Salvage Enzymes in Nucleotide Biosynthesis

Structural Studies on Three Bacterial Thymidine Kinases and Human Uridine-Cytidine Kinase 1

Urszula Kosinska
Faculty of Natural Resources and Agricultural Sciences
Department of Molecular Biology
Uppsala

Doctoral thesis
Swedish University of Agricultural Sciences
Uppsala 2007

Acta Universitatis Agriculturae Sueciae
2007: 49

ISSN 1652-6880
ISBN 978-91-576-7348-0

Tryck: SLU Service/Repro, Uppsala 2007

Abstract

Kosinska, U., 2007. Salvage enzymes in nucleotide biosynthesis: Structural studies on three bacterial thymidine kinases and human uridine-cytidine kinase 1. Doctoral dissertation
ISSN 1652-6880, ISBN 978-91-576-7348-0

Balanced pools of deoxyribonucleoside triphosphates (dNTPs), the building blocks of DNA, and ribonucleoside triphosphates (NTPs), the precursors of RNA, are crucial for a controlled cell proliferation. The dNTPs and NTPs are synthesized *de novo* via energy-consuming reactions involving low-weight molecules, and through a salvage pathway by recycling (deoxy)ribonucleosides originating from food and degraded DNA and RNA. The enzymes described in this thesis catalyze the first reaction in the salvage biosynthesis of dNTPs and NTPs.

The crystal structures of three bacterial thymidine kinases (TKs) are described and the enzymes are investigated as potential targets for antibacterial therapies. TK is a deoxyribonucleoside kinase (dNK) with specificity for thymidine. In addition to the natural substrates, TK can also phosphorylate a number of nucleoside analogs used in antiviral and anticancer therapies. This thesis presents the structures of TKs from three pathogenic microorganisms: *Ureaplasma urealyticum* (*parvum*), *Bacillus anthracis* and *Bacillus cereus*, and compares them to the human thymidine kinase 1 (hTK1). The bacterial TKs and the hTK1 are structurally very similar and have a highly conserved active site architecture, which may complicate structure-based drug design. However, the different complex structures presented in this work provide information regarding the conformational changes of TK1-like enzymes during the time of reaction.

The structure of human uridine-cytidine kinase 1 (UCK1) is also presented. Humans possess two uridine-cytidine kinases, UCK1 and UCK2. The expression pattern of these enzymes is tissue dependent, and despite high sequence as well as structural similarities they possess somewhat diverse substrate specificity. In addition to the natural substrates, uridine and cytidine, UCKs are able to phosphorylate a number of nucleoside analogs. The monomeric structure of UCK comprises four domains: a CORE domain, an NMP-binding domain, a LID domain and a β-hairpin domain, which upon substrate binding undergo dramatic conformational changes. In the structure described in this thesis the enzyme has been trapped in an intermediate conformation between a fully opened and fully closed form, which may represent a sequential mode of substrate binding.

Keywords: thymidine kinase, structural zinc, deoxyribonucleoside kinase, ribonucleoside kinase, uridine-cytidine kinase, X-ray crystallography, transferase, nucleotide biosynthesis, salvage pathway

Author's address: Urszula Kosinska, Department of Molecular Biology, SLU, S-751 24 Uppsala, Sweden. urszula@xray.bmc.uu.se

Contents

Appendix

Papers I-IV

This thesis is based on the following papers, which will be referred to by their Roman numerals:

I. Welin*, M., **Kosinska*, U.**, Mikkelsen, N-E., Carnrot, C., Zhu, C., Wang, L., Eriksson, S., Eklund, H., (2004) Structures of thymidine kinase 1 of human and mycoplasmic origin. *Proc Natl Acad Sci USA* 101, 17970-17975

II. **Kosinska*, U.**, Carnrot*, C., Eriksson, S., Wang, L., Eklund, H., (2005) Structure of the substrate complex of thymidine kinase from *Ureaplasma urealyticum* and investigations of possible drug targets for the enzyme. *FEBS J* 272, 6365-6372

III. **Kosinska, U.**, Carnrot, C., Sandrini, M. P. B., Clausen, A. R., Wang, L., Piskur, J., Eriksson, S., Eklund, H., (2007) Structural studies of thymidine kinases from *Bacillus anthracis* and *Bacillus cereus* provide insights into quaternary structure and conformational changes upon substrate binding. *FEBS J* 274, 727-737

IV. **Kosinska, U.**, Walldén, K., Flodin, S., Hammarström, M., Johansson, I., Moche, M., Nyman, T., Stenmark, P., Eklund, H., Karlsson, A., Nordlund, P., (2007) Structural study of human uridine-cytidine kinase 1. *Manuscript*

* First authorship shared

Abbreviations

Enzymes

dAK	deoxyadenosine kinase
dCK	deoxycytidine kinase
dGK	deoxyguanosine kinase
dNK	deoxyribonucleoside kinase
HSV-1 TK	Herpes Simplex virus type 1 thymidine kinase
NDPK	nucleoside diphosphate kinase
NK	ribonucleoside kinase
NMPK	nucleoside triphosphate kinase
RNR	ribonucleotide reductase
TK	thymidine kinase
TS	thymidylate synthase
UCK	uridine-cytidine kinase

Nucleosides and nucleotides

dN, dNMP, dNDP, dNTP	deoxyribonucleoside, mono, di, and triphosphate
N, NMP, NDP, NTP	ribonucleoside, mono, di, and triphosphate
Ado, dA	adenosine, deoxyadenosine
Cyd, dC	cytidine, deoxycytidine
Guo, dG	guanosine, deoxyguanosine
dT	deoxythymidine = thymidine
Urd, dU	uridine, deoxyuridine
-MP, -DP, -TP	-mono, -di, -triphosphate of nucleosides

Other

3D	three dimensional
MPD	2-methyl-2,4-pentanediol
SAD	single wavelength anomalous diffraction
Se-Met	selenomethionine

Foreword

Proteins play a central role in many biological processes and deserve their name, which comes from the Greek *prota* or *proteios* meaning *of primary importance* or *first place*. Functions such as structural support, storage, transport of chemical compounds, regulation of metabolic processes, signaling from one cell to another, mobility and defense against foreign substances, are all carried out by proteins. Moreover, proteins catalyze most of the chemical reactions that take place in cells. Such proteins are termed enzymes.

The primary sequence, which subsequently determines the three dimensional structure of the protein, is determined by the genetic code in the form of DNA. Furthermore, the function of a protein is strongly dependent on its structure. Thus to establish a complete picture of biological processes we need to combine molecular biology with biochemistry and structural biology. Three methods are mainly used for protein structure determinations; electron microscopy, nuclear magnetic resonance (NMR), and X-ray crystallography. Electron microscopy can be applied in studies of large complexes of biological macromolecules, but the resolution limit is still too low for detailed structural studies. The limiting factor of NMR is the size of the biomolecule of interest, while the major obstacle of crystallography is obtaining well-diffracting crystals. Since no size limit exists in crystallography, this method has an application for structure determination of a variety of large assemblies of biological macromolecules. An example of such an assembly that recently regained large attention is the protein-DNA-RNA complex of the RNA polymerase for which Roger D. Kornberg was awarded the Nobel Prize in Chemistry in 2006. The first applications of X-ray crystallography on biological macromolecules reach 50 years back in time. At that time John C. Kendrew and Max F. Perutz were the foremost leading explorers of protein crystallography. In 1962 they were awarded the Nobel Prize for their determination of the myoglobin and hemoglobin structures. Yet another highly dedicated pioneer who has left deep footprints on protein crystallography is Dorothy C. Hodgkin. Her work resulted in the structure determination of insulin, but also the biologically significant substances like penicillin and vitamin B12, for which she was awarded the Nobel Prize in 1964.

Today there are more than 35 000 structures solved by X-ray crystallography deposited in the Protein Data Bank, reflecting the immense advance of crystallography. In this thesis, the structure determination and biological relevance of six out of the 35 000 structures are described. The work presented here focuses on structural studies using X-ray crystallography on enzymes involved in the synthesis of deoxyribonucleoside triphosphates (dNTPs), the precursors of DNA, and nucleoside triphosphates (NTPs), the precursors of RNA. In particular three bacterial thymidine kinases originating from *Ureaplasma urealyticum*, *Bacillus cereus* and *Bacillus anthracis* were investigated. By combining molecular biology with structural studies the aim was to investigate these enzymes in the light of future design of anti bacterial agents. Additionally, human uridine-cytidine kinase

1 was investigated in collaboration with Structural Genomics Consortium at Karolinska Institutet in Stockholm.

1. Synthesis of (deoxy)ribonucleoside triphosphates

DNA replication and repair strongly depends on a balanced pool of the four deoxyribonucleoside triphosphates (dNTPs). The dNTP pools fluctuate during the cell cycle, increase for the period of DNA synthesis (the S-phase), and reach their lowest concentrations in the post-mitotic G_0 phase. A general estimate is that pyrimidine dNTP pools are often larger than purine pools (Traut, 1994). The dNTP pools are in a way self regulated such that dNTPs control the activities of dNTP synthesizing enzymes via routes of feedback inhibition or allosteric regulation. Additionally, the dNTP pools may be kept balanced by regulating the amounts of enzymes responsible for their synthesis through cell cycle dependent expression and degradation.

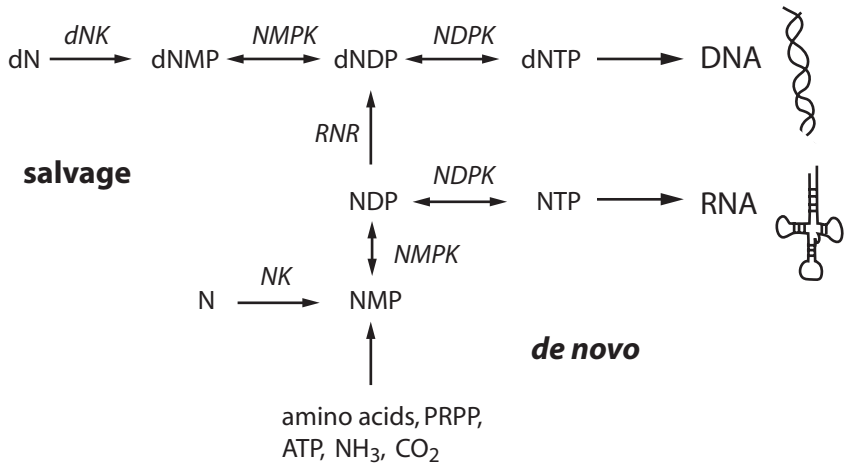


Fig. 1. Overview of nucleotide metabolism. The salvage pathway recycles deoxyribonucleosides (dNs) and ribonucleosides (Ns) from nutrients, and degraded DNA and RNA, whereas low-molecular weight compounds make the starting materials in the *de novo* synthesis.

Most organisms have two pathways for dNTP synthesis, *de novo* and salvage, but some, for example *Ureaplasma urealyticum,* have to solely rely on salvaging enzymes to maintain balanced dNTP pools. The *de novo* synthesis uses low-molecular-weight precursors to build up nucleotides whereas the salvage pathway recycles deoxyribonucleosides (dNs) from degraded DNA. Figure 1 gives a schematic representation of dNTP synthesis and shows how the two pathways meet at the junction when ribonucleoside diphosphates (NDPs) are reduced to deoxyribonucleoside diphosphates (dNDPs) by ribonucleotide reductase (RNR). Before NDPs reach RNR for the irreversible reduction, they are formed from small molecules such as glutamine, asparagine, aspartate, glycine, tetrahydrofolate derivatives, phosphoribosylpyrophosphate (PRPP), ammonia and CO_2. Purine

nucleotides (adenosine and guanosine) are synthesized at the nucleotide level such that the purine ring is assembled on the ribose, while pyrimidine nucleotide (uridine and cytidine) synthesis takes place on the free base level. The final products of the *de novo* synthesis are nucleoside monophosphates (NMPs). As for dNs, there is also a salvage route for the synthesis of NMPs, in which nucleosides from food and degraded RNA are phosphorylated to the corresponding nucleoside monophosphates (NMPs).

1.1 Enzymes in the *de novo* pathway

1.1.1 Ribonucleotide reductase (RNR)

Extensive studies of the mechanism, allosteric regulation, physiology and evolution of RNR have been performed and several reviews are available with the latest by Nordlund and Reichard (Nordlund & Reichard, 2006). RNR plays an important role in converting RNA precursors into DNA precursors by catalyzing a radical dependent reduction of the 2'-hydroxyl of ribonucleoside diphosphates. On the basis of their mode to generate the radical, the RNRs are divided into three classes. Class I RNRs contain a dinuclear iron center which generates and stabilizes a tyrosyl radical, class II enzymes use adenosylcobalamin to produce the radical, while class III enzymes harbor a [4Fe-4S] center, which together with S-adenosyl methionine and reduced flavodoxin generates a glycyl radical. All three classes can be found in microorganisms, where class III is restricted to anaerobic microorganism. Eukaryotes possess only class I RNRs.

