nature
structural biology

Structure of human dCK suggests strategies to improve anticancer and antiviral therapy

Elisabetta Sabini[1], Stephan Ort[2], Christian Monnerjahn[2], Manfred Konrad[2] & Arnon Lavie[1]

Human deoxycytidine kinase (dCK) phosphorylates the natural deoxyribonucleosides deoxycytidine (dC), deoxyguanosine (dG) and deoxyadenosine (dA) and is an essential enzyme for the phosphorylation of numerous nucleoside analog prodrugs routinely used in cancer and antiviral chemotherapy. For many of these compounds, the phosphorylation step catalyzed by dCK is the rate-limiting step in their overall activation pathway. To determine the factors that limit the phosphorylation efficiency of the prodrug, we solved the crystal structure of dCK to a resolution of 1.6 Å in complex with its physiological substrate deoxycytidine and with the prodrugs AraC and gemcitabine. The structures reveal the determinants of dCK substrate specificity. Especially relevant to new prodrug development is the interaction between Arg128 and the hydrogen-bond acceptor at the sugar 2′-arabinosyl position of AraC and gemcitabine. On the basis of the structures, we designed a catalytically superior dCK variant that could be used in suicide gene-therapy applications.

Human deoxycytidine kinase activates the clinically established prodrugs AraC, fludarabine and cladribine and some recently developed chemotherapy agents such as gemcitabine and troxacitabine. In addition, dCK also phosphorylates the antiviral compounds 3TC and ddC used for the treatment of HIV infection. The critical function of dCK for the efficacy of prodrug chemotherapy is evident from a direct correlation between the activity of the enzyme and drug sensitivity in tumor cell lines[1]. Cells lacking dCK activity are resistant to a variety of drugs, including AraC, cladribine, fludarabine and gemcitabine[2,3], and drug sensitivity to AraC can be restored by expressing functional dCK protein[4]. Moreover, *in vivo* studies conducted on animal tumors using gemcitabine showed that increased dCK activity, mediated by dCK gene transfer, results in enhanced accumulation and prolonged elimination kinetics of gemcitabine triphosphate and, ultimately, enhanced tumor response to the prodrug[5].

Despite the determining role played by human dCK in current anticancer prodrug treatment, no structural information on the enzyme was available until now. We have determined the three-dimensional structure of dCK to investigate the substrate specificity of the enzyme and to ultimately overcome the limitations imposed by the low catalytic activity of the enzyme: dCK is a relatively poor catalyst (with turnover rates well below 1 s^{-1}). Our long-term goal is the structure-based design of mutant enzymes with enhanced activity to be used in gene therapy applications. Additionally, the protein structure may be used to guide the development of improved antiviral and anticancer molecules. We report the fully refined structures of human dCK in complex with dC–ADP, AraC–ADP–Mg and gemcitabine–ADP–Mg. These crystal structures, with a natural nucleoside or nucleoside analogs and ADP–Mg, represent 'dead-end' complexes composed of a mixture of substrates and products. Nonetheless, they allow us to observe a closed conformation of the enzyme similar to the one expected in the presence of ATP, as previously demonstrated in the case of other kinases (such as thymidylate kinase, TMPK) when comparing the ATP analog (β,γ-imidoadenosine 5′-triphosphate) to the ADP-bound structures[6]. Furthermore, analysis of distances in our dCK structures between the nucleosides and a modeled ATP molecule suggests that the position observed for the nucleosides in our ADP complex is similar to the one expected with ATP (see below).

These structures explain the different phosphorylation efficiencies observed with various cytidine analogs and suggest specific modifications to our current arsenal of prodrugs that we predict will result in their improved phosphorylation by dCK and, thus, in an improved therapeutic index. Furthermore, the newly gained structural knowledge was exploited in the design of a dCK mutant that has a faster phosphorylation rate than the wild-type enzyme: ~50-fold more efficient in phosphorylating deoxycytidine (dC) and ~4-fold more efficient in activating gemcitabine. This engineered enzyme could have an important role in gene-directed enzyme prodrug therapy strategies. The presence of dCK variants with enhanced activity in neoplastic or virus-infected cells will confront the growing challenge of drug resistance[7].

RESULTS

Overall structure

We solved the three-dimensional structure of human dCK in complex with ADP at the phosphoryl donor site and dC, AraC or gemcitabine at the acceptor site, including the catalytically essential magnesium ion, to a resolution of 1.6 Å. Structural information for human dCK

[1]University of Illinois at Chicago, Department of Biochemistry and Molecular Biology, 1819 West Polk Street, Chicago, Illinois 60612, USA. [2]Max Planck Institute for Biophysical Chemistry, Department of Molecular Genetics, Am Fassberg 11, D-37077 Göttingen, Germany. Correspondence should be addressed to A.L. (lavie@uic.edu).

