Available online at www.sciencedirect.com
SCIENCE DIRECT®
Bioorganic & Medicinal Chemistry 13 (2005) 4160–4167
Bioorganic & Medicinal Chemistry

Docking simulation with a purine nucleoside specific homology model of deoxycytidine kinase, a target enzyme for anticancer and antiviral therapy

Jayaseharan Johnsamuel,[a,*] Staffan Eriksson,[b] Marcos Oliveira[c] and Werner Tjarks[a]

[a]*Division of Medicinal Chemistry, College of Pharmacy, The Ohio State University, Columbus, OH 43210, USA*
[b]*Department of Molecular Biosciences, Section of Veterinary Medical Biochemistry; Swedish University of Agricultural Sciences, The Biomedical Center, SE-75 123 Uppsala, Sweden*
[c]*Department of Pharmaceutical Sciences, College of Pharmacy, University of Kentucky, Lexington, KY 40536-0082, USA*

Received 1 March 2005; revised 14 April 2005; accepted 15 April 2005
Available online 3 May 2005

Abstract—5′-Phosphorylation, catalyzed by human deoxycytidine kinase (dCK), is a crucial step in the metabolic activation of anticancer and antiviral nucleoside antimetabolites, such as cytarabine (AraC), gemcitabine, cladribine (CdA), and lamivudine. Recently, crystal structures of dCK (dCKc) with various pyrimidine nucleosides as substrates have been reported. However, there is no crystal structure of dCK with a bound purine nucleoside, although purines are good substrates for dCK. We have developed a model of dCK (dCKm) specific for purine nucleosides based on the crystal structure of purine nucleoside bound deoxyguanosine kinase (dGKc) as the template. dCKm is essential for computer aided molecular design (CAMD) of novel anticancer and antiviral drugs that are based on purine nucleosides since these did not bind to dCKc in our docking experiments. The active site of dCKm was larger than that of dCKc and the amino acid (aa) residues of dCKm and dCKc, in particular Y86, Q97, D133, R104, R128, and E197, were not in identical positions. Comparative docking simulations of deoxycytidine (dC), cytidine (Cyd), AraC, CdA, deoxyadenosine (dA), and deoxyguanosine (dG) with dCKm and dCKc were carried out using the FlexX™ docking program. Only dC (pyrimidine nucleoside) docked into the active site of dCKc but not the purine nucleosides dG and dA. As expected, the active site of dCKm appeared to be more adapted to bind purine nucleosides than the pyrimidine nucleosides. While water molecules were essential for docking experiments using dCKc, the absence of water molecules in dCKm did not affect the ability to correctly dock various purine nucleosides.

1. Introduction

Human deoxycytidine kinase (dCK) is a key regulatory enzyme required in the initial phosphorylation of endogenous deoxyribonucleosides and many synthetic analogues that are widely used as anticancer and antiviral agents. Cell lines lacking dCK activity were found to be resistant to a variety of drugs including cytarabine (AraC), cladribine (CdA), fludarabine, and gemcitabine.[1,2] The broad specificity of dCK is the basis for its role in activation of nucleosides with substantial structural modifications as found in many antiviral or anticancer agents. Structural information on dCK is vital for understanding the specificity of the enzyme as well as in the design of new drugs using advanced computer aided molecular design (CAMD).

dCK phosphorylates three natural deoxyribonucleosides, the pyrimidine nucleoside deoxycytidine (dC)

Abbreviations: Ara-C, cytarabine; aa, amino acid; CAMD, computer aided molecular design; c, crystal; CdA, cladribine; Cyd, cytidine; dCKc, human deoxycytidine kinase crystal; dCKm, homology model of deoxycytidine kinase; dCKc(ww), dCKc with water molecules; dCKc(wow), dCKc without water molecules; dGKc, human deoxyguanosine kinase crystal; dC, deoxycytidine; dA, deoxyadenosine; dG, deoxyguanosine; dT, deoxythymidine; 3D, three-dimensional; HSV1-TK, herpes simplex virus type-1 thymidine kinase; K_d, substrate binding affinity; K_M, Michaelis–Mentens constant; k_{cat}/K_M, measure of efficiency of substrate phosphorylation; m, model; SVR, structurally variable region; SCR, structurally conserved regions; V_{max}, maximum rate of phosphorylation catalyzed by a fixed amount of enzyme; VZV TK, varicella zoster virus thymidine kinase

Keywords: Deoxycytidine kinase; Computer aided molecular design; Docking simulation; Homology modeling; Antiviral and anticancer therapy.

* Corresponding author. Tel.: +1 614 668 3149; fax: +1 614 292 2435; e-mail: johnsamuel.1@osu.edu

0968-0896/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2005.04.037

and the purine nucleosides deoxyadenosine (dA) and deoxyguanosine (dG). dC has both lower K_M and V_{max} values than dA and dG[3] and there are large discrepancies between the K_i and K_M values of pyrimidine- and purine deoxyribonucleoside substrates.[4] dC is an efficient inhibitor in the phosphorylation of other substrates, while both dA and dG are only very weak inhibitors of dC phosphorylation by dCK. Bohman and Eriksson suggested a model with two different conformational states of dCK to explain this complex kinetic behavior.[5,6] In another study, it was proposed that dCK is a rigid protein with two different basic active sites, one with low affinity and another with high affinity for the ligands, but with both sites being able to bind purine as well as pyrimidine nucleosides.[7–10]

The dCK structure has been the subject of intense research activities.[5,11–16] Several attempts were made to elucidate structural features of dCK using spectroscopic,[13,17] fluorescence,[15] NMR,[8,9] and crystallographic techniques.[16] Only recently, crystal structures of dCK (dCKc) with various pyrimidine nucleosides were obtained,[16] which presumably represent the pyrimidine nucleoside specific conformations of dCK. So far, crystallization of dCK with purine nucleosides has not been successful. This may be difficult to accomplish because of the low affinity of purine nucleosides for the active site of dCK. To the best of our knowledge, no attempts have been described in the literature to obtain a homology model of the purine nucleoside specific form of dCK. Thus, there is no structural information, which could confirm or dispute the hypothesis that there are two distinct substrate sites or substrate dependent conformations of dCK.

