in agreement.

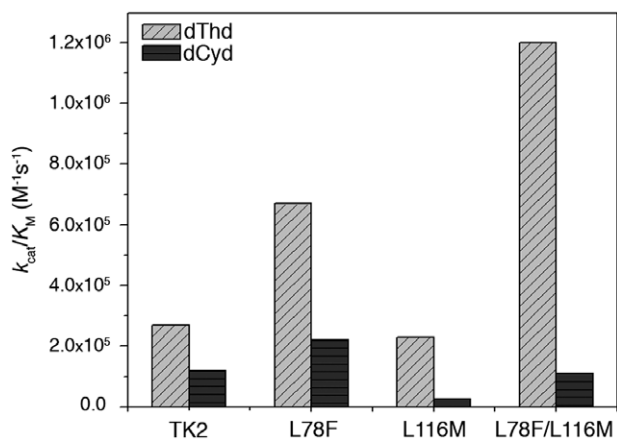


Fig. 3. Summary of catalytic efficiencies.

has a K_M that is improved >6-fold, and now matches the binding affinity of *DmdNK* for dThd (Table 2). These two substitutions, which flank Q79, are apparently crucial in optimizing the active site geometry for dThd binding.

Interestingly, the activity of mutant enzymes towards purine nucleosides remained below the detection limit of the assay ($k_{cat}/K_M < 10^2 M^{-1} s^{-1}$). The L78F substitution generates a TK2 active site that approximates that of *BmdNK* (with the exception that L116 in TK2 is an isoleucine in *BmdNK*), yet it fails to introduce purine activity at levels comparable to the insect kinase [9]. The double mutant L78F/L116M-TK2 has all 29 active site residues identical to *DmdNK*, completely recreating the active site of *DmdNK* on an amino acid level, yet no improvement in purine activity was detected in the engineered enzyme. The combined effect of the L78F and L116M mutations was to narrow the substrate specificity of TK2, making it almost 5-fold more specific for dThd. This change in specificity results from an improved catalytic efficiency for dThd, rather than a decrease in activity towards dCyd (Fig. 3).

Biophysical characterization

The wild-type and mutant TK2 enzymes were characterized by CD spectroscopy, fluorescence spectroscopy, and

size-exclusion chromatography (Fig. 4) to probe the effects of the mutations on the secondary, tertiary, and quaternary structure of the enzymes.

The far-UV CD spectra of all four proteins were essentially identical (Fig. 4A), indicating that the changes in specificity and activity were not brought about by changes in secondary structure. We also probed for changes in tertiary structure, by measuring intrinsic tryptophan fluorescence (Fig. 4B) to detect perturbations in the environment of the five tryptophan side chains in TK2 (W55, W74, W133, W136, and W188). Three of these (W74, W133, and W136) are located in the dimerization interface, W55 stacks with the base of the substrate in the enzyme active site and W188 is located at the end of α -helix 8. The emission spectra of TK2 and L116M-TK2 are identical. In contrast, both enzymes with the L78F mutation exhibit a decrease in fluorescence intensity of 15%, and a 1 nm blue shift in emission maxima. The small changes observed for L78F-TK2 and L78F/L116M-TK2 are most likely due to introduction of an aromatic side chain (L78F) in close proximity to the active site tryptophan (W55).

Size-exclusion chromatography was used to examine possible effects of the mutations on quaternary structure. Untagged wild-type TK2 is known to exist predominantly as a dimer (~70%) in solution, though with tetrameric, hexameric and higher-molecular weight forms also present [7,20]. Similarly, (His)₆-tagged TK2 eluted in two peaks (Fig. 4C), corresponding to estimated molecular masses of 91.1 kDa (24% total protein) and 44.7 kDa (76% total protein). These match the expected masses of compact tetrameric and dimeric forms, respectively (monomer $M_r = 27,600$). In contrast, all three mutant enzymes showed a shift in the dimer-tetramer equilibrium; in each case the fraction of enzyme in the dimeric form increased to 90–97% of total protein with only a small shoulder in the elution profile indicating the presence of the homotetramer (Fig. 4C). Although the dimer-tetramer ratio has no effect on catalytic activity [20], the observed equilibrium shift suggests that the active site mutations impact enzyme topology through long-range conformational changes.