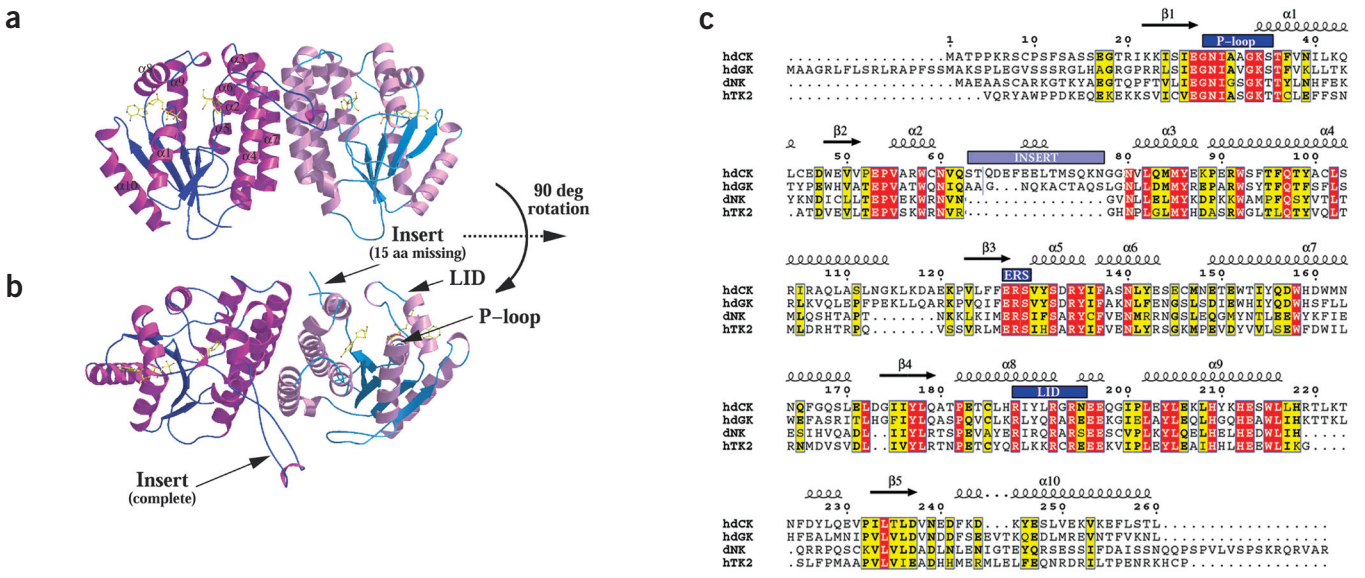

Figure 1 Overall fold of dCK. (**a**) Ribbon diagram of the dCK dimer in the presence of dC and ADP. Human dCK contains a 15-residue insert (Ser63–Asn77) between helix α2 and helix α3. A similar insert (12 residues long) is also found in human dGK (residues Ala80–Ser91) but not in *D. melanogaster* dNK or in human mitochondrial thymidine kinase (TK2). It was possible to model the insert in only one of the monomers. Preliminary kinetic data of insert-deleted variants suggest that the insert has a role in determining the substrate specificity of the enzyme (data not shown). (**b**) View of **a** rotated 90° (dashed arrow is the axis of rotation). The r.m.s. deviation between Cα atoms of dCK and dGK is 1.17 Å over 175 Cα atoms, and between dCK and dNK is 1.20 Å over 178 atoms. (**c**) Sequence alignment of dCK, dGK, dNK and TK2. Although only dCK and dGK contain an insert sequence, all have a P-loop, the ERS motif and a LID region (boxed in blue). Absolutely conserved residues are red; and similar residues are yellow. Figures were made using BobScript[36] and Raster3D[37].

provides the tool required to understand the substrate specificity of the enzyme and has great potential for contributing to the development of improved therapeutic agents that depend on dCK for activation. To date, the only structural knowledge of human nucleoside kinases was limited to a single structure at a resolution of 2.8 Å of deoxyguanosine kinase (dGK) in which ATP was unexpectedly bound to the deoxyguanosine-binding site[8].

Human dCK (**Fig. 1a**,**b**) consists of 260 amino acids with a calculated molecular weight of 30.5 kDa. We solved the structure using MAD[9] on a selenomethionine (SeMet)-containing protein. dCK is a homodimeric globular protein with a fold similar to that described for the *Drosophila melanogaster* nucleoside kinase (dNK) and human dGK[8]. Each monomer consists of a five-stranded, parallel β-sheet core surrounded by ten α-helices. Helices α4 and α7 from each monomer create a dimer interface composed of a four-helix bundle.

Three sequence motifs play a special role in dCK function. The first is the highly conserved P-loop motif located at residues Gly28–Ser35 (consensus sequence GX_4GKS/T; **Fig. 1c**), which binds and positions the α- and β-phosphoryl groups of the phosphoryl donor via interactions between amide backbone hydrogens and phosphate oxygen atoms. The second critical loop contains a glutamic acid (Glu127), an arginine (Arg128) and a serine residue (Ser129), which are strictly conserved within dCK, dGK, dNK and thymidine kinase 2 (TK2) (**Fig. 1c**). Here we observe a direct interaction between Glu127 of the ERS motif and the magnesium ion (**Fig. 2a**), which is the first report of a deoxyribonucleoside kinase structure in which magnesium is present. The hydroxyl group of the P-loop Ser35, a β-phosphate oxygen atom and three water molecules complete the octahedral magnesium coordination.

In the structurally related viral thymidine kinases, such as the herpes simplex virus type 1 enzyme, and in the thymidylate kinase family, the aspartate in a loop containing residues DR(Y/H/F) performs the function analogous to Glu127 in the ERS motif of dCK. Structures of nucleoside/nucleotide kinases with the ERS or DR(Y/H/F) motifs reveal that the arginine is involved in substrate binding and the third residue has a structural role. The ϕ/ψ angles of the arginine residues in the DR(Y/H/F) and the ERS motifs are located in the disallowed area of the Ramachandran plot[10]. Although such commonality suggests an evolutionary relationship between these enzymes, its significance is still not understood. In TMPK, the corresponding arginine (Arg97) acts as a clamp to bring the donor and acceptor nucleotides together by simultaneously interacting with the phosphate group of TMP and the γ-phosphate of ATP[6]. In dCK, Arg128 has an important anchor function: it interacts via its NH2 atom with the 5′-hydroxyl group of deoxycytidine (2.9 Å) and with Glu53 (3.0 Å), which is strictly conserved within dCK, dGK, dNK and TK2 enzymes (**Fig. 1**), and via its NH1 atom with the carbonyl of Gly28 (2.8 Å) in the P-loop and a water molecule (2.9 Å). Modeling an ATP molecule in place of ADP shows that one of the oxygen atoms of the γ-phosphate of ATP would most likely occupy the position of this water molecule, thereby directly interacting with Arg128, the P-loop residues Gly28 and Lys34, and the 5′-hydroxyl group of dC. In this way, Arg128 would play a similar catalytic role to that of Arg97 in the TMPK[6], bringing the 5′-hydroxyl group of dC in close proximity to the γ-phosphate of ATP (distance between modeled γ-phosphorous of ATP and the nucleophilic 5′-hydroxyl of dC is 3.6 Å).