The lack of a purine specific structure for dCKc and the necessity for such a model in the structure-based drug design of anticancer and antiviral purine nucleoside derivatives led us to develop a purine nucleoside specific model (dCKm) using the crystal structure of deoxyguanosine kinase (dGKc) as a template. Comparative docking simulations of dC, cytidine (Cyd), dA, and dG, as well as the anticancer drugs CdA and Ara-C with dCKc and dCKm were carried out to analyze differences in substrate binding modes and to evaluate the suitability of both structures to serve as templates for CAMD.

2. Computational methods

2.1. Crystal and nucleoside structures

dGKc (ID code: 1jag; resolution: 2.8 Å; *R*-value: 0.20), dCKc (ID code: 1p60; resolution: 1.96 Å; *R*-value: 0.165; dC substrate), and dCKc (ID code: 1P5Z; resolution: 1.6 Å; *R*-value: 0.175; Ara-C substrate) were obtained from the protein data bank [Research Collaboratory for Structural Bioinformatics (RCSB) (http://www.rcsb.org/pdb)].

Active site volumes of dCKc, dCKm, and dGKc were calculated by generating the binding pockets of the enzymes using the *Find pocket option* in the SiteID program of SYBYL 6.9 (Tripos, Inc., 1699 South Hanley Road, St. Louis, MO 63144). For these calculations, hypothetical cylindrical shapes of the binding pockets were constructed based on the length and width of the MOLCAD surfaces, generated by SiteID.

Structures dC, dG, dA (Fig. 5A–C), AraC, Cyd, and CdA, were built and minimized using the TRIPOS force field and Gasteiger–Hückel charges until an energy gradient of 0.005 kcal mol^{-1} $Å^{-1}$ was reached using SYBYL 6.9 on a Silicon Graphics O_2 system (Silicon Graphics, Inc., 1600 Amphitheatre pkwy., Mountainview, CA 94043). The volumes of dC, dG, and dA were measured using the *compute* option of HyperChem version 7.5 (Hypercube Inc., Waterloo, Ontario, Canada).

2.2. Homology modeling

Homology modeling was carried out using the 'Homology module'[18] of Insight II (Accelrys Inc., Cerius Modeling Environment, Release 4.0, San Diego: Accelrys Inc. 2001). The sequences of dGKc and dCK were aligned using the multiple sequence alignment methods provided in the homology module of Insight II. Structurally variable regions were assigned using the *generating loop* option of Insight II. SYBYL 6.9 was used for further refinement of dCK by minimization with Tripos force field and Gasteiger–Hückel charges (500 iterations). The final structure of the dCK model was examined with both PROFILES-3D,[19] associated with Insight II, and PROCHECK,[20] which predict the accuracy and stereochemical quality.

2.3. Docking simulations

Docking simulations were carried out using FlexX,™[21] available with SYBYL 6.9, which is a fast algorithm for the flexible docking of small ligands into fixed protein binding sites using an incremental construction process. The active site regions for the comparative FlexX docking simulations of dC, Cyd, dA, dG, CdA, and Ara-C with dCKm and dCKc were constructed based on previously reported structural information.[16,18]

The obtained docking simulations were scored using the CScore™ program, which is a consensus scoring program that integrates multiple well known and extensively applied scoring functions[21] available with SYBYL 6.9. CScore™ scoring functions include root mean square distance (rmsd) values,[22] Chemscore (scores based on a diverse set of 82 ligand–receptor complexes),[23] D_score (scores based on both electrostatic and hydrophobic contributions to the binding energy),[24] G_score (scores ligand–receptor complexes having many polar interactions),[25] FlexX_score (based on empirical functions),[21] and PMF_score (scores based on statistical ligand–receptor atom-pair interaction potentials).[26]

In case of dCKc, docking was carried out with and without active site water molecules using the *Place Particles* option of SYBYL 6.9. Docking with dCKm was carried out without placed water molecules. The terminology used in the following for water-containing and water-

Table 1. Interaction distances between functional groups of substrates and active site aa residues