Human RNR, a class I RNR, is a tetramer composed of two dimers R1 and R2. The R2 dimer is the smaller of the two dimers which generates and stores the radical. Upon substrate binding to the reaction site situated in the R1 dimer, the radical is transferred from R2 to a conserved cysteine residue in the reaction site in R1. The radical abstracts a hydrogen atom from the 3'-carbon of the ribose and produces a substrate radical. Electron transfer then generates a 2'-deoxynucleotide. Since RNR reduces all four ribonucleotides (ADP, GDP, CDP and UDP) a rigorous regulation mechanism is required to control the enzyme in order to prevent dNTP pool imbalance. Thus, in addition to the reaction site, the large subunit has an overall activation site and a specificity site. The overall activation site works like a general on/off switch such that an ATP molecule bound to that site activates the enzyme while a dATP molecule turns it off. The specificity site determines which substrates will be processed by RNR. The type of ribonucleotide to be reduced depends on whether ATP, dATP, dTTP or dGTP is bound at the specificity site. Besides the allosteric regulation of RNR, the enzyme activity is also regulated by cell cycle dependent expression and degradation of the small subunit R2 (Chabes & Thelander, 2000; Chabes *et al.*, 2003). In resting cells a p53-induced R2, p53R2, substitutes for the normal R2 and forms an active complex with R1 to supply the cell with dNTPs required for DNA-damage repair (Guittet *et al.*, 2001).

1.1.2 De novo *synthesis of dTTP involves dCMP deaminase, thymidylate synthase and dUTPase*

The synthesis of three (dATP, dGTP and dCTP) of the four DNA precursors is rather straight forward involving reduction of the ribonucleoside diphosphate by RNR and subsequent phosphorylation to the final deoxyribonucleoside triphosphate. The *de novo* synthesis of dTTP, however, requires additional steps where deoxycytidine and deoxyuridine nucleotides are remodeled to form dTTP. This conversion involves thymidylate synthase (TS), dCMP deaminase and UTPase (Reichard, 1988). TS catalyzes the conversion of dUMP to dTMP, which requires oxidation of 5,10-methylenetetrahydrofolate to dihydrofolate, and thus TS is dependent on an active dihydrofolate reductase. dCMP deaminase is an allosterically regulated enzyme activated by dCTP and inhibited by dTTP. It is the major supplier of dUMP and controls the relative amounts of dCTP and dTTP. A further route to provide TS with dUMP is provided by the reaction of dUTPase which cleaves dUTP to dUMP and a pyrophosphate. Moreover, the action of dUTPase effectively prevents dUTP accumulation which could lead to incorporation of uracil into DNA. In prokaryotic cells formation of dUMP takes place at the dCTP level where a dCTP deaminase produces dUTP which is further hydrolyzed to dUMP by dUTPase.

1.2 Enzymes in the salvage pathway

An alternative route to the *de novo* biosynthesis of (d)NTPs is the less energy demanding salvage pathway, in which (d)Ns are recycled to form their corresponding triphosphates. During the salvage pathway, extracellular (d)Ns originating from degraded cells or nutrients are transported into the cell via non-specific nucleoside carrier proteins (Plagemann, Wohlhueter & Woffendin, 1988) and undergo three successive phosphorylation steps. In the first step deoxyribonucleoside kinases (dNKs) and ribonucleoside kinases (NKs) phosphorylate deoxyribonucleoside or ribonucleosides, respectively. Addition of negative charges in form of phosphates traps the phosphorylated nucleosides inside the cell, and therefore this step is often described as the rate limiting step. This provides substrates for subsequent phosphorylation steps catalyzed by nucleoside monophosphate kinases (NMPKs) followed by nucleoside diphosphate kinases (NDPKs).

1.2.1 Deoxyribonucleoside kinases (dNKs)

1.2.1.1 Human dNKs

There are four dNKs in human cells: deoxyguanosine kinase (dGK), deoxycytidine kinase (dCK), thymidine kinase 1 (hTK1), and thymidine kinase 2 (TK2) (Arnér & Eriksson, 1995; Eriksson *et al.*, 2002). Two of these, dGK and TK2, are found in mitochondria, while the other two, dCK and hTK1, are cytosolic. These enzymes have partially overlapping substrate specificities, but a general division into purine and pyrimidine specific dNKs may be made. The two thymidine kinases, hTK1 and TK2 are strict pyrimidine kinases. The cytosolic

hTK1 only phosphorylates thymidine (dT) and deoxyuridine (dU), whereas the mitochondrial TK2 in addition to dT and dU can also phosphorylate deoxycytidine (dC). Deoxycytidine kinase may also be considered as a pyrimidine kinase although in addition to dC it is also active with deoxyadenosine (dA) and deoxyguanosine (dG). Finally, dGK is a purine kinase, but in addition to dA and dG can also phosphorylate dC. Moreover, a range of clinically important nucleoside analogs are phosphorylated, and thereby activated, by dNKs.

Several nucleoside triphosphates can act as phosphate donors, with ATP being accepted by all dNKs. The activities of dNKs are controlled by feedback regulation. The triphosphate of the most preferred ribonucleoside acts as feedback inhibitor for the corresponding dNK, such that hTK1 and TK2 are inhibited by thymidine triphosphate (dTTP), dCK is inhibited by deoxycytidine triphosphate (dCTP) and dGK by deoxyguanosine triphosphate (dGTP) and deoxyadenosine triphosphate (dATP) (Arnér & Eriksson, 1995; Eriksson*, et al.*, 2002; Ives, Morse & Potter, 1963).

Based on their amino acid sequences dNKs form two distinct groups. One of the groups, the dNK-group, consists of sequences of human dCK, dGK, TK2 and their homologs. The second group only includes sequences similar to the cytosolic hTK1, thus the homologs of hTK1 described in this thesis will be referred to as the TK1-like enzymes.

1.2.1.2 dNKs in other organisms

The set of deoxyribonucleoside kinases as well as their preferences for phosphate acceptors differs among organisms. *Drosophila melanogaster* (Munch-Petersen, Piskur & Søndergaard, 1998), *Bombyx mori* (Knecht *et al.*, 2002a) and *Anopheles gambiae* (Knecht *et al.*, 2003) possess one multisubstrate dNK, with high activities with all four dNs. On the other hand, in *E. coli* only TK activity has been detected and neither dCK, dGK nor deoxyadenosine kinases (dAK) were found (Karlström, 1970). This seems to be a common feature for Gram-negative bacteria where only TK1-like sequences have been identified (Sandrini *et al.*, 2006). Most Gram-positive bacteria, in addition to TK, possess one or two dNKs with specificities for the remaining deoxyribonucleosides. Extensive studies have been performed on dNKs from *Bacillus subtilis* and *Lactobacillus acidophilus*. In addition to TK, *B. subtilis* has two homodimeric enzymes, a strict dGK which only accepts dG as substrate and a deoxyadenosine/deoxycytidine kinase (dAK/dCK) which is equally active with dA and dC (Andersen & Neuhard, 2001; Møllgaard, 1980). A rather interesting set of dNKs is present in *L. acidophilus* strain R26 which seem to lack the RNR, and thus relies on the salvage pathway for dNTP synthesis (Ives & Ikeda, 1998). Besides a TK, *L. acidophilus* has two heterodimeric dNKs, dGK/dAK and dCK/dAK, that assure the salvage of the remaining three deoxyribonucleosides dA, dC and dG. Each subunit of the heterodimeric dNKs is specific for the individual substrate and is feedback inhibited by the triphosphate of its preferred deoxyribonucleoside. In mycoplasmas, offspring of Gram-positive bacteria, the set of dNKs have been reported to contain two enzymes, TK (Carnrot *et al.*, 2003) with specificity for dT

and dU, and dAK showing specificity for dA, dC and dG, with the highest catalytic rate for dA (Wang *et al.*, 2001).

Yet another dNK that should be mentioned is the thymidine kinase from Herpes Simplex Virus type 1 (HSV1-TK). HSV1-TK phosphorylates dT and dU, but is also active with purine deoxyribonucleosides. Interestingly, in contrast to the dNKs described above, HSV1-TK catalyzes two subsequent phosphorylation steps, such that a dN is phosphorylated to a deoxyribonucleoside monophosphate (dNMP) and the dNMP is phosphorylated to deoxyribonucleoside diphosphate (dTDP) (Chen *et al.*, 1979; Chen, Walker & Prusoff, 1979). Moreover, HSV1-TK is a specific activator of nucleoside analogs, of which acyclovir and gancyclovir are commonly used as antiherpetic drugs (Gentry, 1992). HSV-1 TK is extensively studied for its potential to be applied in suicide gene therapy.

1.2.1.3 dNK structures

Intensive structural work has contributed to broadening our understanding of the enzymatic activities of some of the dNKs. For a very long time HSV1-TK was the only dNK with known structure (Brown *et al.*, 1995; Champness *et al.*, 1998; Wild *et al.*, 1995; Wild *et al.*, 1997). The first eukaryotic dNKs to be structurally determined were the multisubstrate *D. melanogaster*-dNK and human dGK, shortly after followed by human dCK (Johansson *et al.*, 2001; Sabini *et al.*, 2003). At the structural, but not at the sequence level, HSV1-TK is related to the enzymes of the dNK-group sharing the overall fold and quaternary structure. However, there are important differences in and around the active sites which explain the substrate specificities harbored by these enzymes (Eriksson*, et al.*, 2002; Knecht *et al.*, 2002b; Sabini*, et al.*, 2003). The first TK1-like structures, the mycoplasmic TK (*Uu*-TK) and human TK1 which are described in Paper I, revealed that TK1-like enzymes are structurally unrelated to other dNKs. In following sections the structures of *Uu*-TK, *Bacillus anthracis* and *Bacillus cereus* TKs (*Ba*-TK and *Bc*-TK) are described in detail and compared to hTK1 as well as other TK1-like structures that are now available *i.e.* TK from *Clostridium acetobutylicum* (*Ca*-TK) (Kuzin, 2004), Vaccinia virus TK (Vv-TK) (El Omari *et al.*, 2006) and an additional, independently determined hTK1 (Birringer *et al.*, 2005).

1.2.2 Ribonucleoside kinases (NKs)

Just as there is a salvage pathway for deoxyribonucleosides, there are several ribonucleoside kinases that phosphorylate ribonucleosides and thereby provide an alternative way to the *de novo* synthesis of nucleotides and deoxyribonucleotides. Only adenosine kinase (AK) and uridine-cytidine kinase (UCK) are known to exist in human cells (Ropp & Traut, 1996; Singh *et al.*, 1996; Spychala *et al.*, 1996; Van Rompay *et al.*, 2001). Adenosine kinase is specific for adenosine, while uridine-cytidine kinase, whose structure is described in this thesis, phosphorylates uridine and cytidine. No human guanosine kinase has been identified, but a guanosine kinase or guanosine-inosine kinase is present in *E. coli, Trichomonas vaginalis* and in potato (Harlow, Nygaard & Hove-Jensen, 1995; Katahira & Ashihara, 2006; Miller & Miller, 1991).

1.2.3 Reaction mechanism of dNKs and NKs

The general mechanism for the phosphorylation reaction catalyzed by dNKs and NKs is presented in figure 2. A glutamic or aspartic acid acts as general base and initiates the reaction by deprotonation of the 5'-hydroxyl group of the ribose or deoxyribose acceptor. The 5'-oxygen is activated to perform a nucleophilic attack on the γ-phosphate of the ATP molecule so that the phosphate may be transferred to create a (d)NMP and ADP. The transition state may be balanced by a Mg-ion and arginines and lysines that commonly occur in the reaction sites of dNKs and NKs (Johansson*, et al.*, 2001; Suzuki *et al.*, 2004; Wild*, et al.*, 1997).