The third loop is the so-called LID region, residues 188–195 (188-RIYLRGRN-195). This flexible segment, containing three conserved arginine residues, closes in on the phosphoryl donor when it binds and supplies critical residues that participate in ATP binding and catalysis. Arg188 provides the hydrophobic platform for the adenine ring of ADP by making π interactions. In the orthorhombic crystal form, Arg192 interacts with the α- and β-phosphates of ADP, an interaction artificially prevented in the tetragonal space group by symmetry-

related residues. In both crystal forms, Arg194 is at a distance of 5.5 Å from the β-phosphate of ADP. A further closing of the LID region is predicted in the presence of a γ-phosphate that would allow Arg194 to interact with the transferred phosphoryl group, making this residue important for transition state stabilization.

Substrate selectivity

In humans, four spatially segregated enzymes accomplish the first step in the salvage pathway of deoxyribonucleoside triphosphate synthesis. The mitochondrial kinases dGK and TK2 supply the precursors for mitochondrial DNA synthesis, and dCK and thymidine kinase 1 (TK1) do the same for nuclear DNA. Because ultimately dCK and TK1 must provide all four DNA triphosphates, it is not surprising that dCK, in addition to phosphorylating dC, also efficiently phosphorylates deoxyguanosine (dG) and deoxyadenosine (dA). In fact, apart from the thymidine-specific TK1, deoxyribonucleoside kinases are characterized by their ability to phosphorylate substrates of different base constituents. Notably, dNK, the only deoxyribonucleoside kinase in *D. melanogaster*, phosphorylates all four physiological deoxyribonucleosides[11].

The ability of dCK to accommodate multiple substrates of opposite hydrogen bonding character is conferred by the conserved Gln97 (**Fig. 1c**). The hydrogen bond–donating and –accepting moieties of the glutamine side chain are rotated and positioned according to the nature of the bound base. In our structures with cytosine nucleosides, the side chain of Gln97 acts as a hydrogen bond donor to the cytosine N3 atom via its amide group and as a hydrogen bond acceptor to the cytosine amino group via its carbonyl group (**Fig. 3**). This cytosine amino group also interacts with the side chain of Asp133. Discrimination by dCK against the pyrimidines thymidine and deoxyuridine is achieved because of the inability of Asp133 to perform favorable hydrogen-bonding interactions with a thymine or uracil base. Adding to the discrimination against thymidine (but not deoxyuridine) is a predicted steric clash between Arg104 and the thymine methyl group. Support for this interpretation comes from dNK mutation experiments[11] and from kinetic results with a dCK triple mutant we designed. In this mutant, the discriminating residues Arg104 and Asp133 are changed to a methionine and an alanine, respectively, and, as a consequence, the dCK mutant gains the ability to phosphorylate thymidine. Similar arguments to the above were formulated to explain the discrimination against thymidine by dGK[8]. Additional interactions between the cytosine base and dCK include a hydrophobic interaction to Phe137 and a hydrogen bond to a water molecule that bridges the base and Tyr86. The interaction with Tyr86 affects deoxyribose versus ribonucleosides discrimination (see below).

It is important to analyze how human dCK and dGK evolved to discriminate between different substrates despite their high conservation (47% sequence identity). Although both dCK and dGK phosphorylate dA and dG, only dCK is capable of phosphorylating dC and its analogs, such as AraC[12]. Can we understand the molecular basis for this difference in substrate selectivity, despite a remarkably conserved active site? Previously, Ser114 of dGK was postulated to be responsible for discriminating against pyrimidine deoxyribonucleosides, but the reason for this was not obvious[11,13]. The structures of dCK with dC reveal that to accommodate dC, Arg104 assumes a different conformation to that of Arg118 observed in the structure of dGK (**Fig. 4**, arrow 1). Why, then, does dCK alter the conformation of Arg104 to accommodate dC while dGK cannot do the same with Arg118? We propose that the answer lies in the different hydrogen-bonding networks occurring in the two active sites. An aspartate residue in both the dCK and dGK structures interacts with the above-mentioned arginine (**Fig. 4**, arrow 2). However, only in dGK is Ser114 able to participate in this network, adding to its stabilization. In contrast, the equivalent position to Ser114 of dGK in dCK contains an alanine (**Fig. 4**, arrow 3). The presence of alanine instead of serine weakens the hydrogen-bonding network that holds the aspartic acid and arginine residues rigid. This allows for a less extended conformation of Arg104, with a concomitant change in conformation for Asp133 upon dC binding. At the same time, when a purine binds to dCK, a change in Arg104 to an extended conformation as observed in the dGK structure is still possible. The conserved tyrosine in position 155 supplies an important hydrogen bond to Asp133 when it adopts the conformation compatible with purine binding. Based on this analysis, we predict