Enzyme	Subst	D133	Q97	Y86	E197	Water-5′-OH	E53	R104	R128	Config.[h]
dCKc	*dC*[a]	*2.9*	*3.1*	*2.7*	*2.6*	*2.7*	*2.6*	*5.1*	*2.9*	*anti*
dCKc(ww[e])[i]	dC[b]	2.8	2.8	2.8	2.9	2.3	3.1	6.8	4.5	*anti*
dCKc(ww[e])	**dC**[c]	**2.9**	**2.9**	**2.8**	**2.9**	**2.5**	**3.4**	**6.8**	**4.6**	***anti***
dCKc(wow[f])	dC[b]	2.5	2.9	2.7	2.8	—	5.0	7.0	6.0	*anti*
dCKc(ww[e])	dA[b]	2.3	5.7	3.0	3.4	3.0	3.1	6.5	4.0	*syn*
dCKc(wow[f])	dA[b]	4.0	2.6	3.5	2.2	—	4.6	3.8	2.7	*anti*
dCKc(ww[e])	dG[b]	13.3	18.4	12.5	8.9	8.5	3.8	12.0	5.4	*syn*
dCKc(wow[f])	dG[b]	9.1	6.1	9.1	6.2	—	5.1	9.1	2.9	*syn*
dCKm[i]	dC[d]	5.0	2.7	4.3	3.0	—	2.9	2.7	3.5	*syn*
dCKm	dA[d]	4.2	3.0	3.9	2.9	—	2.9	3.1	3.9	*anti*
dCKm	dG[d]	6.3	3.0	2.6	2.8	—	2.6	2.7	2.8	*anti*
dGKc	*ATP*[g]	*4.0 (D147)*	*2.7 (Q111)*	*2.7 (Y100)*	*2.6 (E211)*	*3.8*	*6.6 (E70)*	*2.7 (R118)*	*4.0 (R142)*	*anti*

[a] Data taken from Ref. 16 with dC as the substrate of dCKc.
[b] Data for dC, dA, and dG docked into dCKc.
[c] Data of dC docked into dCKc using Autodock3.0.
[d] Data for dC, dA, and dG docked into dCKm.
[e] Docking was carried out with crystal water molecules.
[f] Docking was carried out without crystal water molecules.
[g] Data taken from Ref. 34 with ATP as substrate in the nucleoside binding site of dGKc.
[h] *syn* and *anti* refers to the conformation of the base portion of the dC, dA, and dG. The ribose conformation was in all cases C3′-*endo*.
[i] Details of interaction distances between specific groups of ligands and active site aa's of dCKc(ww) and dCKm are shown in Figure 5A–C.

Table 2. FlexX_score, G_score, and rsmd data of docking simulations and the corresponding K_d values

Enzyme	Substr.	rmsd	FlexX	G_score	K_d (μM)[17]
dCKc(ww[a])	dC	2.1	−22	−211	1.25 ± 0.10
dCKc(ww[a])	AraC	2.8	−17	−191	1.60 ± 0.15
dCKc(ww[a])	Cyd	5.0	−16	−178	2.70 ± 0.20
dCKm	CdA	2.5	−17	−180	0.95 ± 0.10
dCKm	dA	2.2	−15	−175	1.90 ± 0.10
dCKm	dG	0	−11	−114	3.90 ± 0.20

[a] Docking was carried out with water molecules.

free dCKc are dCKc(ww) and dCKc(wow) [see also footnote of Tables 1 and 2].

2.4. Biological data

K_d values of dC, AraC, Cyd, CdA, dA, and dG with dCK, obtained from previous studies,[17] were used for correlation with the CScore™ scoring functions of the nucleosides obtained from the docking simulations. The binding affinity (K_d) of a substrate to dCK is here defined as minimum concentration of a nucleoside (in μM) required to obtain half maximum fluorescence intensity at 335 nm with dCK (0.4 μM).[13,17] As determined previously,[27,28] k_{cat}/K_M values, which define the efficiency of phosphorylation of a nucleoside by dCK, are not indicative of the affinity or binding strengths between nucleosides and an enzyme such as dCK.

3. Results and discussion

3.1. Homology modeling

Our approach was to use the homologous structure of the purine nucleoside specific conformation of dGKc (with ATP bound) as a template in order to develop a purine nucleoside specific model of dCK for CAMD. It should be noted that the resolution of dGKc is 2.8 Å, which is marginal for modeling studies.[29] However, dGKc, having 53% sequence identity with dCK, is currently the only purine nucleoside specific kinase available for the development of a purine nucleoside specific dCK.

The sequence alignment of dCK and dGKc is shown in Figure 1. The boxed regions are structurally conserved (SCR) and the unboxed regions are structurally variable regions (SVR). The sequence alignments showed that

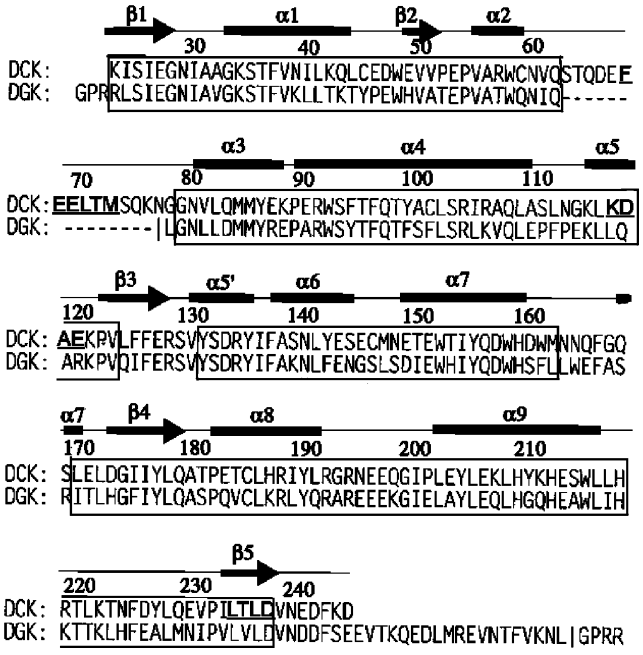


Figure 1. Sequence alignment of dCK and dGKc. The sequences inside the boxes are SCRs and the loop sequences are SVRs. The defective aa's are underlined.