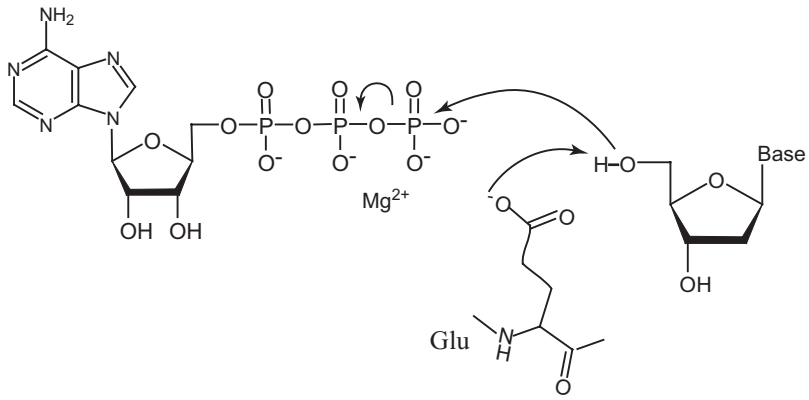


Fig.2. General reaction mechanism for phosphate transfer as catalyzed by dNKs. A similar mechanism takes place in NKs with the exception that the phosphate acceptor is a ribonucleoside instead of a deoxyribonucleoside and the catalytic base is an aspartate. Because of its high abundance ATP is considered as the biological phosphate donor, but other nucleoside triphosphates are also accepted as phosphate donors. The figure was reproduced from Eriksson *et al*. 2002.

1.2.4 Nucleoside monophosphate kinases (NMPKs)

The second phosphorylation step is catalyzed by nucleoside monophosphate kinases (NMPKs) present in most organisms. The requirement for the sugar moiety to be a deoxyribose is not as strict as in dNKs, and thus NMPKs are able to phosphorylate NMPs as well as dNMPs. On the other hand, as with dNKs, NMPKs are base specific. Consequently, human cells possess a thymidylate kinase (dTMPK), uridylate-cytidylate kinase (UMP/CMPK), several isoforms of adenylate kinase (AMPK) and several guanylate kinases (GMPKs) (Van Rompay, Johansson & Karlsson, 2000). As with hTK1, dTMPK has strict phosphate acceptor requirements and can only phosphorylate thymidine monophosphate (dTMP) and deoxyuridine monophosphate (dUMP). Moreover, dTMPK is cell

cycle regulated, as is the case with hTK1 (Huang *et al.*, 1994). Low or absent dTMPK activity has been detected in nonproliferating tissues while fast replicating tumor cells show elevated activities. Structurally the NMPKs are classified in the nucleoside monophosphate kinase family (Yan & Tsai, 1999) with a general monomeric structure comprising CORE, LID and NMP-binding domains. These domains are also present in human uridine-cytidine kinase (UCK) thus they will be described in more detail in the section presenting the UCK1 structure. The bacterial set of NMPKs differs from the human NMPKs by containing two separate enzymes for phosphorylation of (d)UMP and (d)CMP. Especially, bacterial UMPKs are interesting deviants from the nucleoside monophosphate kinase family, acting as hexamers and being structurally more similar to the carbamate kinase-like superfamily (Briozzo *et al.*, 2005).

1.2.5 Nucleoside diphosphate kinases (NDPKs)

The third and last phosphorylation of (deoxy)ribonucleoside diphosphates to corresponding triphosphates is catalyzed by nucleoside diphosphate kinases (NDPKs). NDPKs are non-specific and readily phosphorylate ribose and deoxyribose purine and pyrimidine nucleotides. In contrast to (d)NKs and NMPKs which form ternary complexes of enzyme, phosphate donor and phosphate acceptor where the phosphoryl group is transferred from the donor directly to the acceptor, NDPKs follow a ping-pong mechanism where the phosphoryl transfer occurs via a phosphoenzyme intermediate (Gilles *et al.*, 1991). A crystal structure of NDPK in complex with transition state mimicking Al_3 describes the catalytic mechanism of phosphate transfer from the ATP onto an active site histidine (Xu *et al.*, 1997). As reviewed by Lacombe and colleagues, humans possess eight NDPK encoding genes with tissue dependent expression and subcellular localization of the protein. Besides the classical NDPK activity, NDPKs have been assigned functions associated with cell proliferation, development, differentiation and metastasis suppression (Kimura *et al.*, 2000; Lacombe *et al.*, 2000). Eukaryotic NDPKs are homohexamers while some bacterial NDPKs are homotetramers. Nevertheless they are built up from similar dimers. The subunits of NDPKs from different organisms have highly similar 3D structures with the core of the ferredoxin fold (Janin *et al.*, 2000). However, formation of heterohexamers has been observed in human erythrocytes where the two NDPK polypeptides, A and B, form several isoenzymes by random association of these isoforms (Gilles*, et al.*, 1991).

1.3 Medical applications of the salvage enzymes

In addition to playing an important function in synthesis of DNA and RNA precursor, dNKs and NKs are also activators of a number of nucleoside analogs used in cancer therapies and treatment of viral infections. Nucleoside analogs are derivatives of natural ribonucleosides and deoxyribonucleosides with modifications on the base and sugar moieties. They are harmless prodrugs which have to be phosphorylated in order to display their pharmacological effect. The first phosphorylation step is catalyzed by nucleoside and deoxyribonucleoside

kinase. This is an important step in the activation process of the analogs since addition of the negative phosphate to the analogs traps them inside the cell. Phosphorylated nucleoside analogs act as antimetabolites and compete with natural nucleotides and deoxynucleotides. They interact with and inhibit a range of enzymes involved in the synthesis of nucleic acids. Ultimately, the nucleoside analog triphosphates are incorporated into the DNA and RNA where they cause strand termination, which leads to apoptosis.

1.3.1 Nucleoside analogs

The oldest known nucleoside analogs are the fluorinated uracil and deoxyuracil analogs, 5-flourouracil (FU) and 5-fluorodeoxyuridine (FdU) (Heidelberger *et al.*, 1957). These chemical compounds are efficient inhibitors of DNA synthesis and are still common anticancer drugs. Combination of biochemistry and structural biology helped to elucidate the mode of action for FU and FdU as tumor killers and it is now known that these compounds act, via their metabolite 5-fluorodeoxyuridine monophosphate (FdUMP), as irreversible inhibitors of thymidylate synthase (Ichikawa, 2006). Better understanding of the basis of action of nucleoside analogs has led to a great expansion of cytotoxic nucleoside analogs. Some of the currently used analogs in chemotherapies against cancer are fludarabine, cladirabine and gemcitabine. All three are initially phosphorylated by dCK while the above described FdU is activated by hTK1 phosphorylation (Galmarini, Mackey & Dumontet, 2002).

Nucleoside analogs are also used in chemotherapy of virus infections (De Clercq, 2001). Several anti-HIV drugs are nucleoside analogs which after activation by human dNKs, NMPKs and NDPKs inhibit the reverse transcriptase reaction of HIV which results in viral DNA chain termination. One of the first nucleoside analogs to be approved for treatment of HIV-infections was zidovudine (3´-azidothymidine or AZT). Acyclovir, also known as zovirax, and gancyclovir are two commonly used and highly specific anti-herpes virus agents. Their high specificity stems from the fact that the first phosphorylation step is catalyzed by the virus-encoded HSV-TK. The triphosphates of the antiherpetic analogs target the viral DNA polymerase and act as DNA chain terminators.

1.3.2 Suicide gene therapy

Suicide gene therapy or gene-directed enzyme prodrug therapy (GDEPT) aims to increase the selectivity of prodrugs for cancer cells (Niculescu-Duvaz & Springer, 2005). This is achieved by transfecting cancer cells with a foreign gene, which encodes a highly specific drug-activating protein, followed by administration of the prodrug. Only the transfected cells will be able to activate the prodrug and will finally get killed by its cytotoxic effect. This method, which is mainly suitable for treatment of solid tumors, is currently under development, where delivery of the gene to the target cells and control of the gene expression are the largest obstacles to be overcome. A number of enzyme/prodrug combinations have been explored, and the most studied system is the HSV-TK/gancyclovir combination where the gene is delivered in a replication deficient adenovirus.

2. Bacteria of interest

The enzymes whose structures are presented in this thesis originate from three microorganisms; *Ureaplasma urealyticum* and two related bacillus species *Bacillus anthracis* and *Bacillus cereus*. These pathogenic organisms are described bellow.

2.1 *Ureaplasma urealyticum*

Ureaplasma urealyticum is a member of the class Mollicutes, commonly referred to as mycoplasmas. Members of this class evolved from Gram-positive bacteria, but lack a cell wall. They are the smallest-known free-living organisms. The capability of autonomous growth, simple cell structure and small genome make mycoplasmas an interesting tool for exploration of the minimal gene set required to sustain bacterial life. A characteristic of ureaplasmas, which divides them from other Mollicutes, is their ability to hydrolyze urea to ammonia and CO_2. This reaction is catalyzed by urease and contributes to 95% of the ATP production, which is driven by the electrochemical gradient produced by the ammonium ions (Smith *et al.*, 1993).

U. urealyticum is an opportunistic pathogen commonly found in the urogenital tract of humans. It is described as causing urethritis, premature spontaneous delivery, meningitis and pneumonia in neonates (Waites, Katz & Schelonka, 2005). *U. urealyticum* can be transmitted from the mother to the developing fetus in the uterus or to the neonate at delivery. The infections are especially severe for neonates and the treatment options available at present have a series of problems. The first major dilemma results from the risks of medical therapy applied to neonates because of potential drug toxicities. Furthermore, the therapy requires frequent isolation of the microorganism from the mucosal surfaces, which is difficult to do without causing pain to the newborn. The second problem linked with therapeutic treatment of *U. urealyticum* infections is antibiotic resistance. Because of the absence of a cell wall in mycoplasmas and their simple biosynthetic mechanism the range of potential antibiotics against *U. urealyticum* is very limited. β-lactam drugs that inhibit cell wall synthesis and are thereby frequently used against Gram-positive bacteria have, in the case of *U. urealyticum,* no effect. Chloramphenicol, which inhibits the elongation step in protein synthesis, is not recommended for use in medical treatment of neonates because of bone marrow toxicity. For the same reason, and additionally because of the frequent development of resistance, tetracycline is excluded as potential drug candidate. Additional difficulties to finding potent treatment, for mycoplasma infections in pregnant women and neonates are due to limited clinical studies of the efficacy of antibiotics. Nevertheless, macrolides, for instance erythromycin, are currently suggested to be the best choice of antibiotics (Waites, Katz & Schelonka, 2005).

The genome of *U. urealyticum* (serovar 3), the most common serovar in humans, has been fully sequenced by Glass *et al.* (Glass *et al.*, 2000). *U. urealyticum* has a small genome comprising ~750 kb. Biological roles have been assigned to 53% of the 613 predicted protein-coding genes. Interestingly, no genes encoding for the *de novo* biosynthesis of purines and pyrimidines, and no gene for RNR have been identified in the *U. urealyticum* genome. This brings about the necessity to import nucleosides and to rely of the salvage pathway to keep up a stable dNTP pool for the DNA synthesis. Surprisingly, no transporters for bases or nucleotides have been identified, but these may still be found among the hypothetical *U. urealyticum* proteins that have 5 or more predicted transmembrane regions. In addition, there are genes missing in order to fulfill the salvage pathway of dNTPs *i.e.* no nucleoside diphosphate kinase or thymidylate synthase has been identified. However, glycolytic kinases such as 6-phosphofructokinase, phosphoglycerate kinase and pyruvate kinase have been suggested to be possible replacements for the missing nucleoside diphosphate kinases in mycoplasmas (Pollack *et al.*, 2002).

Two years after the genomic characterization, it was agreed that *U. urealyticum* should be divided into two species: *U. urealyticum* and *U. parvum* (Robertson *et al.*, 2002). As a consequence of the reclassification *U. urealyticum* (serovar 3) was renamed to *U. parvum*. For consistency with early literature we have kept the old name, thus whenever in this thesis or in the papers attached to it *U. urealyticum* is mentioned it refers to *U. parvum* according to the new classification.

2.2 *Bacillus anthracis* and *Bacillus cereus*

Bacillus anthracis, *Bacillus cereus* and *Bacillus thuringiensis* belong to the *Bacillus cereus* group of organisms commonly found in soil and the gut of soil-dwelling invertebrates (Jensen *et al.*, 2003). These species are genetically highly similar and their functional differences are mainly assigned to the genes carried on plasmids (Rasko *et al.*, 2005). For instance *B. thuringiensis* possesses plasmids carrying genes coding for insecticidal toxins often produced as intracellular protein crystals. Loss of these plasmids would make *B. thuringiensis* indistinguishable from *B. cereus*. Similarly, the virulence of *B. anthracis* is due to the genes present on two plasmids pX01 and pX02 (Okinaka *et al.*, 1999) which are absent in *B. cereus* or non-pathogenic *B. anthracis* strains. The plasmid pX01 harbors genes coding for three toxic factors while pX02 encodes a poly-D-glutamic acid capsule protecting the bacterium from phagocytosis (Jensen, *et al.*, 2003). As a means of survival through environmental stress the members of the *B. cereus* group form endospores which after up-take by the host germinate into vegetative, virulent bacteria, eventually killing the host. In particular spores from *B. anthracis* possess properties that make them attractive biological weapons *i.e.* they are resistant to heat, ultraviolet and ionizing radiation, high pressure and chemical agents, and in addition they are easy to produce in large quantities.