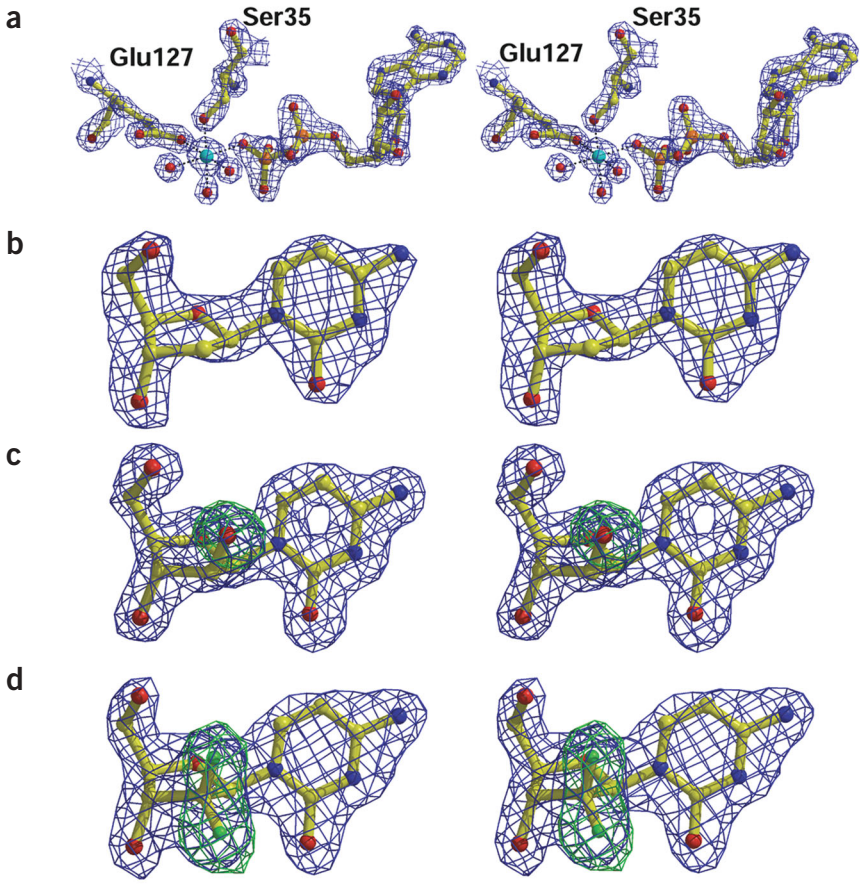

Figure 2 Electron density of the dCK-bound substrates. (**a**) The ADP molecule with an octahedrally coordinated magnesium ion (cyan), which is the first structure of a nucleoside kinase with the catalytically essential magnesium ion. (**b**) Deoxycytidine. (**c**) AraC. (**d**) Gemcitabine. The simulated annealing omit maps are shown contoured at 2.8 σ (blue), and the extra electron density for AraC (3.8 σ) and gemcitabine (3.5 σ) when refined in the presence of dC is indicated (green).

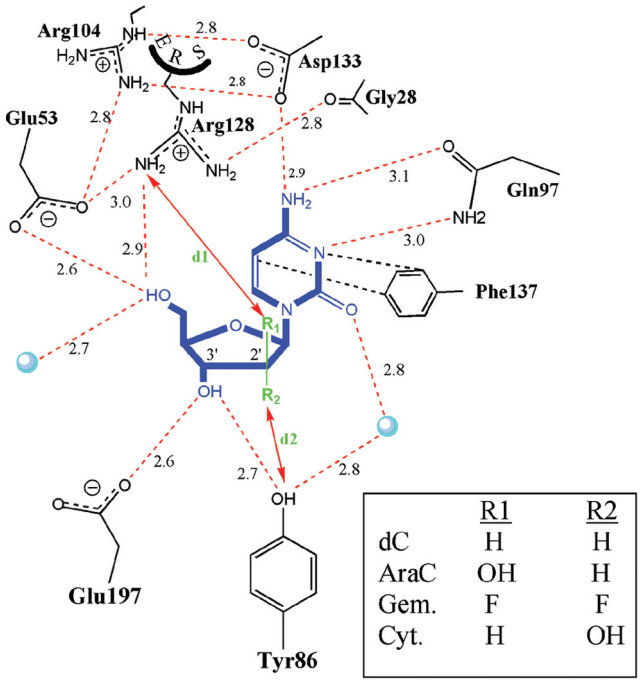

Figure 3 Schematic representation of the interactions made by the bound nucleoside with dCK. Interactions present only with the prodrugs are labeled d1 and d2. All distances are in Å.

that replacing Ser114 with an alanine in dGK will result in a mutant able to phosphorylate dC and, possibly, AraC. We are now testing this prediction.

A common feature among the nucleoside kinases is the stabilization of the sugar 3′-hydroxyl group by a conserved Tyr-Glu pair (dCK, Tyr86-Glu197; dGK, Tyr101-Glu211; dNK, Tyr70-Glu172)[8]. Because of the proximity of the tyrosine hydroxyl group to the 2′-sugar position, this residue favors deoxyribonucleosides over ribonucleosides. The isosteric substitutions of protons by fluorines at the 2′-position in gemcitabine puts one of these fluorines at a distance such that it is able to interact with tyrosine hydroxyl group. Even a slightly larger substituent at this position will result in some steric repulsion, as indicated by the high K_m of cytidine in comparison to deoxycytidine (**Table 1**).

Based on comparison of *in vitro* kinetic data for dCK with UTP and ATP, the preferred *in vivo* phosphoryl donor was proposed to be UTP[14,15]. Moreover, the kinetic behavior changes from a random bi-bi mechanism in the presence of ATP to an ordered mechanism of substrate addition in the case of UTP[16]. This suggests that the nature of the phosphoryl donor influences a conformational transition from a partially closed state, present when a single substrate is bound, to the fully closed state, as observed in our structure when both substrate-binding sites are occupied. Our lack of success in forming dCK crystals in complex with UDP, in contrast to the ADP complexes that resulted in high-quality crystals, is consistent with the predicted enzyme conformational dependency on the phosphoryl donor. Analysis of the phosphoryl donor–binding site in the closed conformation does not provide us with an obvious explanation for the preference of UTP over ATP, and this awaits a crystal structure of dCK in the presence of UTP or UDP.

DISCUSSION

Determinants of prodrug activation

The structure of dCK in complex with dC–ADP is similar to the complexes of dCK with AraC–ADP–Mg (r.m.s. deviation of 0.66 Å) and with gemcitabine–ADP–Mg (r.m.s. deviation of 0.63 Å) (**Fig. 2**). Compared with the structure with dC bound in the active site, the hydroxyl group in the 2′-position of the arabinofuranosyl sugar in AraC provides an additional hydrogen bond to the conserved Arg128 of the ERS motif (d1 = 3.0 Å, **Fig. 3**). A similar interaction (2.9 Å) with Arg128 is formed by the fluorine atom (R_1) of gemcitabine, while the other fluorine (R_2) makes a hydrogen bond with Tyr86 (d2 = 2.7 Å, **Fig. 3**). Steady-state kinetic experiments using AraC, gemcitabine and dC as substrates show that there is a 10-fold increase in the k_{cat}, but only a slight increase in K_m for the two prodrugs as compared with dC (**Table 1**). Overall there is a roughly four-fold higher enzymatic efficiency for dCK with AraC and gemcitabine.