84% (excluding the tail region) of the dGKc sequence was conserved in the dCK sequence, indicating a close structural resemblance between purine nucleoside specific dGKc and dCK.

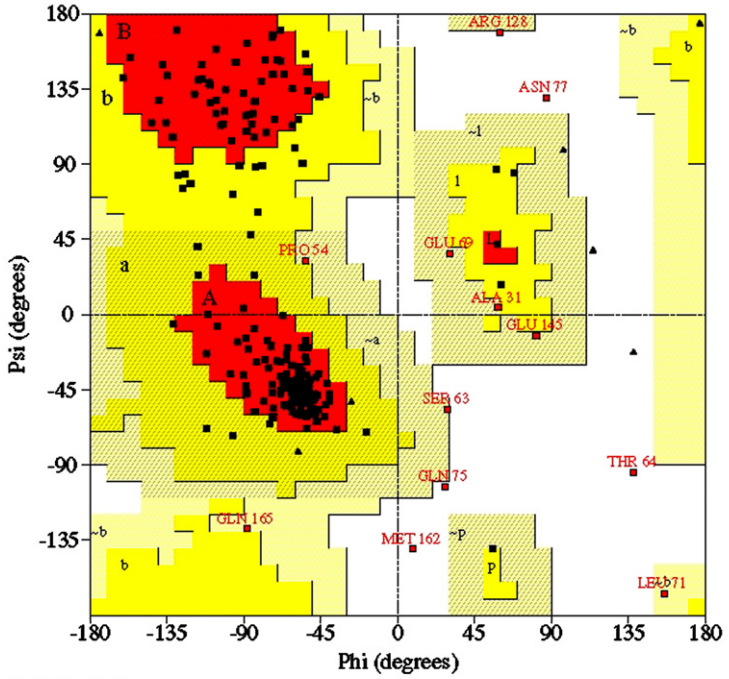
The Ramachandran plot of dCKm (K22–F256) (Fig. 2), produced with PROCHECK, indicated 172 aa (81%) in the most favored region (A, B, L), 30 aa (14%) in the additionally allowed region (a, b, l, p), 6aa (3%) in the generously allowed region (~a, ~b, ~l, ~p), and 5 aa (2%) in the disallowed region. Thus, 98% of all torsion angles are in favorable regions. A good homology model should show >90% of the data points in the favorable region[30] and torsion angle values for various homology models found in the literature ranged from 90% to 96%,[30–33] supporting that dCKm is sufficiently accurate.

The PROFILES-3D identified three imperfect structural regions in dCKm (F68-M73, K117-E120, and L235-D238) indicated in Figure 1 by underlined letters. These were located on the surface of the enzyme about 30 Å away from the active site and did not seem to interfere with the accuracy of the docking studies described below.

3.2. Docking simulations

3.2.1. Analysis of nucleoside binding patterns in the active sites of dCKc and dCKm. The structural features of dCKc have been described in detail elsewhere.[16] Our focus in this paper is the differences between the active sites of dCKc and dCKm, which showed structural variations in the active site aa residues, notably Y86, Q97, D133, R104, R128, and E197 (Fig. 3A and B). These variations created entirely different binding patterns for pyrimidine and purine nucleosides in dCKc and dCKm.

Differences in binding between dC and the active site aa residues of dCKc and dCKm are shown in Figure 3A and B. In dCKc there was no interaction between R104 and N3 (5.1 Å) while D133 was strongly interacting with 4-NH_2 (2.9 Å). Q97 formed two hydrogen bonds with N3 and 4-NH_2, and R128 interacted via hydrogen bonding with 5′-OH (2.9 Å). In contrast, the strong bond between D133 and 4-NH_2 is absent in dCKm (5.0 Å). The resulting lack of stability allows R104 to extend further into the binding pocket of dCKm to interact effectively with N3 and 4-NH_2 impos-



Plot Statistics

Residues in most favorable regions [A, B,L]	172	81%
Residues in additional allowed regions [a,b,l,p]	30	14%
[a] Residues in generously allowed regions [~a,~b,~l,~p]	6	3%
[a] Residues in disallowed regions	5	2%
a :- *Marked with residue number*		
Number of non-glycine and non-proline residues	213	100%
Number of G (shown in triangles) and P residues	16	
Number of end residues	6	
Number of chain end residues not included (K1-K22 & T235-L60)	25	
Total number of residues	260	

Figure 2. Ramachandran plot of dCKm. A good stereochemical quality of dCKm is demonstrated by the presence of 98% aa residues in favored regions.