B. anthracis is the etiological agent of anthrax, a disease that is primarily found among herbivores. The infection is communicated to man via contact with infected

animals causing cutaneous anthrax, or from eating inappropriately prepared meat from anthrax-infected animals resulting in gastrointestinal anthrax. The latter is very rare and occurs mainly in Africa and in parts of Asia but it is considered to be one of the most severe forms of the disease with fatal prognosis. There is yet a third type of the disease, the pulmonary or inhalation anthrax. As with the intestinal anthrax, this form has very high mortality rate (>80%) and is difficult to treat because the infection may stay unrecognized until it is too late for treatment (Baillie & Read, 2001).

Penicillin and doxycycline are currently used for the treatment of anthrax, but also chloramphenicol, erythromycin, tetracycline and ciprofloxacin can be used (Dixon *et al.*, 1999). Doxycylcline and ciprofloxacin are also used in prophylactic treatment of patients with suspected exposure to anthrax spores, while persons at risk for exposure to anthrax spores are vaccinated. Patients suffering from cutaneous anthrax may require intravenous therapies. However, these therapies mainly focus on elimination of the bacteria and need to be combined with antitoxin therapies.

B. cereus is often associated with food poisoning. Two types of food poisonings are caused by *B. cereus*. One is characterized by diarrhea and abdominal pain and the other is associated with nausea and vomiting. The incubation time is 8-16 h before the symptoms break out and usually last for 12-24 h. *B. cereus* infections are rarely reported or diagnosed, and thus it is difficult to estimate the extent of the bacteria. The most common food product that is contaminated by *B. cereus* and its heat-resistant spores is rice, but also other dried products such as spaghetti, dried milk and spices may be infected. In addition to food poisoning *B. cereus* causes a number of non-gastrointestinal infections. These are usually associated with patients with impaired immune defense, neonates, drug addicts or patients with surgical wounds or catheters (Kotiranta, Lounatmaa & Haapasalo, 2000).

There is a variety of *B. cereus* virulent factors that cause food poisoning or play a role in non-gastrointestinal infections (Kotiranta, Lounatmaa & Haapasalo, 2000). Both multi-component as well as single protein enterotoxins with hemolytic, cytotoxic, dermonecrotic, vascular permeability and fluid accumulation effects have been found in food poisoning associated strains, whereas strains causing non-gastrointestinal infections have been reported to express collagenase and proteases *e.g.* with an ability to hydrolyze hemoglobin. Furthermore different forms of β-lactamase have been detected in different strains of *B. cereus* explaining its residence to β-lactam antibiotics.

3. Thymidine kinase

3.1 Human TK1

The monomer of human TK1 (hTK1) comprises 234 amino acids and has a molecular weight of 25.5 kDa. Discrepant data regarding the native weight of the hTK1 is presented in the literature with some reports stating hTK1 to be a dimer whereas others provide evidence for a tetrameric form of the enzyme (Berenstein *et al.*, 2000; Birringer *et al.*, 2006; Munch-Petersen *et al.*, 1991; Munch-Petersen, Tyrsted & Cloos, 1993; Sherley & Kelly, 1988a). This issue will be further discussed in the forthcoming sections. In contrast to other human dNKs, hTK1 is cell cycle regulated. The TK1 activity increases during the G1/S phase of the cell cycle when DNA is replicated, but strongly decreases during late mitosis when cell division takes place (Bello, 1974). The molecular mechanism for the cell cycle dependent activity of hTK1 was ascribed to posttranslational regulation, but is not dependent on the hTK1 mRNA levels as was previously suggested (Coppock & Pardee, 1987; Sherley & Kelly, 1988b). Moreover, it was demonstrated that the fluctuation in TK1 activity corresponds to changes in the levels of hTK1 protein rather than being due to protein inhibition, and a signal for mitotic degradation was localized to a region between 10 and 40 amino acids from the C terminus of the hTK1 (Kauffman & Kelly, 1991; Sherley & Kelly, 1988b). Kelly and co-workers correctly hypothesized that the degradation of human TK1 could be mediated by a highly specific protease that was active only during the early G1 phase of the cell cycle. Recently a detailed study has shown that the rapid degradation of hTK1 during the end of mitosis is dependent on anaphase-promoting complex/cylosome (APC/C) mediated ubiquitinylation, activated by Cdh1 (Ke & Chang, 2004). The degradation of hTK1 is initiated through binding of the APC/C-Cdh1 complex to a KEN box motif at the C terminus of hTK1. Similar KEN box motives are found in other mammalian TK1-like sequences, which consequently also exhibit cell cycle dependent TK1 activities, but are absent in most bacterial and viral TK1-like sequences where the C termini are usually shorter.

As already mentioned, hTK1 phosphorylates dT and dU, but also some nucleoside analogs: FdU, 3'-fluorodeoxythymidine (FLT) and AZT (Eriksson *et al.*, 1991; Munch-Petersen*, et al.*, 1991). A further medical application of hTK1 is as a biomarker for cancer diagnosis and post-operative follow-up studies. Elevated hTK1 activity and concentration have been observed in blood serum of patients suffering from various solid cancers as well as leukemia and lymphoma (He *et al.*, 2005). The concentration of serum hTK1 has been shown to be a particularly sensitive marker for detection and post-surgical monitoring of breast, gastric, non-small-cell lung and bladder cancer (Gronowitz *et al.*, 1984; He *et al.*, 2000; Li *et al.*, 2005; Zhang *et al.*, 2006; Zou *et al.*, 2002).

3.2 Bacterial TKs

Most bacteria, but not all, posses TK1-like enzymes (Saito & Tomioka, 1984; Saito, Tomioka & Ohkido, 1985; Sandrini*, et al.*, 2006). The first bacterial TK to be biochemically characterized was from *Escherichia coli* where thymidine kinase is the only deoxyribonucleoside kinase present (Karlström, 1970). Recently TKs from *Ureaplasma urealyticum* (*Uu*-TK) and *Bacillus anthracis* (*Ba*-TK) were cloned, purified and characterized (Carnrot*, et al.*, 2003; Carnrot *et al.*, 2006). *Uu*-TK and *Ba*-TK consist of 223 and 195 amino acids respectively, and share 44% sequence identity. Both proteins share some enzymatic properties with human TK1 *e.g.* they are strictly pyrimidine specific and only phosphorylate dT and dU of the natural substrates and are feedback inhibited by dTTP. Compared with hTK1, *Uu*-TK and *Ba*-TK are less fastidious in their choice of phosphate donor, and besides ATP also accept GTP and other (d)NTPs.

In the following section the structural studies of three bacterial thymidine kinases, *Uu*-TK, *Ba*-TK and *Bc*-TK are described. These enzymes have been trapped in different structural arrangements that are interpreted as conformational stages during the reaction pathway. The initiative to this work was based on expectations that the structural knowledge of TK1-like enzymes would contribute to future structure based drug design. As already described, lack of a *de novo* pathway for dNTP synthesis in *U. urealyticum* triggered the idea to use the salvage pathway in order to disturb the dNTP pools and consequently the DNA synthesis in the bacteria. Therefore, the possibilities of using *Uu*-TK in future anti bacterial therapies are evaluated. Finally, the sequential and structural relationship with hTK1 as well as other TK1-like structures is summarized.

3.3 *Uu*-TK in complex with dTTP (Paper I)

When our group commenced work on structure determination of thymidine kinase no homologous structures were known. Concurrent crystallization trials of human thymidine kinase 1 (hTK1) and thymidine kinase from *U. urealyticum* (*Uu*-TK) were initiated. The hTK1 and *Uu*-TK share 29% sequence identity, thus there was a possibility that once the structure of one thymidine kinase was solved it could provide a model for molecular replacement structure determination of the other. Very often things sound easier than they turn out to be and the most interesting proteins are often recalcitrant to crystallization. After two years of persistent attempts to grow well-diffracting crystals, hTK1 and *Uu*-TK, both in complex with their feedback inhibitor dTTP, were solved independently. The structure of hTK1 was solved with single isomorphous replacement with anomalous scattering (SIRAS) using gold as anomalous scatter, while the *Uu*-TK structure was determined with single wavelength anomalous diffraction (SAD) on selenomethionine (Se-Met) substituted enzyme.

3.3.1 Crystallization and structure determination

The *Uu*-TK was studied in collaboration with the department of Molecular Biosciences at SLU, Uppsala. Expression, purification and biomolecular characterization were performed by C. Carnrot (Carnrot*, et al.*, 2003). *Uu*-TK crystallized easily in different kinds of PEGs, with or without addition of salts at neutral pH. Since the protein often formed micro crystals or crystal clusters, the difficulty was to make the enzyme form single-growing, sufficiently big, diffracting crystals. The *Uu*-TK was prone to precipitate in solution and the ability to form crystals diminished with time of storage at -80°C. Things were further complicated by the fact that Se-Met substituted protein had to be crystallized fresh. Screening with Hampton Research Additive Screen indicated that dithiothreitol (DTT) increased the size of the crystals. Se-Met substituted *Uu*-TK crystallized under very similar conditions to the native protein. In order to increase the size of Se-Met crystals streak seeding (Bergfors, 2003; Ducruix & Giegé, 1992; Stura & Wilson, 1990) was carried out. A seeding stock, prepared from a crushed crystal, was used to streak seed a drop that had been left to equilibrate for 2h. This gave rather large crystals, but as before, no single crystals were obtained. Using a micro-spatula, it was possible to separate the largest crystals and the structure of *Uu*-TK could be determined.

3.3.1.1 How selenium turned out to be zinc – SAD phasing of Uu-TK

Two data sets, one native and one SAD, were collected of *Uu*-TK in complex with dTTP. The crystals of native and Se-substituted *Uu*-TK were isomorphous and belonged to the monoclinic space group $P2_1$ comprising four subunits in the asymmetric unit.

A highly redundant SAD data set was collected at the wavelength corresponding to the absorption energy at K-edge for selenium (λ=0.979Å, 12.66 keV). The crystals diffracted to approximately 3 Å and the anomalous signal was detectable to 4 Å resolution. Using data to 4 Å *SHELXD* (Schneider & Sheldrick, 2002) localized 12 heavy atom positions in the asymmetric unit with a sharp drop in occupancy between position 12 and 13. These sites where further refined in *AutoSHARP* (Bricogne *et al.*, 2003; de La Fortelle & Irwin, 1997). Initially, all 12 sites were thought to correspond to the selenium atoms, three for each monomer. The sequence of *Uu*-TK contains three methionine residues including the very first N-terminal residue. In the beginning of tracing of the polypeptide chain two of the heavy atom sites were assigned to Met21 and Met82, but the position of the third did not fit with the sequence. The first residue that was ordered was Ile11, which consequently strongly indicated that the third heavy atom site could not correspond to Met1. As the modeling proceeded it became clear that the third site did not correspond to a selenomethionine, but rather to a zinc ion, which was coordinated by three cysteines and one histidine. The anomalous contribution of zinc was possible since the absorption K edge for zinc is at a lower energy (9.66 keV, λ=1.284) than for selenium. The presence of zinc ion was later on verified by atomic absorbance spectroscopy.

3.3.2 The structure of Uu-TK-dTTP

The monomer is arranged in two domains; an α/β domain and a lasso domain (Fig. 3A). The α/β domain is the larger of the two domains comprising 150 residues of the N-terminal part of the enzyme. The α/β domain is build up from a central, six-stranded, parallel β-sheet flanked by a long α-helix on one side, and three shorter α-helices on the other side. The turn between β1 and α1 harbors a conserved phosphate binding motif, GXXXXGKS/T. The region connecting β2 with β3 forms a β-hairpin which is perpendicular to the β-sheet, and is situated on the same side as the long α-helix. However, in two subunits this region is disordered and could not be traced.

From the last β-strand of the central sheet, the chain continues to the lasso domain. The last 70 C-terminal residues fold into two perpendicular β-ribbons. The second ribbon, the longer one, opens up to form a big loop whose shape resembles a lasso. As will be described in the following section this domain mediates substrate binding, and because of its shape and function it was given the name lasso domain. The lasso loop is stabilized by two conserved residues, an arginine and a tyrosine. In contrast to the other residues found in the lasso loop, the arginine and tyrosine are the only two whose side chains are pointing into the middle of the loop and not towards the surface. The positioning of the β-ribbons with respect to each other is supported by a zinc ion. The metal is tetrahedrally coordinated by two cysteine residues of the first β-ribbon and one cysteine and one histidine residue situated after the second β-ribbon. The very last 10-15 residues form into a C-terminal α-helix, which as will be shown, participates in tetramerization of *Uu*-TK.