We can reconcile these kinetic data by considering a reaction mechanism in which the proton from the sugar 5′-hydroxyl is accepted by a nearby base (either before the O5′ nucleophilic attack on the ATP γ-phosphate or after the formation of the O–P bond). A likely candidate for this role as the general base is the conserved Glu53 (ref. 17), which is at hydrogen-bonding distance to the sugar 5′-hydroxyl group (2.6 Å via OE2) and to NH1 of Arg128 (3.0 Å via OE1) (**Fig. 3**). In the three complexes, the atomic positions of Glu53, Arg128 and the nucleoside or prodrug are identical within experimental error. This suggests that the increased dCK activity towards the two cytosine analogs results from the interactions of Arg128 to R1 and of Tyr86 to R2. Because both AraC and gemcitabine possess the Arg128–R1 interaction and the increase in enzymatic efficiency for the two prodrugs is similar, the Arg128–R1 interaction presumably plays a more predomi-

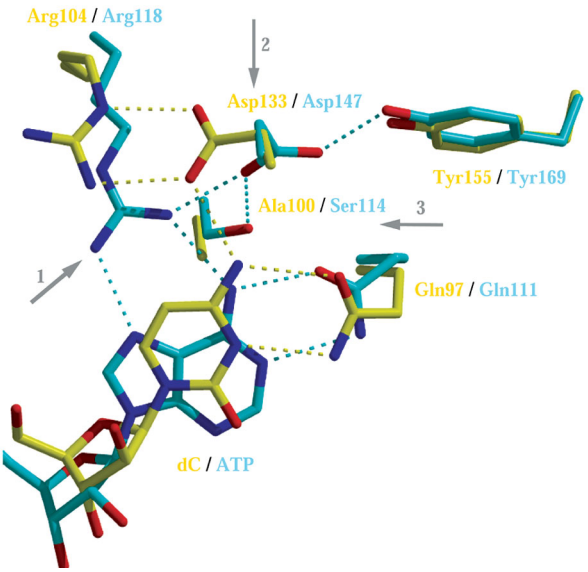

Figure 4 Superposition of the dCK (yellow) and dGK (cyan) active-site residues. Ser114 in dGK is the key residue for discrimination against pyrimidine bases. The corresponding Ala100 in dCK allows Asp133 to swing between the conformation adopted by the asparagine in the dCK–pyrimidine structure (hydrogen bonded to Arg104) and the one adopted by Asp147 in the dGK–purine structure (hydrogen bonded to Tyr155). In this way, dCK can accept both dC and dA/dG as substrates. Labels and hydrogen bonds in yellow correspond to dCK and those in cyan to dGK. For explanation of the arrows 1, 2 and 3, see text.

nant role than the Tyr86–R2 interaction. Assuming that the observed steady-state kinetic rate reflects the phosphoryl transfer step, the increased rate for AraC and gemcitabine could be explained as follows: in the presence of dC, the interaction between Glu53 and Arg128 is not weakened by the substrate. However, a hydrogen-bond acceptor at the 2′-arabinosyl position, as present in AraC and gemcitabine, could compete with Glu53 for the Arg128 interaction. A weaker Glu53–Arg128 interaction would potentiate the proton-accepting ability of the carboxylic acid group from the O5′-hydroxyl. As a result, the nucleoside 5′-hydroxyl group would become more nucleophilic; hence, the k_{cat} for AraC and gemcitabine phosphorylation would be increased. The positive influence of hydrogen-bonding acceptor substituents at the 2′-arabinosyl position on dCK activity is also supported by the recent discovery that the purine drug 2-chloro-2′-arabino-fluoro-2′-deoxyadenosine (clofarabine), a new derivative of cladribine that is currently in phase II clinical trial for treatment of pediatric leukemia, is about three-fold more efficient as a substrate for dCK than is cladribine[18–20]. On the basis of the structures presented here, we predict that this improvement is caused by the interaction of the 2′-arabino-fluorine atom of clofarabine with Arg128, thus improving the ability of Glu53 to function as a base.

Another important interaction involved in the stabilization and positioning of the nucleoside sugar is made by the 3′-hydroxyl group. The complexes with AraC and gemcitabine were crystallized under the same conditions as that of dC (see Methods). However, we were unable to obtain crystals with the antiviral drug 2′,3′-dideoxycytidine (ddC). This may suggest that weak ddC binding (K_m of 407 μM versus 6.2 μM for dC, **Table 1**) did not induce the same enzyme conformation necessary to promote crystal growth. A likely explanation for such behavior may be the absence of the 3′-hydroxyl group in ddC. Indeed, in all the dC, AraC and gemcitabine structures, the 3′-hydroxyl group is held in place by Glu197 and Tyr86 (**Fig. 3**), residues that are strictly conserved within the deoxyribonucleoside kinases dCK, dGK, dNK and TK2 (**Fig. 1c**). To counter the low phosphorylation efficiency of ddC and increase its antiviral effect, we propose a modified prodrug molecule that includes a hydrogen bond acceptor at the 2′-position—for example, a hydroxyl group or fluorine in the 2′-arabinosyl position or fluorines in both 2′-positions. When we searched for such molecules in the literature, we found that 2′,3′-dideoxy-2′-arabino-fluorocytosine has been synthesized and shown to be active in an HIV antiviral screen (http://dtp.nci.nih.gov/docs/static_pages/compounds/628495.html). Our results support the testing of this compound as an improved replacement over ddC.

Table 1 Steady state kinetic data

Wild-type dCK			
Nucleoside	K_m (μM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}\,s^{-1}$)
dC	6.2	0.03	4.8×10^3
ddC	406.8	0.18	0.4×10^3
Cytidine	382.7	0.74	1.9×10^3
AraC	15.5	0.26	16.8×10^3
Gemcitabine	22.0	0.37	16.8×10^3
A100V/R104M/D133A mutant dCK			
dC	4	0.93	232.5×10^3
AraC	102	1.73	17.0×10^3
Gemcitabine	50	3.52	70.4×10^3

The enzyme dCK is also responsible for the initial phosphorylation of non-natural L-nucleoside prodrugs[21]. These include 3TC, which is clinically approved for treating HIV infection[22], and the newly developed troxacitabine, which is currently in clinical trials against human malignancies[23–25]. Modeling of troxacitabine in the dCK active site by overlaying of the common base moiety with the physiological substrate dC demonstrates ample space for the dioxolane ring, and rotations around the torsion angles brings the 5′-OH to the same position as observed with dC. However, although this binding mode is suggestive of a rationale for the ability of dCK to phosphorylate these non-natural nucleosides, it assumes a similar enzyme and nucleoside conformation, which must still be verified experimentally.