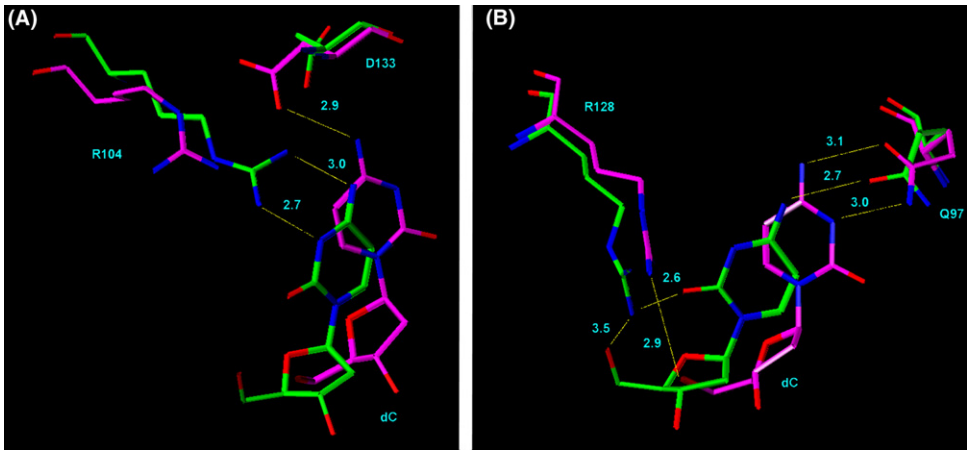


Figure 3. (A) Alignment of active site residues, R104 and D133, of the dCKc/dC [magenta] and dCKm/dC [green] complexes; (B) alignment of active site residues, R128 and Q97, of the dCKc/dC [magenta] and dCKm/dC [green] complexes.

ing *syn* conformation on cytosine, which was further stabilized by hydrogen bonding between the C2 carbonyl bond and R128. This *syn* conformation could be considered as a 'pseudo-purine' conformation, in which hydrogen bonding of N3 with R104 imitates the interaction between N7 and R104 in the case of the binding of dG in dCKm (see Fig. 4).

When dC was docked into dCKc(ww), binding interactions were generated that are similar to those within the dC–dCK crystal structure[16] except for the interactions between R128 and 5′-OH (4.5 Å) as well as R104 and N3 (6.8 Å) [Table 1 and Fig. 5A]. We do not know whether this deviation is due to computational or crystallographic errors. However, docking of dC into dCKc(ww) with another docking program, AutoDock 3.0, produced the same deviated pattern (R128/5′–OH = 4.6 Å and R104/N3 6.8 Å) [Table 1 and Fig. 5A]. These differences in interaction distances between the FlexX docked dC–dCKc(ww) complex and the dC–dCK crystal structure also may have contributed to the relatively large rmsd value of 2.1 Å between both structures (see Section 3.2.2 for more details).

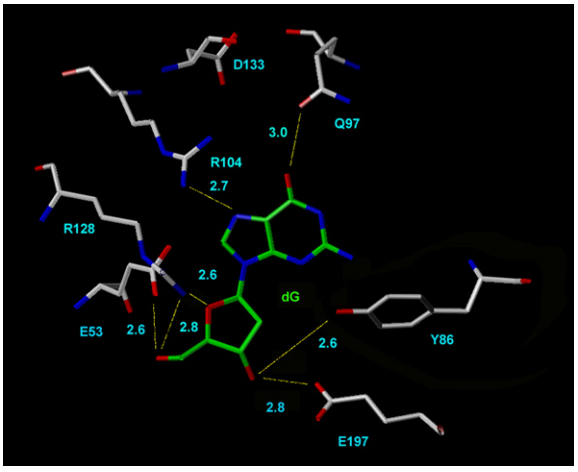


Figure 4. The dG/dCKm complex. The yellow lines represent bonding interactions in Ångstrom.

When dC was docked into dCKc(wow), the distances between R128/5′-OH and R104/N3, respectively, increased further and the interaction distance between E53 and 5′-OH increased from 3.1 Å [dCKc(ww)] to 5.0 Å [dCKc(wow)]. This strongly indicates an important role of active site water molecules in the accuracy of docking simulations with pyrimidine nucleosides in case of dCKc. dC showed *anti*-conformation in dCKc(ww) and dCKc(wow), which was also found in the dC–dCK crystal structure.[16]

The binding pattern of dG with dCKm is shown in Figures 4, 5B, and Table 1. The 3′-OH group showed hydrogen bonding with Y86 (2.6 Å) and E197 (2.8 Å), while 5′-OH interacted with E53 (2.6 Å) and R128 (2.8 Å) via hydrogen bonds. The interactions of the deoxyribose portion of dG with dCKm were similar to those found in the dC–dCKc crystal structure reported by Lavie and co-workers[16] (Table 1). The extended conformation of R104 in dCKm accommodated strong hydrogen bonding (2.7 Å) with N7 while interaction of the amino group of Q97 stabilized the *anti*-conformation of dG via effective hydrogen bonding with the C-6 carbonyl group (3.0 Å). dG assumes *anti*-conformation in dCKm with optimal interaction distances to almost all active site aa residues. These interactions correspond closely with those found for ATP in substrate binding site of dGKc[34] with the exception of E53/E70 with 5′-OH (2.6 vs 6.6 Å). In case of ATP in dGKc the interaction between E70 and 5′-OH seems to be mediated by a water molecule, which is situated halfway between both entities while there is a direct hydrogen bond between E53 and 5′-OH in the case of dG in dCKm.