3.3.2.1 Catching the base with the lasso – the active site

The active site is situated at the carboxy end of the central β-sheet, in between the α/β domain and the lasso domain. The thymidine base of the feedback inhibitor is buried in a deep, hydrophobic pocket, covered by the lasso loop. All the interactions between the base and the enzyme are with main-chain atoms of residues in the lasso and α/β domains. The side chain of the tyrosine, which was mentioned earlier as a lasso stabilizing residue, stacks against the thymine. Also the sugar moiety interacts only with the main chain of the enzyme. The 3'-hydroxyl group makes hydrogen bonds with a conserved glycine in the lasso domain. The phosphates are more exposed and interact with the residues from the P-loop, where the only specific interaction with the ligand is provided by the side chain of a lysine residue. The negative charges of the phosphates are balanced by arginines, and in two subunits a Mg-ion is present in close proximity to the phosphates.

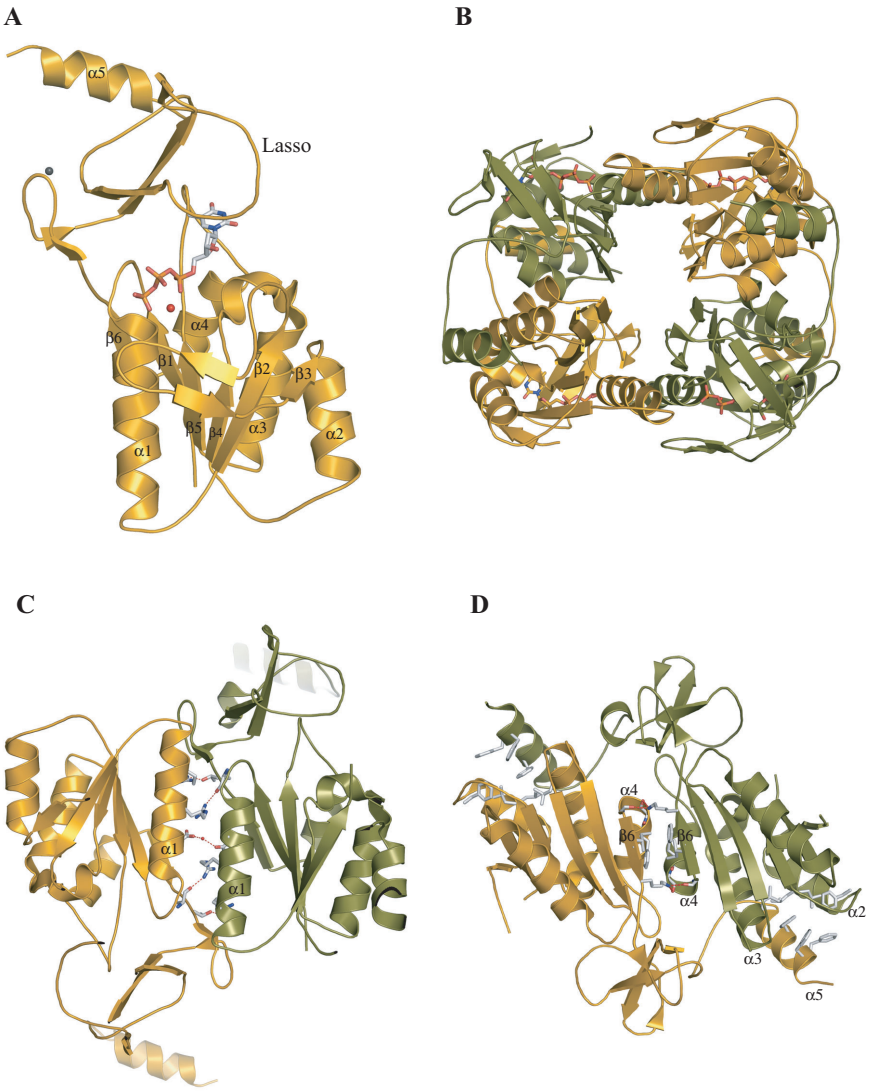


Fig. 3. The structure of *Uu*-TK. **A)** The monomer of *Uu*-TK with the feedback inhibitor dTTP. The dTTP binds as a bisubstrate analog with the base and sugar moieties in the acceptor site and the phosphates partly occupying the donor site. Structural Zn-ion is colored in gray and Mg-ion is in red. **B)** The tetramer of *Uu*-TK. **C)** The subunit-subunit interaction formed by an antiparallel helix pair. The subunits mainly interact via H-bonds formed between side-chain and main-chain atoms. **D)** The subunit-subunit interaction formed between the neighboring β-sheets. In addition to the interaction formed between the β–strands, helix α5 makes hydrophobic interactions with helices α2 and α3 of the adjacent subunit.

3.3.2.2 Quaternary structure

The *Uu*-TK forms a tetramer with 222 fold non crystallographic symmetry (Fig. 3B). Two subunit-subunit interactions exist in the tetramer. One is formed between the long α-helices (α1) between neighboring subunits such that the helices make an antiparallel helix pair (Fig. 3C). The interactions are mainly provided through hydrogen bonds between side-chain and main-chain atoms. The second interaction, also formed in an antiparallel manner, is between the β-strands β6 and the helices α4 (Fig. 3D). The edge of the adjacent β-sheets interact through conserved hydrophobic interactions, salt bridges formed between an aspartic acid and an arginine, and hydrogen bonds between the side-chain nitrogen from a lysine residue in the strand of one subunit to the backbone oxygen in the other subunit. Moreover, a chain of water-mediated hydrogen bonds holds the subunits together. This subunit-subunit interaction is further strengthened by hydrophobic interactions formed between the C-terminal helix α5, which is stretched out and interacts with helices α2 and α3 in the neighboring subunit.

3.4 The substrate complex of *Uu*-TK (Paper II)

3.4.1 Crystallization and structure determination

After determining the 3D structure of thymidine kinase, the next goal was to study the enzyme as potential drug activator and to further investigate the mechanism of phosphate transfer catalyzed by this enzyme. Attempts were made to co-crystallize *Uu*-TK together with a number of different nucleoside analogs, bisubstrate analogs such as P^1-(5'-adenosyl)-P^4-(5'-(2'-deoxythymidyl))-tetraphosphate, (AP_4dT) and P^1-(5'-adenosyl)-P^5-(5'-(2'-deoxythymidyl))-pentaphosphate (AP_5dT), as well as combinations of ADP and ATP analog adenosine 5'-[β,γ-methyleno]-triphosphate (AppCp) with dT (Fig. 4). Only the AppCp-dT crystals diffracted suitably for structure determination. As in *Uu*-TK-dTTP crystals, there were four subunits in the asymmetric unit and the space group was $P2_1$ with comparable cell parameters. The structure of this new *Uu*-TK complex was solved by molecular replacement using the inhibited *Uu*-TK as search model.

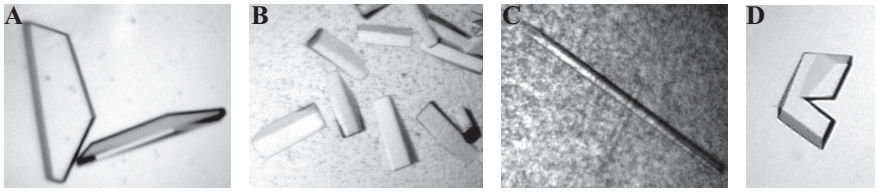


Fig. 4. Crystals of *Uu*-TK co-crystallized with different ligands. **A)** ADP + 5-FdU **B)** AP_4dT **C)** AP_5dT **D)** AppCp + dT

3.4.2 The structure of Uu-TK-dT

Only the substrate was bound to the *Uu*-TK and no density for the ATP analog was detectable. The phosphate donor site is occupied by water molecules and in one subunit by a Tris molecule. The overall structure of *Uu*-TK in complex with its substrate, dT, is highly similar to the inhibited enzyme. The main difference is focused to the phosphate binding area, where the flexible β-hairpin is absent or adopts a highly different conformation than observed in the dTTP complex. The substrate binds in the same manner as the base and deoxyribose moiety of the feedback inhibitor, with the thymine in a hydrophobic pocket between the α/β domain and the lasso domain. As for the dTTP molecule, hydrogen bond donors and acceptors of the thymine interact with the main-chain atoms of the enzyme (Fig. 5). The methyl group points towards the Cβ of a serine residue and is surrounded by other hydrophobic residues. In contrast to the dTTP complex, the side chain of Glu97 is directed into the active site and is hydrogen-bonded to the 5'-hydroxyl group of the substrate. This conformation is in agreement with the proposed mechanism for phosphate transfer, where the glutamic acid plays the role of a catalytic base, as described in the introduction. A similar conformational change of the catalytic base, where the phosphates of the feedback inhibitor repel the side chain of the negatively charged glutamic or aspartic acid, has been described for the *D. melanogaster*-dNK (Mikkelsen *et al.*, 2003) and UCK2 (Suzuki*, et al.*, 2004).

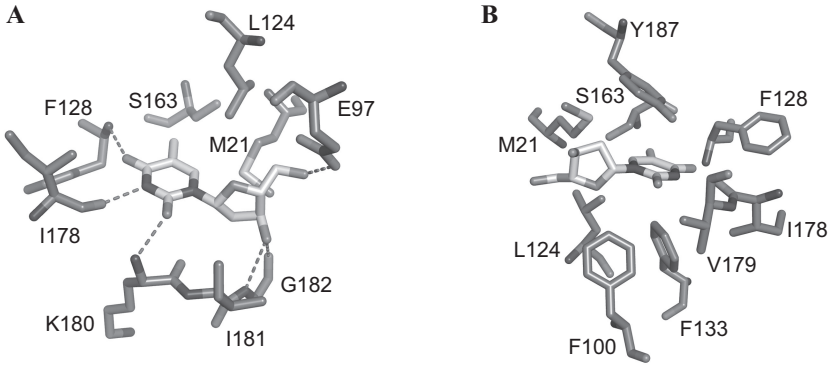


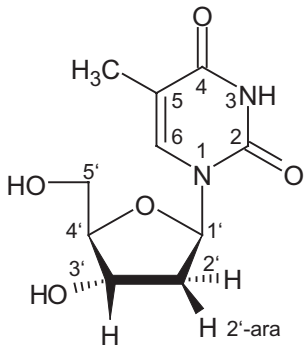
Fig. 5. Two views of the acceptor binding site of *Uu*-TK.
A) Thymidine interacts only with main-chain atoms of the enzyme. The methyl group is in a hydrophobic environment surrounded by Cβ of Ser163, Leu124 and Met21. The side chain of Glu97 is H-bonded with 5'-OH of the dT and may play the role as catalytic base. **B)** The thymidine base is surrounded by hydrophobic residues from the α/β domain and lasso domain.

Based on this structure we can confirm that the dTTP acts as a bisubstrate inhibitor, where the base and sugar moieties of the inhibitor are bound in the phosphate acceptor site and the phosphates partly occupy the phosphate donor site.

3.4.3 Uu-TK as drug target

Uu-TK was investigated as potential prodrug activator. Several nucleoside analogs were tested as substrates or inhibitors of *Uu*-TK and the results were compared with hTK1 and put in light of the substrate complex of *Uu*-TK.

The analogs that were studied included modifications either of the base (positions 3 or 5) or of the sugar (positions 2' or 3') (Fig. 6). The highest activity was obtained with 5-halogenated analogs. Substitutions as big as a cyclopropyl group at this position were accepted. Larger groups such as 5-(2-bromovinyl)-dU caused a significant decrease of activity, probably due to steric hindrance with residues that line the narrow pocket perfectly suited for a methyl group of dT. Analogs with alkyl modification at N3 position all showed lower activity compared with dT. Normally N3 is hydrogen-bonded to the carbonyl oxygen of a backbone residue in the lasso loop, but this interaction is lost with alkyl substitution. This prevents proper binding of the lasso loop which might be important for effective phosphorylation.



Position	Substitution	Activity (%) *Uu*-TK
3	Small alkyl groups	15-20
5	Halogens, ethyl, or no groups	50-100
	Bulky goups	<0.1
2'	F	<2
2'-ara	F	35-45
	Hydroxyl	<2
3'	F	50
	N3	35
	No group	<5

Fig. 6. The activity of *Uu*-TK with various substrate analogs. The percentage values indicated correspond to the enzyme activity with 100 μM analog as compared to the activity with 100 μM dT.