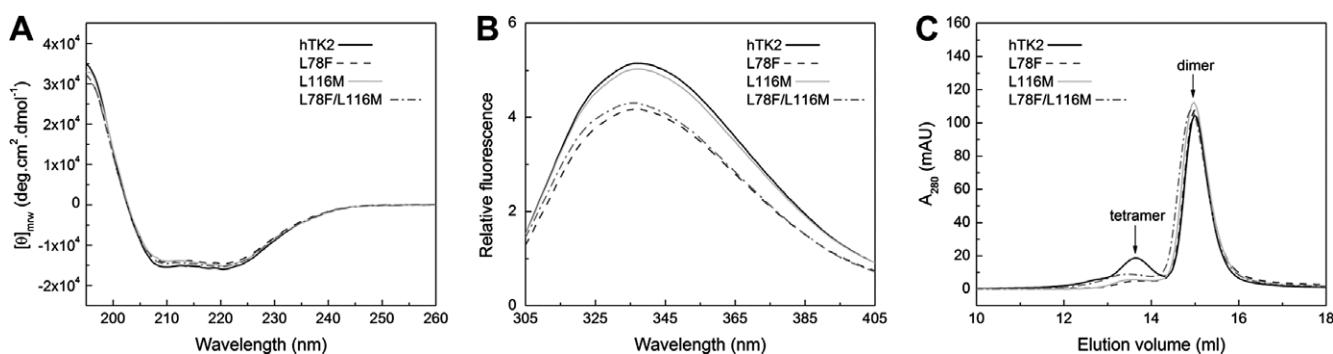


Fig. 4. Biophysical characterization of wild-type and mutant TK2 enzymes. (A) Far-UV CD spectra. (B) Intrinsic tryptophan fluorescence spectra. (C) Size-exclusion chromatography profiles.

Summary

We have generated a mutant, L78F/L116M-TK2, that is 4.4-fold more efficient than wild-type TK2 at phosphorylating dThd, making it the first reported example of an improved TK2 variant. We anticipate that this more active variant will be an ideal starting point for engineering kinases with improved activity towards nucleoside analogs. Our improved protocol for protein expression and purification will also facilitate future engineering, in addition to enabling structural studies.

Surprisingly, L78F/L116M-TK2 showed no change in catalytic efficiency with dCyd, nor any detectable activity towards purine nucleosides. Rather than broadening specificity as hypothesized, mutagenesis at positions 78 and 116 of TK2 actually narrowed it. It has generally been assumed that active site residues determine the substrate specificity in dNKs [13,21–23]. Here, we have provided direct evidence that active site residues are not the sole determinant of the narrow specificity of TK2, but rather that more distant residues must play equally important roles. This demonstrates the limited ability of rational design to predict subtle changes, including those that are mediated by long-range interactions within the structure of the protein. Other probes of protein structure/function relationships, such as homologous or non-homologous recombination, will be required to search beyond the active sites of dNKs for determinants of their activity and specificity.

Acknowledgments

We acknowledge financial support in part by the National Institutes of Health (GM69958), as well as by a grant to the Emory Center for AIDS Research (AI050409) from the National Institutes of Health and by institutional funding from the Emory University Health Science Center. M.L.G. is an ARCS Fellow. DNA sequencing was performed at the Center for Fundamental and Applied Molecular Evolution at Emory University (NSF-MRI 0320786).