Improving dCK catalytic efficiency

dCK is an inefficient enzyme as a result of a slow k_{cat} (0.03 s^{-1}; **Table 1**). In contrast, dNK exhibits a 2,500-fold higher efficiency for dC phosphorylation resulting mainly from its faster k_{cat} (16.5 s^{-1})[11]. With the goal of making dCK more active, we followed a strategy of mutating key active site residues in dCK to those found in dNK. We hypothesized that Arg104 bends away from the substrate dC to minimize an otherwise detrimental interaction between an extended conformation of Arg104 (as seen in the structure of dGK) and the amino group of dC, and that a methionine in its place, as found in dNK, would result in an increased phosphorylation rate. Because Arg104 is involved in the already mentioned active-site hydrogen-bonding network (**Fig. 4**), its mutation to the uncharged methionine was combined with the compensatory D133A mutation. In addition, we made the A100V change to match the residue found in dNK. The kinetic data for the dCK triple mutant A100V R104M D133A are encouraging: efficiency toward dC increased by 50-fold and that toward gemcitabine by four-fold, with no change for AraC (**Table 1**). However, for suicide gene-therapy applications, an enzyme with higher efficiency in relation to nucleoside analogs as compared to dC would be advantageous.

Concluding remarks

The high-resolution structures of dCK in complex with dC and two clinically used prodrugs reveal determinants of substrate specificity. In addition, the structures provide an ideal starting point for the design of superior nucleoside analogs, taking into account the structural constraints imposed by the dCK active site. Because it is the dCK phosphorylation step that limits the efficacy of clinically used prodrugs in most cases, more efficient prodrug phosphorylation by dCK is predicted to result in potentiated antiviral and anticancer response. Lastly, we have shown that the structure of the enzyme can be used as a first step toward designing a dCK mutant of improved activity that can be used in suicide gene-therapy applications.

METHODS

Protein expression, purification and crystallization. Human dCK was amplified from a cDNA library and cloned into the pET14b vector (S.O., unpublished data). BL21(DE3) *Escherichia coli* carrying the recombinant plasmid coding for His-tagged dCK was grown in 2YT media at 37 °C, induced with 0.1 mM IPTG and harvested after 8 h. The cell pellet was lysed by sonication and loaded onto a TALON Co^{2+}-affinity column (Clontech). After washing and elution with imidazole, the His-tag was cleaved by thrombin. The protein was further purified on an ion-exchange column and an S-200 gel filtration column (Amersham Biosciences). SeMet-substituted protein was produced using an established protocol[26].

Crystals of human dCK in complex with nucleosides and nucleotides (dC and ADP, AraC and ADP, gemcitabine and ADP) were grown by the vapor diffusion method using either sitting-drop or hanging-drop geometry. Nucleotides were from Sigma except for gemcitabine, which was a gift from Eli

Table 2 Data collection, phasing and refinement statistics

		MAD		dC–ADP	dC–ADP	AraC–ADP–Mg	gemcitabine–ADP–Mg
Data collection statistics							
Beamline		14-BM-D		14-BM-C	14-BM-C	22-ID	22-ID
Wavelength (Å)	0.97997	0.97973	0.95705	0.9	0.9	1.0	1.0
Temperature (K)		100		100	100	100	100
Resolution (Å)[a]		2.3		1.96 (1.96–2.0)	2.2 (2.2–2.3)	1.6 (1.6–1.7)	1.9 (1.9–1.95)
Reflections							
Observed	257,787	257,933	258,286	270,973	219,344	487,680	326,066
Unique	24,032	23,928	24,055	39,237	15,406	41,313	27,499
Completeness (%)[a]	94.5 (64.0)	94.0 (61.8)	94.6 (64.8)	97.4 (67.8)	95.9 (78.2)	99.0 (99.6)	99.5 (99.4)
R_{sym} (%)[a,b]	4.5 (31.1)	4.7 (32.1)	4.6 (33.1)	4.5 (29.1)	6.4 (15.6)	5.2 (51.1)	5.4 (53.3)
$I / \sigma (I)$[a]	29.5 (3.3)	27.8 (3.1)	28.9 (2.8)	23.7 (3.3)	28.2 (7.5)	23.1 (4.9)	24.2 (4.7)
Space group		$P4_32_12$		$C222_1$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Unit cell (Å)							
a		79.64		52.74	80.00	80.72	80.20
b		79.64		132.92	80.00	80.72	80.20
c		93.71		157.64	93.95	94.28	94.60
Molecules per a.u.		1		2	1	1	1
Phasing statistics							
Heavy atom sites (Se)	4						
Figure of merit		0.54					
Refinement statistics							
R_{cryst} (%)[c]				16.1	22.1	17.3	17.9
R_{free} (%)[d]				20.3	28.0	19.7	20.7
Resolution range (Å)				20–1.96	20–2.2	20–1.6	20–1.9
Number of atoms							
Protein				1,990 (A); 1,830 (B)	1,869	1,982	1,892
Nucleoside				16 × 2	16	17	18
ADP				27 × 2	27	27	27
Water				333	131	173	141
R.m.s. deviation							
Bond length (Å)				0.020	0.021	0.018	0.021
Bond angles (°)				1.922	1.885	1.770	1.966
Average *B*-factor (Å^2)							
Protein							
Main chain				27 (A); 28 (B)	43	27	34
Side chain				29 (A); 31 (B)	44	30	37
ADP				22 (A); 23 (B)	37	22	30
Nucleoside				20 (A); 22 (B)	33	20	26
Waters				37	47	37	41

[a]Numbers in parentheses are for the highest resolution shell. [b]$R_{sym} = \Sigma_{hkl} |I(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl} \langle I(hkl)\rangle$. [c]$R_{cryst} = \Sigma_{hkl} |F_o(hkl) - F_c(hkl)| / \Sigma_{hkl} |F_o(hkl)|$. [d]$R_{free} = R_{cryst}$, calculated for 10% randomly selected reflections not included in the refinement.