When it comes to modifications of the sugar moiety, it was shown that analogs with small, polar groups at the 3'-position were accepted as substrate with about 50% of activity retained, whereas nonpolar substitutions resulted in loss of activity. Similarly as for the N3 substitution this is connected with the possibility of hydrogen bond formation between the backbone of the lasso loop and the substrate. The 2'-position turned out to be the least suited for modification. Analogs altered at this position demonstrated the lowest activities of all analogs tested. The binding site is rather crowed around the 2'-position so the great loss of activity with these analogs was expected.

Two nucleotide analogs, 3'-fluoro-dT (FLT) and 3'-fluoro-5-cyclopropyl-dU (FCPU), showed slightly higher relative activities with *Uu*-TK than with hTK1. However, subsequent kinetic experiments showed that hTK1 was more efficient with these analogs as substrates than *Uu*-TK.

Moreover, some of the nucleotides were tested for their ability to act as inhibitors for thymidine phospharylation. The *Uu*-TK and hTK1 show diverse abilities to withstand inhibition. For instance 5-fluoro-2'-deoxyuridine (FdU) gave a 20 times higher IC_{50} value (the concentration needed to inhibit 50% of the enzyme activity) for *Uu*-TK compared with hTK1. On the other hand, 2'-fluoro-arabinosyl-5-iodo-2'-deoxyuridnie (FIAU) and 3'-azido-methyl-2'-deoxythymidine (AZMT) showed approximately 5 times lower IC_{50} values for *Uu*-TK than for hTK1.

To summarize, we have seen that analogs with modification at 5- and 3'-position were readily phosphorylated by *Uu*-TK, but at the same time these analogs were also equally well accepted as substrates by hTK1. In order for a compound to be considered as a potential therapeutic agent it must show a significantly different enzymatic pattern with the target enzyme than with the corresponding enzyme of the host. These results show that we are far from finding a good antibiotic that either would act as an exclusive substrate for *Uu*-TK or would be a selective inhibitor against *Uu*-TK.

3.5 Thymidine kinase from *B. anthracis* and *B. cereus* (Paper III)

The project on bacillus TKs was carried out in collaboration with the group of Prof. J. Piskur currently at the department of Cell and Organism Biology at Lund University, Sweden (*Bc*-TK) and with the group of Prof. S. Eriksson at the department of Molecular Biosciences at SLU, Uppsala, Sweden (*Ba*-TK). These groups kindly provided us with the material for structural studies.

Thymidine kinases from *B. cereus* and *B. anthracis* have 96% sequence identity. In addition to five C-terminal amino acids, the enzymes differ by three residues. These residues are situated far from the active site, they are not part of the tetramerization regions, and do not, in any other aspect, affect the structures of the enzymes. Hence, *Bc*-TK and *Ba*-TK may be considered as structurally identical.

3.5.1 Crystallization and structure determination

The *Bc*-TK crystallized in a variety of conditions, forming different crystal forms (Fig. 7), but it was only the bipyramidal crystals, grown in a high concentration of 2-methyl-2,4-pentanediol (MPD), that diffracted. Similarly, the crystals of *Ba*-TK adopted a bipyramidal morphology and likewise *Bc*-TK grew in an organic precipitant, 1,2-propanediol.

Both the *Ba*-TK and the *Bc*-TK crystals belonged to the tetragonal space group $I4_122$ with the asymmetric unit comprising one subunit. However, the crystal packing differs significantly for the *Ba*-TK and *Bc*-TK, which is mirrored in the difference in the cell parameters. The structure of *Ba*-TK was solved by molecular replacement using a polyalanine model of *Uu*-TK. Subsequently, the structure of *Bc*-TK was solved using the structure of *Ba*-TK.

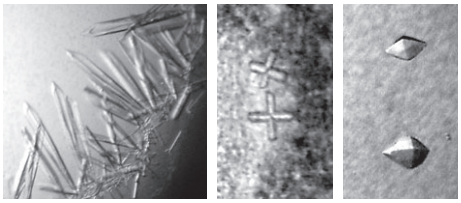


Fig. 7. Different crystal forms of *Bc*-TK. The data were collected from the bipyramidal crystals, which resemble the *Ba*-TK crystals.

3.5.2 The structure of Ba-TK

The monomer of *Ba*-TK is highly similar to the *Uu*-TK monomer (Fig. 3). The structures can be superimposed with root-mean-square distance (r.m.s.d.) of 0.8 Å for 165 Cα atoms. The substrate, dT, is bound in the same way as described for *Uu*-TK, with the base in a hydrophobic environment created by residues from the α/β and the lasso domains. The side chain of the catalytic base, Glu89, takes the same conformation as the corresponding residue in the substrate complex of *Uu*-TK, towards the active site, and forms a hydrogen bond with the 5' hydroxyl group of the deoxyribose. There was no electron density for the β-hairpin (residues 46-68) situated in the α/β domain and this is the only part of the structure missing in the model. A tetramer, equivalent to the *Uu*-TK tetramer, is formed by application of symmetry operators.

3.5.3 The structure of Bc-TK

The overall topology of *Bc*-TK is not surprisingly highly similar to *Ba*-TK and *Uu*-TK. However, a few interesting differences exist which need to be highlighted. The *Bc*-TK was co-crystallized with the feedback inhibitor, dTTP, hence, it was very intriguing to find the acceptor site empty and instead discover clear density for a nucleoside triphosphate at the phosphate donor site. Since a high concentration of ATP was used during one of the purification steps, the density was initially thought to represent an ATP molecule. However, the possibility that the density represented the dTTP molecule bound backwards compared to its binding as a feedback inhibitor was not ruled out, and both ligands were considered during the refinement process. Judging from the electron density corresponding to the base and sugar moieties, the ligand was interpreted as a dTTP molecule. According to the biochemical data available for *Uu*-TK (Carnrot, *et al.*, 2003) and hTK1 (Ives, Morse & Potter, 1963), dTTP is strictly an inhibitor and does not act as phosphate donor in TK1-like enzymes. Thus, it was surprising to find the dTTP mimicking a phosphate donor in the *Bc*-TK structure. However, later characterization of the *Ba*-TK showed that this enzyme may use dTTP as

phosphate donor, but to a lower degree than ATP (Carnrot*, et al.*, 2006). Shortly, it became clear why dTTP was bound as a phosphate donor and not as a feedback inhibitor despite the fact that it has a much higher affinity for the acceptor site than for the donor site. During the final runs of refinement, electron density, too big to correspond to a water molecule, became visible in the acceptor site in the exact position for the base of a substrate or feedback inhibitor. The density was assigned to represent an MPD molecule, which was used as crystallizing agent. Finding MPD at or near the active sites is not unusual and has been reported for a number of structures. A review of MPD as crystallizing agent is available (Anand, Pal & Hilgenfeld, 2002).

Although the acceptor site was occupied by an MPD molecule, the ligand did not provide sufficient hydrogen bonding partners to stabilize the lasso loop. Consequently, the lasso loop is in an open conformation. Similar flexibility in the lasso loop is present in the structure of thymidine kinase from *C. acetobutylicum* (*Ca*-TK) in complex with ADP (Kuzin, 2004), where parts of this region are missing in the model. In *Bc*-TK the open conformation of the lasso was stabilized by crystal contacts which enabled tracing the entire loop.

Let us now move to the phosphate donor binding site, where the most interesting conformational differences are found. An occupied donor site influences the conformation of the tetramer. The tetramer of *Bc*-TK is not as compact as the tetramer of *Ba*-TK, or of any other TK1-like enzyme with an empty donor site. The distance between two subunits at the helical interaction is 3 Å wider in *Bc*-TK, and this allows the base of the dTTP to stack between two phenylalanine residues, one from each subunit (Fig. 8). The separation of the subunits and the stacking interaction of the phosphate donor base between hydrophobic residues of neighboring helices are in agreement with the *Ca*-TK structure in complex with ADP. The opening of the tetramer, or formation of a less tight tetramer, is necessary to create space for the base of the nucleotide to stack between the adjacent helices. Moreover, the open tetramer prevents steric clashes that otherwise would occur between the P-β-hairpins of neighboring subunits, and between the residues situated on the tip of the P-β-hairpin and the neighboring α-helix. Moreover, the phosphates of the donor provide stabilizing binding possibilities for the residues in the P-β-hairpin. Accordingly, the P-β-hairpin could be fully traced in the *Bc*-TK structure, and the tip of the hairpin hydrogen bonds with phosphates of the dTTP.

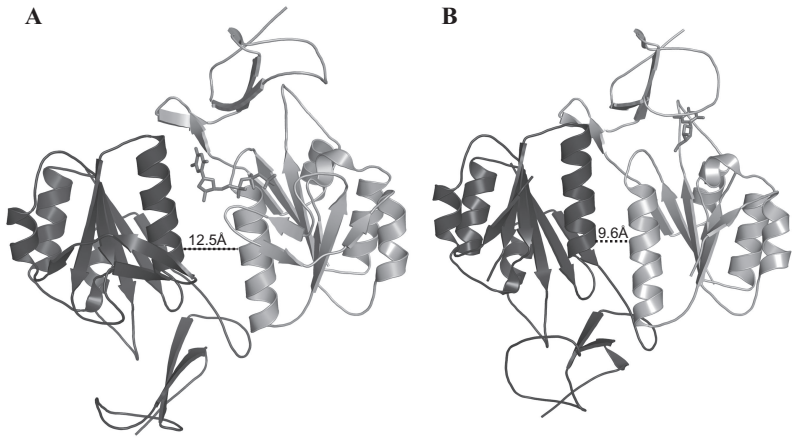


Fig. 8. The distance between two subunits at the helical interaction is 3 Å wider when the phosphate donor is bound as in *Bc*-TK with dTTP (A), compared to when the donor site is unoccupied as in *Ba*-TK with dT (B). In *Ba*-TK with dT in the acceptor site the lasso loop is closed down over the active site, whereas in *Bc*-TK the loop has an open conformation. The distances in the figure are measured between Cα–atoms of residues 23 and 26 of the neighboring helices.

3.6 Comparison of TK1-like enzymes

3.6.1 Amino acid comparison

Amino acid sequence alignment divides bacterial TKs in two groups: the TKs from Gram-negative bacteria and TKs from Gram-positive bacterial including mycoplasmas (Fig. 9). The sequence identity is 60-70% within each group but only around 20-30% between the groups. Some of the conserved motifs, both of enzymatic and structural importance, were identified first after the TK-structures were determined. Below, the TK sequences of Gram-positive and Gram-negative bacteria are compared and discussed. The analysis is based on 11 TK sequences from Gram-positive bacteria, 10 mycoplasmic TKs and 5 TKs from Gram-negative bacteria. An alignment of selected sequences is provided in figure 9. A similar sequence relationship in TK1-like kinase of bacterial origin has been reported (Sandrini*, et al.*, 2006).

The phosphate binding motif, P-loop (GXXXXGKS/T)(Saraste, Sibbald & Wittinghofer, 1990) is the only conserved region among all dNKs including enzymes of both the dNK-family and TK1-like enzymes. It is situated on the loop connecting the first β-strand, β1, with the first α-helix, α1. The residues of this region take part in positioning the phosphates of the phosphate donor and ligate the phosphates of the feedback inhibitor. The significance of the lysine has been ascertained by numerous mutation studies of a number of GTP/ATP binding proteins where the mutants demonstrated decreased enzymatic activity and reduced substrate affinity (Reinstein, Schlichting & Wittinghofer, 1990; Sigal *et al.*, 1986; Tian *et al.*, 1990). In Vv-TK, substitution of the conserved lysine by a

hydrophobic isoleucine decreased the ATP affinity and led to loss of enzymatic activity (Black & Hruby, 1990a). The first glycine in the P-loop is substituted with a serine in Gram-negative bacteria. A small, polar residue at this position should not cause any major steric hindrance for ATP binding and the hydroxyl group of the serine residue should be able to interact with the oxygen atoms of the phosphate groups. Changing any of the glycines to valine residues, however, was shown to significantly lower the enzyme activity, probably because the larger valine residues hinder the access of ATP to the binding site or cause conformational changes of the protein (Black & Hruby, 1990a; Liu & Summers, 1988; Reinstein, Brune & Wittinghofer, 1988).

The region of the P-β-hairpin contains five conserved residues (D51, R53, S60, R61 and G63 using the numbering of the *Uu*-TK sequence). In those TK structures where this region is visible, it can be observed that the aspartic acid (D51) hydrogen bonds with the 3'-hydroxyl group of the acceptor deoxyribose. The side chain of one of the arginines (R53) points toward the active site and it is likely that this residue will stabilize the transition state. The second arginine (R61) is in most Gram-positive bacteria substituted by a histidine. In the structure of *Bc*-TK this residue is positioned at the tip of the P-β-hairpin and together with the preceding serine hydrogen bond to the phosphate donor.