The interaction distances of dG with active site aa residues of dCKc(wow) clearly indicated docking to the protein in some distance from the active site, which becomes even more pronounced when dG is docked to dCKc(ww) (Table 1). The active site volume of dGKc (ID code: 1jag) [1108 Å] is slightly larger than that of

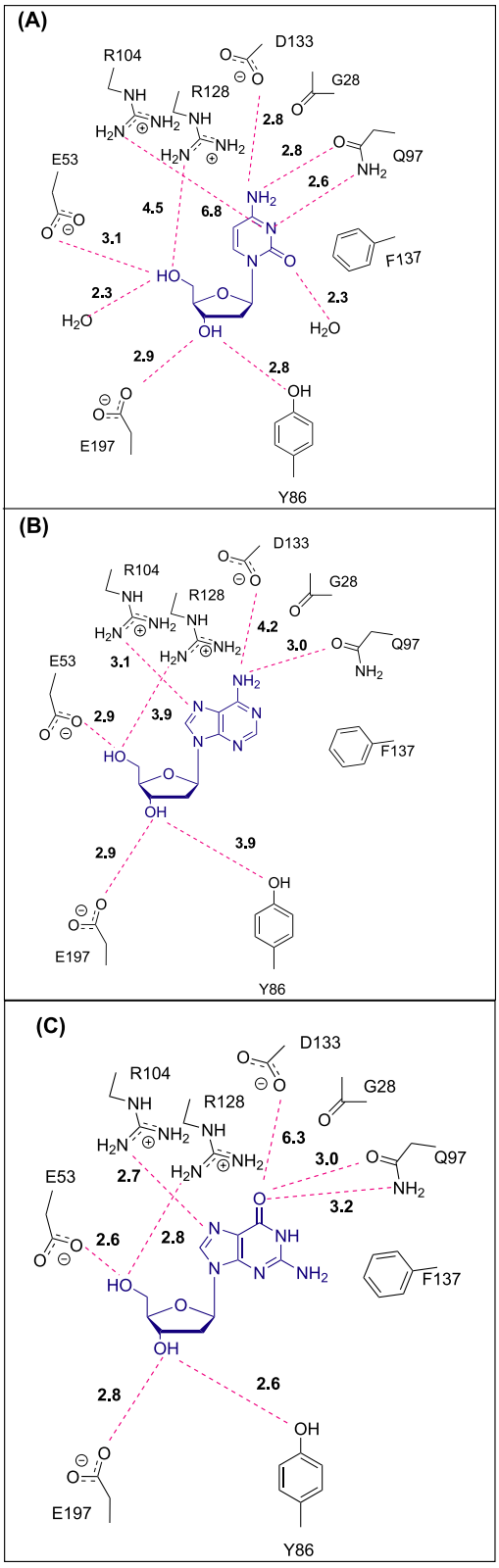


Figure 5. (A) Schematic representation of the docked pose of dC in the dCKc(ww) active site; (B) schematic representation of the docked pose of dA in the dCKm active site; (C) schematic representation of the docked pose of dG in the dCKm active site. The dotted lines represent interaction distance in Ångstrom.

dCKm [1070 Å], which is significantly larger than that of dCKc (ID code:1p60) [858 Å]. The van der Waals volumes of dC[634 Å^3], dA[681 Å^3], and dG[713 Å^3] reflect these differences in active site volumes of dCKc, dCKm, and dGKc, respectively, to some extent. dG has the largest volume of all three nucleosides and the active site of dCKc is significantly smaller than those of dCKm and dGKc. Consequently, the unfavorable binding orientation of dG in dCKc may be attributed to the steric interference.

Docking of dA into dCKm (Table 1, and Fig. 5C) resulted in *anti* conformations with binding interaction comparable to those of dG in dCKm. Overall, the interaction distances between dA in both dCKc(ww) and dCKc(wow) are significantly improved compared with those of dG in the same structures with slightly more favorable values for dCKc(wow).

3.2.2. Comparison of scoring functions and K_d values. For the comparison of scoring functions with experimental data, the nucleosides dC, AraC, Cyd, CdA, dA, and dG were docked into dCKc and dCKm. Experimental data were previously published K_d values of these nucleosides with dCK (Table 2). The scoring functions were FlexX_score and G_score, which are included in the CScore™ module of FlexX.

In case of the purine nucleosides CdA, dA, and dG, we explored overall four different docking approaches to correlate FlexX_scores and G_scores with experimental K_d values: (a) FlexX/G_scores were obtained by docking the purine nucleosides into dCKc with and without water molecules. (b) Rebuilding of the structure of dGK by substituting the active site aa's of dGKc with those of dCK, which produced an active site with an identity score of 54% and rmsd of 6.8 Å compared with the dCKm active site. FlexX/G_scores were then obtained by docking the purine nucleosides into this structure with and without water molecules. (c) FlexX/G_scores were obtained by docking the purine nucleosides into dCKm with water molecules generated by adding the coordinates of the active site water molecules of dGKc. (d) FlexX/G_scores were obtained by docking the purine nucleosides into dCK without water molecules (dCKm).

Only in case of (d), FlexX and G_score scores did correlate with the experimental K_d values of the purine nucleoside as shown in Table 2. Also, in some of models including water molecules, substrate docking outside of the binding pocket was observed. This excludes the possibility of utilizing the models described under a–c for CAMD. Docking of the three pyrimidine nucleosides dC, AraC, and Cyd into dCKc(ww) also produced high correlation of FlexX and G_scores with experimental K_d values (Table 2).