Lilly & Co. After the respective complex was formed by mixing of the nucleoside or nucleoside analog (dC, AraC or gemcitabine) and ADP (final concentrations 5 mM each) and the dialyzed enzyme solution (12 mg ml^{-1} dCK, 5 mM $MgCl_2$, 20 mM HEPES, pH 7.5, 5 mM DTT and 100 μM dC, AraC or gemcitabine), 2 μl of the premixed solution was added to 2 μl of the reservoir solution and left to equilibrate at 20 °C against the reservoir. Tetragonal crystals were obtained from sitting drops using a reservoir solution that contained 0.95–1.0 M trisodium citrate dihydrate and 100 mM HEPES, pH 7.5, or from hanging drops using 20% (w/v) PEG 1000, 100 mM magnesium acetate and 100 mM Tris, pH 9.0. In some cases, the citrate condition also gave orthorhombic crystals. Typically, crystals grew within 1 week to dimensions of 200 × 100 × 100 μm^3.

Data collection and processing. Crystals were transferred to a cryoprotectant solution, consisting of mineral oil (Sigma) for the citrate condition and consisting of the mother liquor with 10% (w/v) xylitol and 10% (w/v) sucrose for the crystals grown in PEG. Once the crystals were mounted in loops, they were frozen by directly immersing them in liquid nitrogen. X-ray data (**Table 2**) were collected at the Advanced Photon Source using the BioCARS beamlines BM-C and BM-D and the SERCAT beamline ID-22. The data were indexed, scaled and merged using XDS and XSCALE[27], or DENZO and SCALEPACK[28]. Data collection statistics for all data sets are shown (**Table 2**).

Structure determination and refinement. The structure of human dCK was solved using the MAD method[9]. Using data from the inflection, peak and remote wavelengths, we located four selenium atoms using SOLVE[29]. The map calculated from the experimental phases allowed us to build a model of dCK in O[30]. Refinement was carried out using CNS[31] and REFMAC[32]. Data for the MAD data set were collected to a resolution of 2.3 Å. A partially built model

into the MAD electron density was used to solve the tetragonal crystals. The electron density for dCK extended to the last residue (Leu260), whereas the N terminus (residues 1–19) was flexible and could not be modeled. Deoxycytidine and ADP were modeled in the electron density map after most of the protein main chain and side chain atoms were built and refined. The ligands were refined without conformational torsion-angle restraints to prevent bias toward a particular ring conformation for the sugars. The final model of the complex of dCK with dC and ADP was used as starting model for the refinement of the other structures in the tetragonal space group and was also used to solve the structure of dCK that crystallized in the orthorhombic space group by molecular replacement (using AMoRe[33]).

Data for the AraC–ADP–Mg and the gemcitabine–ADP–Mg complexes were collected to resolutions of 1.6 Å and 1.9 Å, respectively, on the crystals grown in PEG. The calculation of difference maps clearly showed the presence of substituents on the second position of the deoxyribose moiety of dC: a hydroxyl group in the case of AraC and two fluorine atoms in the case of gemcitabine. The ligands were modeled into the electron density maps after simulated annealing and refinement of the protein model. Water molecules were then automatically added using ARP/wARP[34].

Steady-state kinetic assay. Deoxycytidine kinase activity was determined using a colorimetric assay[35] in 50 mM Tris-HCl, pH 7.5, 100 mM KCl and 5 mM $MgCl_2$ at 37 °C. dCK was at a concentration of 0.4 μM and ATP–Mg at 1 mM, while for the nucleosides and the nucleoside analogs we used 10 μM–1 mM concentrations. All experiments were performed in duplicate.

Coordinates. Coordinates and structure factors have been deposited with the Protein Data Bank (accession codes 1P5Z (AraC–ADP–Mg), 1P60 (dC–ADP, orthorhombic space group), 1P61 (dC–ADP, tetragonal space group), and 1P62 (gemcitabine–ADP–Mg).

ACKNOWLEDGMENTS

We acknowledge the contributions made by K. Storie, L. Shuvalova and A. Arockiasamy in the initial stages of the project, and thank B. Beck and M. Godsey for careful reading of the manuscript. We thank Eli Lilly & Co. for providing gemcitabine as a gift. We thank the staff of BioCars and SerCAT for help in data collection at the Advanced Photon Source. E.S. and A.L. were supported by the US National Institutes of Health (NIH), and S.O., C.M. and M.K. were supported by grants from the Deutsche Forschungsgemeinschaft and the Max Planck Society.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 11 March; accepted 21 May 2003
Published online 15 June 2003; doi:10.1038/nsb942