Approximately 80 residues from the P-loop there is a conserved DEXQF motif is in all bacterial TKs. Generally there are four hydrophobic residues before the aspartic acid in Gram-positive bacteria but three in Gram-negative. It has been suggested that the corresponding aspartic acid residue in Vv-TK (D82) is involved in Mg-binding (Black & Hruby, 1992). The available TK1-like structures reveal that the corresponding aspartic acid indeed is oriented toward the metal, but the distance between the magnesium ion and the side chain is rather long to be considered as a metal-ligand bond. The glutamic acid which follows the asparagine has been suggested to act as the catalytic base which deprotonates the 5'-hydroxyl group of the deoxyribose (Papers I & II), while the phenylalanine is one of the hydrophobic residues of the acceptor binding site. The glutamine has a structural role.

Further conserved motifs among bacterial kinases are situated at the C-terminal part of the sequence. Black and Hruby made an early observation (Black & Hruby, 1990a; Black & Hruby, 1992) that all the TK1-like sequences that were available at that time, contained a conserved cysteine residue (C170 in VvTK, C191 in *Uu*-TK and C183 in *Ba*-TK/*Bc*-TK). Incorrectly, they speculated that this residue was involved in nucleophilic attack on ATP or in ligating the Mg ion. The structures of TK1-like enzymes have made us attentive to two Zn-binding motifs, CXXC and CXXC/H similar to those in other well known, zinc-containing proteins such as the tumor repressor p53 (Cho *et al.*, 1994) or the glucocorticoid receptor (Härd *et al.*, 1990), as well as the adenylate kinases of Gram-positive bacteria (Berry & Phillips, 1998; Gilles *et al.*, 1994). In between the two zinc pairs there are two additional, structurally important and highly conserved residues - an arginine and a tyrosine. These are the residues holding up the lasso loop.

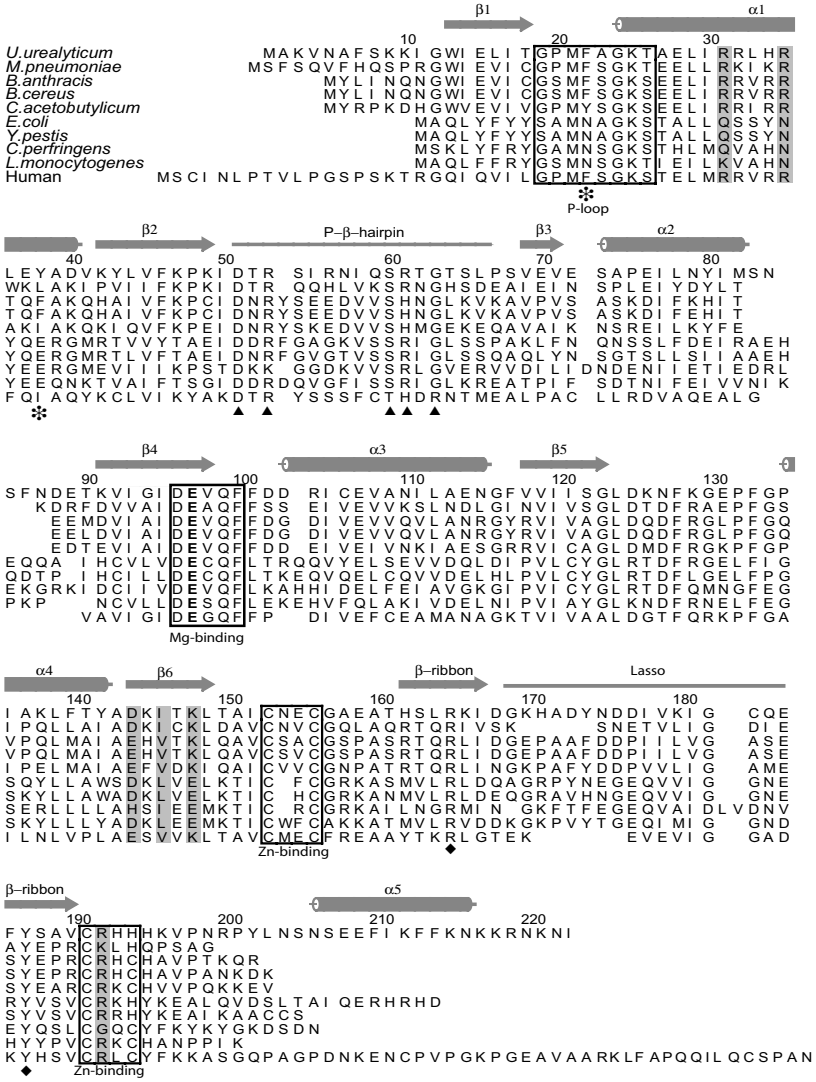


Fig. 9. Sequence alignment of TK1-like sequences from mycoplasmas (*U. urealyticum* (Q9PPP5), *M. pneumoniae* (NP_109732.)), Gram-positive bacteria (*B. anthracis* (Q81JX0), *B. cereus* (DQ384595), *C. acetobutylicum* (NP_349490), *C. perfringens* (NP_563120), *L. monocytogenes* (NP_466067)), Gram-negative bacteria (*E. coli* (NP_415754), *Y. pestis* (NP_405720)) and human (P04183). The secondary structure elements and the sequence numbering shown above the alignment correspond to *Uu*-TK. The P-loop, the Mg-binding and the Zn-binding motifs are boxed. Residues indicated with a star stack the base of the phosphate donor, a triangle indicates the conserved residues in the P-β-hairpin, while the squares indicate the Tyr-Arg couple, which stabillizes the lasso loop. The catalytic base is in bold. Residues that take part in the subunit-subunit interactions are shown against gray background. It is interesting to notice that TKs from the Gram-positive *C. perfringens* and *L. monocytogenes* align with TKs from Gram-negative bacteria. The alignment was made using ClustalW (Chenna *et al.*, 2003), and Alscript (Barton 1993).

So far we have not come across any major discrepancies between the Gram-positive and Gram-negative TK sequences. However, the residues involved in the binding of the adenine moiety of the ATP differ. As is observed in the structures of *Bc*-TK (Paper III) and *Ca*-TK (Kuzin, 2004), the base of the phosphate donor is stacked between two hydrophobic residues, one from each adjacent subunit. At the equivalent positions in Gram-negative bacteria there is an asparagine and a glutamic acid. There is no structure of any TK from Gram-negative bacteria available, but it is clear that these organisms will bind the phosphate donor in a different way.

Also residues involved in subunit–subunit interactions differ between Gram-positive and Gram-negative bacteria, but these differences will most likely not affect the oligomeric form of the TKs. The β-sheet interaction involves the E/DXX$_h$XK/E motif, situated on strand β6, and an arginine situated in the second Zn-binding pair, CRXC/H. The hydrophobic residue, X$_h$, is predominantly a valine in Gram-positive bacteria, whereas Gram-negative prefer a leucine or isoleucine at this position. The subunits interact through van der Waals interactions between the hydrophobic residues and by hydrogen bonds formed between the glutamic or aspartic acid and arginine, and between lysine or glutamic acid and main chain atoms. The subunit interaction formed between α-helices involves hydrogen interaction between side chains and the main-chain atoms. In Gram-positive bacteria there are two arginines that interact with the backbone of the neighboring α-helix. In Gram-negative bacteria, these arginines are substituted by glutamine and asparagine, thus the possibility to form similar subunit interactions as in Gram-positive bacteria is retained.

Finally, the hydrophobic active site residues will be explored. In general all the phenylalanines that build up the active site are well conserved, but the hydrophobic residue situated on the lasso loop (I178 in *Uu*-TK) is in Gram-negative bacteria a polar glutamine. Furthermore, the position near the methyl group of the thymine differs among TKs. In *Uu*-TK there is a serine residue (S163) whereas other Gram-positive bacteria, as well as hTK1, have a threonine at this position. In Gram-negative bacteria, valine is the most frequent residue at this position. Interestingly, *C. perfringens* has an asparagine facing the methyl group of the thymidine. The effect of these alterations on the enzymatic activity has not yet been explored.

3.6.2 Structural comparison

3.6.2.1 Comparison of bacterial TKs with hTK1

Despite moderate sequence identity between the bacterial TKs and hTK1, below 30%, the 3D structures are highly similar, indicating a common ancestor. The subunits of hTK1 and *Uu*-TK may be superimposed with an r.m.s.d of 1.2 Å for 147 Cα atoms. Like the bacterial TKs described in this thesis, the crystal structure of hTK1 is a tetramer with comparable subunit-subunit interactions as illustrated for the bacterial TKs (Paper I; Birringer*, et al.*, 2005). Also the topology of the monomer is similar to the bacterial TKs (Fig. 3). The P-β-hairpin region is only visible in one monomer of the available hTK1s (Birringer*, et al.*, 2005) and is very

dissimilar from the β-hairpin observed in bacterial TKs. In hTK1 this region folds into two short α-helices that are involved in crystal packing, thus the unusual conformation might be a result of crystallization. The lasso loop, which has not been found in any other protein, is slightly shorter than in *Uu*-TK and other bacterial TKs. Since a C-terminally truncated hTK1 was used for structure determination, there is no structural data for the last 41 residues of the hTK1, which in hTK1 harbors the KEN sequence predicted to serve as ubiquination ligase signal involved in mitotic degradation of hTK1 (Ke & Chang, 2004). A model for this part of the enzyme is provided by the full-length *Uu*-TK structure.

3.6.2.2 *Bc*-TK compared to *Ca*-TK

The structures of *Bc*-TK with dTTP and *Ca*-TK in complex with ADP (but empty acceptor site) (Kuzin, 2004) are so far the only TK1-like structures with occupied phosphate donor sites. In agreement with what we have observed in the *Bc*-TK, the lasso loop of *Ca*-TK is in an open conformation and the tetramer shows a similar, loose conformation as in *Bc*-TK. The distance between the α-helices that make up the helical subunit interface is equivalent to the *Bc*-TK, but longer than in *Uu*-TK, *Ba*-TK and hTK1.

3.6.2.3 Search for structural relatives to TK1-like enzyme family

A DALI (Holm & Sander, 1993) search reveals that the α/β domain of TKs is structurally more related to enzymes from the RecA-F_1ATPase family (Story, Weber & Steitz, 1992), in particular CobA (Bauer *et al.*, 2001), rather than to other dNKs (Fig. 10A,B). The α/β domain of TK1-like enzymes differs from the α/β domain of dNKs in strand order, the number and size of α-helices and in the architecture of the active site. The lasso loop is substituted by a helical domain and an arginine rich LID domain, which cover the active site (Fig. 10C). Despite the structural differences pointing to a different evolutionary origin, TK1-like enzymes and dNKs catalyze the phosphorylation of deoxyribonucleosides in a similar way and all are feedback inhibited by the end products of the salvage pathway.

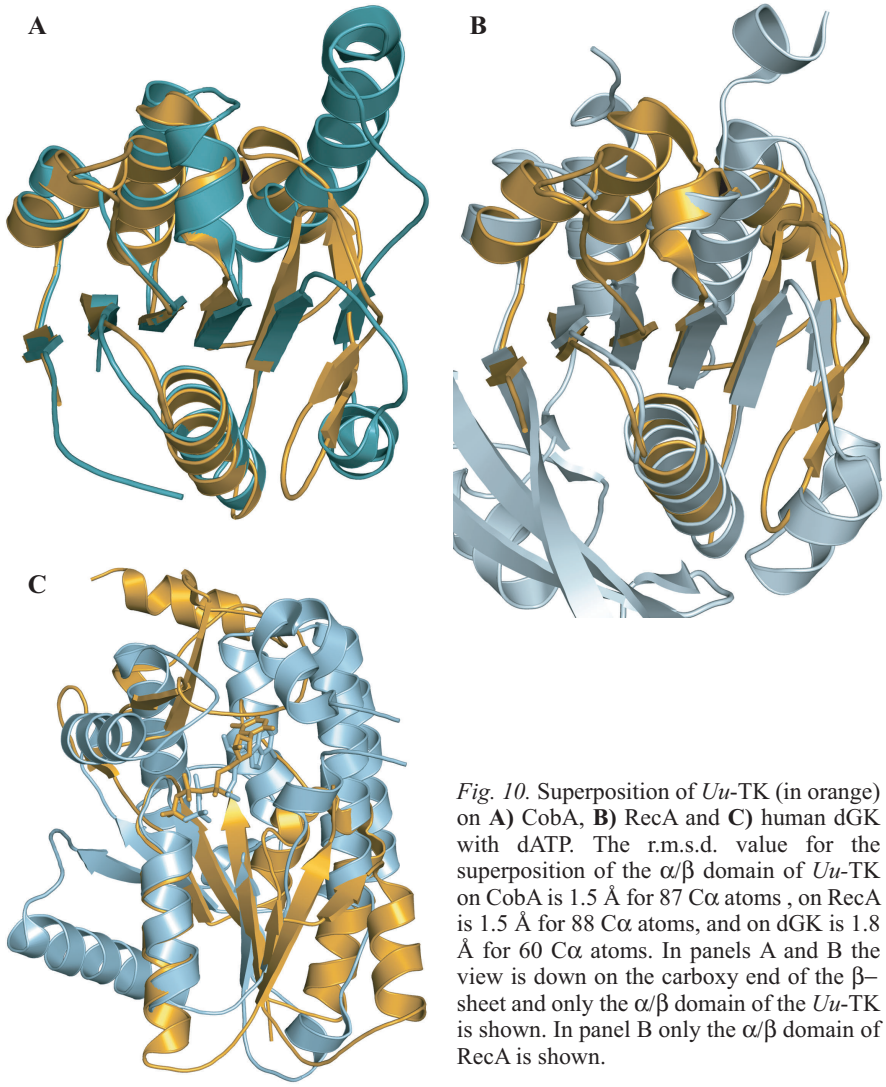


Fig. 10. Superposition of *Uu*-TK (in orange) on **A)** CobA, **B)** RecA and **C)** human dGK with dATP. The r.m.s.d. value for the superposition of the α/β domain of *Uu*-TK on CobA is 1.5 Å for 87 Cα atoms , on RecA is 1.5 Å for 88 Cα atoms, and on dGK is 1.8 Å for 60 Cα atoms. In panels A and B the view is down on the carboxy end of the β–sheet and only the α/β domain of the *Uu*-TK is shown. In panel B only the α/β domain of RecA is shown.