References

- [1] E.S. Arnér, S. Eriksson, Mammalian deoxyribonucleoside kinases, *Pharmacol. Ther.* 67 (1995) 155–186.
- [2] A.R. Van Rompay, M. Johansson, A. Karlsson, Substrate specificity and phosphorylation of antiviral and anticancer nucleoside analogues by human deoxyribonucleoside kinases and ribonucleoside kinases, *Pharmacol. Ther.* 100 (2003) 119–139.
- [3] L. Wang, B. Munch-Petersen, A. Herrström Sjöberg, U. Hellman, T. Bergman, H. Jörnvall, S. Eriksson, Human thymidine kinase 2: molecular cloning and characterisation of the enzyme activity with antiviral and cytostatic nucleoside substrates, *FEBS Lett.* 443 (1999) 170–174.
- [4] M. Johansson, A. Karlsson, Cloning of the cDNA and chromosome localization of the gene for human thymidine kinase 2, *J. Biol. Chem.* 272 (1997) 8454–8458.
- [5] E.S. Arnér, T. Spasokoukotskaja, S. Eriksson, Selective assays for thymidine kinase 1 and 2 and deoxycytidine kinase and their activities in extracts from human cells and tissues, *Biochem. Biophys. Res. Commun.* 188 (1992) 712–718.
- [6] A. Saada, A. Shaag, H. Mandel, Y. Nevo, S. Eriksson, O. Elpeleg, Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy, *Nat. Genet.* 29 (2001) 342–344.
- [7] J.F. Barroso, M. Elholm, T. Flatmark, Tight binding of deoxyribonucleotide triphosphates to human thymidine kinase 2 expressed in *Escherichia coli*. Purification and partial characterization of its dimeric and tetrameric forms, *Biochemistry* 42 (2003) 15158–15169.
- [8] W. Knecht, G.E. Petersen, M.P. Sandrini, L. Søndergaard, B. Munch-Petersen, J. Piskur, Mosquito has a single multisubstrate deoxyribonucleoside kinase characterized by unique substrate specificity, *Nucleic Acids Res.* 31 (2003) 1665–1672.
- [9] W. Knecht, G.E. Petersen, B. Munch-Petersen, J. Piskur, Deoxyribonucleoside kinases belonging to the thymidine kinase 2 (TK2)-like group vary significantly in substrate specificity, kinetics and feed-back regulation, *J. Mol. Biol.* 315 (2002) 529–540.
- [10] B. Munch-Petersen, W. Knecht, C. Lenz, L. Søndergaard, J. Piskur, Functional expression of a multisubstrate deoxyribonucleoside kinase from *Drosophila melanogaster* and its C-terminal deletion mutants, *J. Biol. Chem.* 275 (2000) 6673–6679.
- [11] T. Schwede, J. Kopp, N. Guex, M.C. Peitsch, SWISS-MODEL: An automated protein homology-modeling server, *Nucleic Acids Res.* 31 (2003) 3381–3385.
- [12] N.E. Mikkelsen, K. Johansson, A. Karlsson, W. Knecht, G. Andersen, J. Piskur, B. Munch-Petersen, H. Eklund, Structural basis for feedback inhibition of the deoxyribonucleoside salvage pathway: studies of the *Drosophila* deoxyribonucleoside kinase, *Biochemistry* 42 (2003) 5706–5712.
- [13] K. Johansson, S. Ramaswamy, C. Ljungcrantz, W. Knecht, J. Piskur, B. Munch-Petersen, S. Eriksson, H. Eklund, Structural basis for substrate specificities of cellular deoxyribonucleoside kinases, *Nat. Struct. Biol.* 8 (2001) 616–620.
- [14] M.L. Gerth, W.M. Patrick, S. Lutz, A second-generation system for unbiased reading frame selection, *Protein Eng. Des. Sel.* 17 (2004) 595–602.
- [15] S.C. Gill, P.H. von Hippel, Calculation of protein extinction coefficients from amino acid sequence data, *Anal. Biochem.* 182 (1989) 319–326.
- [16] P. Valentin-Hansen, Uridine-cytidine kinase from *Escherichia coli*, *Methods Enzymol.* 51 (1978) 308–314.
- [17] J.G. Thomas, F. Baneyx, Protein misfolding and inclusion body formation in recombinant *Escherichia coli* cells overexpressing heat-shock proteins, *J. Biol. Chem.* 271 (1996) 11141–11147.
- [18] J. Steczko, G.A. Donoho, J.E. Dixon, T. Sugimoto, B. Axelrod, Effect of ethanol and low-temperature culture on expression of soybean lipoxygenase L-1 in *Escherichia coli*, *Protein Express Purif.* 2 (1991) 221–227.
- [19] J.G. Thomas, F. Baneyx, Divergent effects of chaperone overexpression and ethanol supplementation on inclusion body formation in recombinant *Escherichia coli*, *Protein Express Purif.* 11 (1997) 289–296.
- [20] J.F. Barroso, R.N. Carvalho, T. Flatmark, Kinetic analysis and ligand-induced conformational changes in dimeric and tetrameric forms of human thymidine kinase 2, *Biochemistry* 44 (2005) 4886–4896.
- [21] W. Knecht, M.P. Sandrini, K. Johansson, H. Eklund, B. Munch-Petersen, J. Piskur, A few amino acid substitutions can convert deoxyribonucleoside kinase specificity from pyrimidines to purines, *EMBO J.* 21 (2002) 1873–1880.
- [22] M.P. Sandrini, J. Piskur, Deoxyribonucleoside kinases: two enzyme families catalyze the same reaction, *Trends Biochem. Sci.* 30 (2005) 225–228.
- [23] E. Sabini, S. Ort, C. Monnerjahn, M. Konrad, A. Lavie, Structure of human dCK suggests strategies to improve anticancer and antiviral therapy, *Nat. Struct. Biol.* 10 (2003) 513–519.
- [24] O. O'Sullivan, K. Suhre, C. Abergel, D.G. Higgins, C. Notredame, 3DCoffee: combining protein sequences and structures within multiple sequence alignments, *J. Mol. Biol.* 340 (2004) 385–395.