1. Hapke, D.M., Stegmann, A.P. & Mitchell, B.S. Retroviral transfer of deoxycytidine kinase into tumor cell lines enhances nucleoside toxicity. *Cancer Res.* **56**, 2343–2347 (1996).
2. Owens, J.K., Shewach, D.S., Ullman, B. & Mitchell, B.S. Resistance to 1-β-D-arabinofuranosylcytosine in human T-lymphoblasts mediated by mutations within the deoxycytidine kinase gene. *Cancer Res.* **52**, 2389–2393 (1992).
3. Ruiz van Haperen, V.W. *et al.* Development and molecular characterization of a 2′,2′-difluorodeoxycytidine-resistant variant of the human ovarian carcinoma cell line A2780. *Cancer Res.* **54**, 4138–4143 (1994).
4. Stegmann, A.P., Honders, W.H., Willemze, R., Ruiz van Haperen, V.W. & Landegent, J.E. Transfection of wild-type deoxycytidine kinase (dCK) cDNA into an AraC- and DAC-resistant rat leukemic cell line of clonal origin fully restores drug sensitivity. *Blood* **85**, 1188–1194 (1995).
5. Blackstock, A.W. *et al.* Tumor uptake and elimination of 2′,2′-difluoro-2′-deoxycytidine (gemcitabine) after deoxycytidine kinase gene transfer: correlation with *in vivo* tumor response. *Clin. Cancer Res.* **7**, 3263–3268 (2001).
6. Ostermann, N. *et al.* Insights into the phosphoryltransfer mechanism of human thymidylate kinase gained from crystal structures of enzyme complexes along the reaction coordinate. *Structure* **8**, 629–642 (2000).
7. Bergman, A., Pinedo, H. & Peters, G. Determinants of resistance to 2′,2′-difluorodeoxycytidine (gemcitabine). *Drug Resist. Updat.* **5**, 19–33 (2002).
8. Johansson, K. *et al.* Structural basis for substrate specificities of cellular deoxyribonucleoside kinases. *Nat. Struct. Biol.* **8**, 616–620 (2001).
9. Hendrickson, W.A. & Ogata, C.M. Phase determination from multiwavelength anomalous diffraction measurements. *Methods Enzymol.* **276**, 494–523 (1997).
10. Ramachandran, G.N., Ramakrishnan, C. & Sasisekharan, V. Stereochemistry of polypeptide chain configuration. *J. Mol. Biol.* **33**, 491–497 (1963).
11. Knecht, W. *et al.* A few amino acid substitutions can convert deoxyribonucleoside kinase specificity from pyrimidines to purines. *EMBO J.* **21**, 1873–1880 (2002).
12. Herrstrom Sjoberg, A., Wang, L. & Eriksson, S. Substrate specificity of human recombinant mitochondrial deoxyguanosine kinase with cytostatic and antiviral purine and pyrimidine analogs. *Mol. Pharmacol.* **53**, 270–273 (1998).
13. Usova, E.V. & Eriksson, S. Identification of residues involved in the substrate specificity of human and murine dCK. *Biochem. Pharmacol.* **64**, 1559–1567 (2002).
14. White, J.C. & Capizzi, R.L. A critical role for uridine nucleotides in the regulation of deoxycytidine kinase and the concentration dependence of 1-β-D-arabinofuranosylcytosine phosphorylation in human leukemia cells. *Cancer Res.* **51**, 2559–2565 (1991).
15. Shewach, D.S., Reynolds, K.K. & Hertel, L. Nucleotide specificity of human deoxycytidine kinase. *Mol. Pharmacol.* **42**, 518–524 (1992).
16. Hughes, T.L., Hahn, T.M., Reynolds, K.K. & Shewach, D.S. Kinetic analysis of human deoxycytidine kinase with the true phosphate donor uridine triphosphate. *Biochemistry* **36**, 7540–7547 (1997).
17. Eriksson, S., Munch-Petersen, B., Johansson, K. & Eklund, H. Structure and function of cellular deoxyribonucleoside kinases. *Cell. Mol. Life Sci.* **59**, 1327–1346 (2002).
18. Lotfi, K. *et al.* Biochemical pharmacology and resistance to 2-chloro-2′-arabino-fluoro-2′-deoxyadenosine, a novel analogue of cladribine in human leukemic cells. *Clin. Cancer Res.* **5**, 2438–2444 (1999).
19. Mansson, E. *et al.* Down-regulation of deoxycytidine kinase in human leukemic cell lines resistant to cladribine and clofarabine and increased ribonucleotide reductase activity contributes to fludarabine resistance. *Biochem. Pharmacol.* **65**, 237–247 (2003).
20. Kantarjian, H.M. *et al.* Phase I clinical and pharmacology study of clofarabine in patients with solid and hematologic cancers. *J. Clin. Oncol.* **21**, 1167–1173 (2003).
21. Wang, J., Choudhury, D., Chattopadhyaya, J. & Eriksson, S. Stereoisomeric selectivity of human deoxyribonucleoside kinases. *Biochemistry* **38**, 16993–16999 (1999).
22. Kewn, S., Veal, G.J., Hoggard, P.G., Barry, M.G. & Back, D.J. Lamivudine (3TC) phosphorylation and drug interactions *in vitro*. *Biochem. Pharmacol.* **54**, 589–595 (1997).
23. Gourdeau, H. *et al.* Mechanism of uptake and resistance to troxacitabine, a novel deoxycytidine nucleoside analogue, in human leukemic and solid tumor cell lines. *Cancer Res.* **61**, 7217–7224 (2001).
24. Townsley, C.A. *et al.* Phase II study of troxacitabine (BCH-4556) in patients with advanced and/or metastatic renal cell carcinoma: a trial of the National Cancer Institute of Canada-Clinical Trials Group. *J. Clin. Oncol.* **21**, 1524–1529 (2003).
25. Giles, F.J. *et al.* Randomized phase I/II study of troxacitabine combined with cytarabine, idarubicin, or topotecan in patients with refractory myeloid leukemias. *J. Clin. Oncol.* **21**, 1050–1056 (2003).
26. Doublie, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* **276**, 523–530 (1997).
27. Kabsch, W. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Crystal.* **24**, 795–800 (1993).
28. Otwinowski, Z. & Minor, W. Processing of X-ray data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
29. Terwilliger, T.C. Reciprocal-space solvent flattening. *Acta Crystallogr. D* **55**, 1863–1871 (1999).
30. Jones, T.A., Zhou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).
31. Brünger, A.T. *X-PLOR: A System for X-ray Crystallography and NMR* (Yale University Press, New Haven, 1993).
32. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of macromolecular structures by the maximum likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
33. Navaza, J. AMoRe: an automated package for molecular replacement. *Acta Crystallogr. A* **50**, 157–163 (1994).
34. Perrakis, A., Morris, R. & Lamzin, V.S. Automated protein model building combined with iterative structure refinement. *Nat. Struct. Biol.* **6**, 458–463 (1999).
35. Agarwal, K.C., Miech, R.P. & Parks, R.E. Jr. Guanylate kinases from human erythrocytes, hog brain, and rat liver. *Methods Enzymol.* **51**, 483–490 (1978).
36. Esnouf, R. An extensively modified version of MOLSCRIPT that includes greatly enhanced coloring capabilities. *J. Mol. Graph.* **15**, 133–138 (1997).
37. Merrit, E.A. & Murphy, M.E.P. Raster3D version 2.0—a program for photorealistic molecular graphics. *Acta Crystallogr. D* **50**, 869–873 (1994).