Data from the scoring functions Chemscore, D_score, and PMF_score were not presented because they appeared to be random. Only data from FlexX_score and G_score correlated adequately with the K_d values of various nucleosides probably because they are based to a substantial degree on polar receptor–ligand interactions,[35] which dominate in the protein–substrate com-

plexes analyzed here. In general, the rmsd value of a docked ligand is calculated using the coordinates of the same ligand in the crystal structure as a reference and optimal ligand–protein interactions are defined by rmsd values $\leqslant 2$ Å.[18,36] The rmsd value obtained in this way for dC and AraC in dCKc(ww) were 2.1 Å and 2.8 Å, respectively (Table 2). These relatively high rmsd values could be due to the fact that both reference nucleoside structures were imperfect within dCKc, with for example, missing bonds between N-1 and C′-1, missing C-2 carbonyl, and a missing proton at 4-NH_2. For the calculation of the rmsd values of Cyd, the coordinates of dC in dCKc were used since the corresponding reference ligand was not available. For the same reason, the coordinates of the best-docked pose of dG in dCKm was used as the reference in the case of dCKm for the calculation of all rmsd values. We adopted the criteria that 'relative' rmsd values $\leqslant 3.0$ Å are indicative of appropriately docked ligands whereas values $\geqslant 3.0$ Å indicate no docking. In previous CAMD studies with estrogen receptor ligands, rmsd values obtained in this way still allowed a fairly accurate estimation of the degree of ligand–protein interaction.[37,38] Nevertheless, such relative rmsd values have to be interpreted with caution.

4. Summary and conclusions

A purine nucleoside specific homology model of dCK was created using dGKc as the template. The active site of dCKm was considerably larger than that of dCKc, which, in combination with structural alterations including in particular aa residues Y86, Q97, D133, R104, R128, and E197 resulted in preferential binding of the purine nucleosides dA and dG as substrates compared with the presumably pyrimidine nucleoside specific dCKc, which showed preference for dC.

The obtained results from our docking studies indicated that in case of docking of dC into dCKc, water molecules appear to be important for accurate docking while in particular for docking of dG into dCKm, the absence of water molecules seems to be an important factor for accuracy. Similar scenarios have been observed for HSV1 TK and VZV TK, which belong to the same deoxynucleoside kinase family as dCK.[34,39] Data from crystallographic studies as well as from docking studies with homology models indicated that the presence and absence of active site water molecules is a crucial factor in the binding modes of pyrimidine nucleosides such as dT versus purine nucleoside such as acyclovir and gancyclovir with HSV1 TK and VZV TK.[40,41]

The results of our studies suggest a model with two different conformational states of dCK, as proposed by Bohman and Eriksson,[5,6] which may be caused by specific interaction of substrates with different structural features. A so-called induced fit of a substrate in the active site of an enzyme leading to a specific conformation of that site has been observed in various enzyme systems[34,42,43] including HSV-1 TK.[44] We hypothesize that the difference in pyrimidine and purine specific conformations of dCK could be the result of a similar process, which could encompass the inclusion and exclusion of active site water molecules caused by conformational rearrangements of the active site.

Scoring functions of the pyrimidine nucleosides, dC, AraC, and Cyd with dCKc(ww) and the purine nucleosides, CdA, dA, and dG with dCKm correlated with the corresponding K_d values. This supports the conclusion that dCKc and dCKm may be suitable for the prediction of the inhibitory capacity of novel nucleosides for antiviral and anticancer therapy.

Acknowledgements

This work was supported by the US Department of Energy Grant DE-FG02-90ER60972 and The Swedish Medial Research Council Grant K5104-1085.

References and notes

1. Owens, J. K.; Shewach, D. S.; Ullman, B.; Mitchell, B. S. *Cancer Res.* **1992**, *52*, 2389–2393.
2. Ruiz van Haperen, V. W.; Veerman, G.; Eriksson, S.; Boven, E.; Stegmann, A. P.; Hermsen, M.; Vermorken, J. B.; Pinedo, H. M.; Peters, G. J. *Cancer Res.* **1994**, *54*, 4138–4143.
3. Arner, E. S. J.; Eriksson, S. *Pharmacol. Ther.* **1995**, *67*, 155–186.
4. Kierdaszuk, B.; Bohman, C.; Ullman, B.; Eriksson, S. *Biochem. Pharmacol.* **1992**, *43*, 197–206.
5. Bohman, C.; Eriksson, S. *Biochemistry* **1988**, *27*, 4258–4265.
6. Bohman, C.; Eriksson, S. *Biochemistry* **1990**, *9*, 11–35.
7. Kierdaszuk, B.; Krawiec, K.; Kazimierezuk, Z.; Jacobsson, U.; Johansson, N. G.; Munch-Petersen, B.; Eriksson, S.; Shugar, D. *Adv. Exp. Med. Biol.* **1998**, *431*, 623–627.
8. Maltseva, T.; Usova, E.; Eriksson, S.; Milecki, J.; Foldesi, A.; Chattopadhayaya, J. *Perkin 2* **2000**, 2199–2207.
9. Maltseva, T.; Usova, E.; Eriksson, S.; Milecki, J.; Foldesi, A.; Chattopadhayaya, J. *Nucleosides Nucleotides Nucleic Acids* **2001**, *20*, 1225–1228.
10. Johansson, N. G.; Eriksson, S. *Acta Biochim. Pol.* **1996**, *43*, 143–160.
11. Eriksson, S.; Munch-Petersen, B.; Johansson, K.; Eklund, H. *Cell. Mol. Life Sci.* **2002**, *59*, 1327–1346.
12. Van den Heuvel-Eibrink, M. M.; Wiemer, E. A. C.; Kuipers, M.; Pieters, R.; Sonneveld, P. *Haematol. Blood Transfus.* **2003**, *41*, 232–238.
13. Mani, R. S.; Usova, E. V.; Eriksson, S.; Cass, C. E. *Nucleosides Nucleotides Nucleic Acids* **2003**, *22*, 175–192.
14. Mitchell, B. S.; Song, J. J.; Johnson, E. E., II; Chen, E.; Dayton, J. S. *Adv. Enzyme Regul.* **1993**, *33*, 61–68.
15. Shafiee, M.; Gosselin, G.; Imbach, J.-L.; Divita, G.; Eriksson, S.; Maury, G. *Eur. J. Med. Chem.* **1999**, *34*, 423–431.
16. Sabini, E.; Ort, S.; Monnerjahn, C.; Konrad, M.; Lavie, A. *Nat. Struct. Biol.* **2003**, *10*, 513–519.
17. Mani, R. S.; Usora, E. V.; Eriksson, S.; Cass, C. E. *Nucleosides Nucleotides Nucleic Acids* **2004**, *23*, 1343.
18. Greer, J. *Proteins* **1990**, *7*, 317–334.
19. Bowie, J. U.; Luthy, R.; Eisenberg, D. *Science* **1991**, *253*, 164–170.
20. Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Cryst.* **1993**, *26*, 283–291.

21. Rarey, M.; Kramer, B.; Lengauer, T.; Klebe, G. *J. Mol. Biol.* **1996**, *261*, 470–489.
22. Gohlke, H.; Hendlich, M.; Klebe, G. *J. Mol. Biol.* **2000**, *295*, 337–356.
23. Eldridge, M. D.; Murray, C. W.; Auton, T. R.; Paolini, G. V.; Mee, R. P. *Comput. Aided Mol. Des.* **1997**, *11*, 425.
24. Ewing, T. J. A.; Kuntz, I. D. *J. Comput. Aided Mol. Des.* **2001**, *15*, 1175.
25. Jones, G.; Willett, P.; Glen, R. C.; Leach, A. R.; Taylor, R. *J. Mol. Biol.* **1997**, *267*, 727–748.
26. Muegge, I.; Martin, Y. C. *J. Med. Chem.* **1999**, *42*, 791.
27. Bissantz, C.; Folkers, G.; Rognan, D. *J. Med. Chem.* **2000**, *43*, 4759–4767.
28. Bohm, H. J. *J. Comput. Aided Mol. Des.* **1994**, *8*, 243–256.
29. Davis, A. M.; Teague, S. J.; Kleywegt, G. J. *Angew. Chem., Int. Ed.* **2003**, *42*, 2718–2736.
30. Ettrich, R.; Melichercik, M.; Teisinger, J.; Ettrichova, O.; Krumscheid, R.; Hofbauerova, K.; Kvasnicka, P.; Schoner, W.; Amler, E. *J. Mol. Model.* **2001**, *7*, 184–192.
31. Hesson, D. P.; Sturgess, M. A. *Bioorg. Med. Chem. Lett.* **1997**, *7*, 1437–1442.
32. Sticht, H.; Gallert, K.-C.; Krauss, G.; Rosch, P. *J. Biomol. Struct. Dyn.* **1997**, *14*, 667–675.
33. Fengler, A.; Brandt, W. *J. Mol. Model.* **1999**, *5*, 177–188.
34. Johansson, K.; Ramaswamy, S.; Ljungcrantz, C.; Knecht, W.; Piskur, J.; Munch-Petersen, B.; Eriksson, S.; Eklund, H. *Nat. Struct. Biol.* **2001**, *8*, 616–620.
35. Clark, R. D.; Strizhev, A.; Leonard, J. M.; Blake, J. F.; Matthew, J. B. *J. Mol. Graph. Model.* **2002**, *20*, 281–295.
36. Good, A. C.; Cheney, D. L. *J. Mol. Graph. Model.* **2003**, *22*, 23–30.
37. Johnsamuel, J.; Byun, Y.; Jones, T. P.; Endo, Y.; Tjarks, W. *J. Organomet. Chem.* **2003**, *680*, 223–231.
38. Johnsamuel, J.; Byun, Y.; Jones, T. P.; Endo, Y.; Tjarks, W. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 3213–3216.
39. Pospisil, P.; Pilger, B. D.; Marveggio, S.; Schelling, P.; Wurth, C.; Scapozza, L.; Folkers, G.; Pongracic, M.; Mintas, M.; Malic, S. R. *Helv. Chim. Acta* **2002**, *85*, 3237–3249.
40. Pospisil, P.; Kuoni, T.; Scapozza, L.; Folkers, G. *J. Receptor Signal Transd.* **2002**, *22*, 141–154.
41. Spadola, L.; Novellino, E.; Folkers, G.; Scapozza, L. *Eur. J. Med. Chem.* **2003**, *38*, 413–419.
42. O'Brien, J. R.; Schuller, D. J.; Yang, V. S.; Dillard, B. D.; Lanzilotta, W. N. *Biochemistry* **2003**, *42*, 5547–5554.
43. Ragona, L.; Fogolari, F.; Catalano, M.; Ugolini, R.; Zetta, L.; Molinari, H. *J. Biol. Chem.* **2003**, *278*, 38840–38846.
44. Bennett, M. S.; Wien, F.; Champness, J. N.; Batuwangala, T.; Rutherford, T.; Summers, W. C.; Sun, H.; Wright, G.; Sanderson, M. R. *FEBS Lett.* **1999**, *443*, 121–125.