3.7 Concluding remarks

3.7.1 Conformational changes through the reaction

The different complex structures of thymidine kinases presented in this thesis allow us to speculate about the conformational changes that occur during the course of reaction. The monomer may be considered to consist of three domains: a rigid α/β domain and two flexible domains, the P-β-hairpin and the lasso. The flexible domains become stabilized upon ligand binding. When the enzyme is empty, awaiting the substrates to bind, both the lasso loop and the P-β-hairpin are unordered and the active site is accessible for the substrates (Fig 11A). As soon as

the phosphate acceptor binds, the lasso becomes ordered and closes down over the substrate, while the phosphate donor stabilizes the P-β-hairpin (Fig 11B). The flexible domains are also stabilized by the feedback inhibitor, such that the lasso is closed down over the active site (Fig 11C).

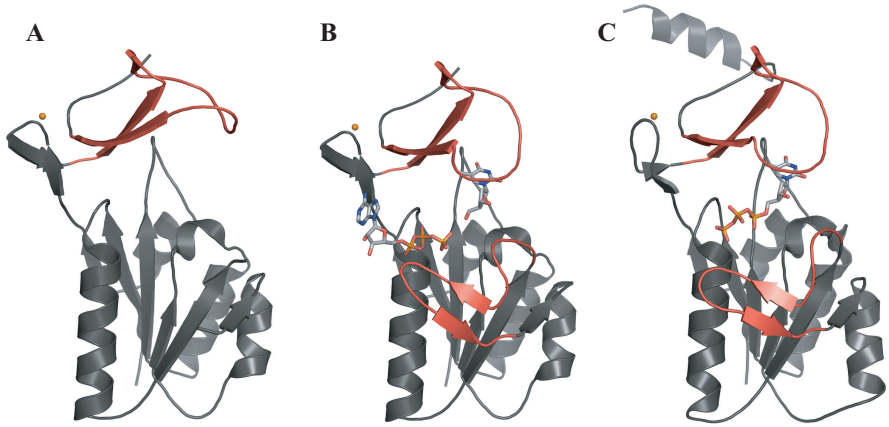


Fig. 11. Different conformations of TK1-like enzymes. **A)** Awating the substrates - lack of stabilizing interactions from the ligands leaves the lasso loop open. The P-β-hairpin could not be traced. **B)** Ready for the reaction to take place - both substrates, ATP and dT, are bound and stabilize the lasso and the P-β-hairpin in their closed conformations. **C)** Inhibited - the feedback inhibitor occupies both the phosphate acceptor and phosphate donor sites. The lasso is closed and the P-β-hairpin has a slightly different conformation than in the active form. The model in panel A is a combination of the α/β domain from *Ba*-TK and the open lasso domain of *Bc*-TK. In panel B the α/β domain from *Bc*-TK was combinde with the lasso domain of *Ba*-TK. ATP was modeled based on the dTTP binding in *Bc*-TK. Panel C shows *Uu*-TK with dTTP.

3.7.2 Quaternary structure

The quaternary structure of TK1-like enzymes has for a long time been a matter of debate. In the beginning of the 1980s, Hruby and Ball suggested that the vaccinia virus TK acts as a tetramer (Hruby & Ball, 1982), which some years later was verified by a number of biomolecular techniques (Black & Hruby, 1990b). However, contradictory results exist regarding the quaternary structure of the hTK1. It was shown that hTK1 forms dimers in the absence and tetramers in the presence of ATP (Munch-Petersen, Tyrsted & Cloos, 1993). Both the dimeric and the tetrameric forms were active, but showed different enzymatic properties. However, the hTK1 has also been reported to form tetramers regardless of the ATP incubation (Birringer*, et al.*, 2006). Considering bacterial thymidine kinases, the *Uu*-TK and *Ba*-TK have been reported as dimers in solution (Carnrot*, et al.*, 2003; Carnrot*, et al.*, 2006), but *Ba*-TK and *Bc*-TK have also been shown to form tetramers (Paper III). Nevertheless, all available crystal structures of TK1-like enzymes demonstrate similar tetrameric arrangement. The formation of the tetramer is independent of the space group and we can rule out that the tetramer is an artifact of crystallization. The oligomeric state of TK1-like enzymes might be a matter of concentration, but so far no unambiguous experiment has proven this to

be a correct assumption. From the structural point of view both the tetramer and the dimer should be able to exhibit enzymatic activity. The most probable dimer would in this case be the one formed by the β-sheet subunit-subunit interaction. This dimer would comprise two separate active sites with the bases of the phosphate donors partly exposed.

4. Uridine-cytidine kinase

Uridine-cytidine kinase (also called uridine kinase, uridine monophosphokinase, cytidine monophosphokinase; EC 2.7.1.48) catalyzes the phosphorylation of uridine or cytidine to their corresponding nucleoside monophosphates, UMP and CMP respectively. UCK exists in most cells and tissues, and it has been observed that tumor cells generally express higher levels of UCK than normal cells (Herzfeld & Raper, 1979). Ehrlich ascites cells showed the highest activities of the enzyme, and therefore, most of the early work on UCK, was done with isolates from this particular type of tumor cells.

UCK was proposed to be allosterically regulated, involving inhibition by CTP and UTP, and activation by ATP (Cheng, Payne & Traut, 1986). It was suggested that the feedback inhibitors, CTP and UTP, through binding to a regulatory site different from the active site, promoted dissociation of the active tetramer into inactive monomers. One of the substrates, ATP, was believed to reverse the effect of the feedback inhibitors, and to promote tetramer formation. Moreover, the transformation between the different polymeric states was proposed to be temperature dependent. However, new biochemical data (Ropp & Traut, 1998) and structural studies (Suzuki*, et al.*, 2004) proved the hypothesis of allosteric activation wrong. Instead, it was shown that the feedback inhibitors act as bisubstrate analogs by occupying the phosphate acceptor site along with part of the phosphate donor site, and no dissociation of the tetramer upon inhibitor binding could be detected (Suzuki*, et al.*, 2004).

Cloning and expression of cDNA encoding UCK has facilitated the molecular characterization of the enzyme. Firstly, cDNA encoding UCK from mouse brain was cloned and expressed (Ropp & Traut, 1996), which was followed by cloning and expression of human UCK cDNA (Koizumi *et al.*, 2001; Van Rompay*, et al.*, 2001). Two separate enzymes were identified and named UCK1 and UCK2 (Van Rompay*, et al.*, 2001). UCK1 consists of 277 amino acids while UCK2 is somewhat shorter and contains 261 amino acids. This gives predicted molecular weights of 31 and 29 kDa for UCK1 and UCK2 subunits, respectively. The amino acid sequence identity between UCK1 and UCK2 is about 70%.

Kinetic properties have been determined showing preferred phosphate acceptors and donors of the recombinant UCKs (Van Rompay*, et al.*, 2001). Both enzymes are strict pyrimidine kinases, only phosphorylating uridine and cytidine, with higher V_{max} with cytidine. Moreover, UCK1 exhibited 3- to 6-fold higher K_M values for uridine and cytidine, and based on the k_{cat}/K_M values, was less efficient than UCK2. Both ATP and GTP act as efficient phosphate donors, and it has been reported that also deoxyribose triphosphates may act as phosphate donors.

Interestingly, it was found that the expression of UCK1 and UCK2 is tissue dependent. High levels of UCK1 were detected in liver, kidney, skeletal muscle and heart and low amounts were found in brain, placenta, small intestine and spleen. On the other hand only low levels of UCK2 were measurable in placenta

(Van Rompay, *et al.*, 2001). There are no clear explanations to this tissue specific expression but it is speculated to be coupled with regulatory processes.

In addition to the natural substrates, UCK1 and UCK2 phosphorylate a number of nucleoside analogs, but with different phosphorylation efficiency (Van Rompay, *et al.*, 2001). Recent studies reveal that UCK2, but not UCK1, is responsible for the phosphorylation of nucleoside analogs. Especially interesting for their antitumor activities are cyclopentenylcytosine (CPEC), which was undertaken for clinical trials in 2000 (Verschuur *et al.*, 2000), and 3'-ethynylcytidine (ECyd), which entered clinical trials in 2004 (Murata *et al.*, 2004). In order to display their pharmacological effect, both these nucleoside analogs require activation through phosphorylation by UCK and subsequent kinase. ECyd-triphosphate targets RNA Polymerase (Kang *et al.*, 1989), while CPEC-triphosphate has been shown to contribute to decreased CTP pools probably by acting as a CTP synthetase inhibitor (Tabata *et al.*, 1997). Inhibition of RNA synthesis makes these analogs more efficient against slow growing solid tumors, where only a minor fraction of the cells is in the S phase and inhibition of the DNA synthesis is not possible.

4.1 Structure determination of UCK1 (Paper IV)

As part of the structure determination strategy practiced at Structural Genomics Consortium (SGC) a variety of UCK1 constructs were undertaken for crystallization trials. The constructs were designed based on domain boundary analysis using software created at SGC, and the UCK2 structure. Several constructs in the presence and absence of natural ligands (phosphate acceptors, phosphate donors, products and combinations thereof) gave rise to needle shaped crystals. These crystals belonged to the monoclinic space group C2 encompassing eight monomers in the asymmetric unit, giving rise to low quality diffraction data with the best resolution reaching 3 Å.

The structure of UCK1 described here is based on a construct with N-terminal and C-terminal deletion of 21 and 35 amino acids, respectively. Particularly, the removal of the C-terminal residues facilitated better crystal packing, which resulted in hexagonal crystals belonging to the space group $P6_222$, with one monomer in the asymmetric unit. The protein was crystallized in the presence of CMP. To some crystals the second product, ADP, was soaked in.

The structure of UCK1 was solved by the molecular replacement method using one subunit of the product bound UCK2 as the search model. After several cycles of model building and refinement TLS refinement (Winn, Isupov & Murshudov, 2001) with 5 groups was applied, which lowered the R and R_{free} by approximately 4% each. The TLS groups were obtained from the TLS Motion Determination server based on thermal motion of the model (Painter & Merritt, 2006).

4.2 The structure of UCK1

4.2.1 Overall structure

Human UCK1 is a member of the NMPK fold family, and is build up from four domains: the CORE, the NMP-binding domain and the LID domain (Fig. 12A). In addition to the classic NMPK domains, UCK1 has a fourth domain, which folds to a long β-hairpin and aligns with one edge of the β-sheet. The CORE domain is an α/β domain with a central, five stranded parallel β-sheet with two helices on each side. The CORE domain comprises the phosphate binding loop, or P-loop (**G_{30}**GTAS**GKS_{37}**), which is situated in the junction between the first strand and the first helix of the enzyme. Two helical domains cover the CORE at the carboxy end of the central β-sheet. One of these domains is the NMP-binding domain, which includes one α-helix with surrounding loops. The second domain that covers the CORE is the LID domain consisting of two long helices joined by a loop making an antiparallel helix pair. Together, these domains make up the active site as described by Suzuki *et al.* (Suzuki*, et al.*, 2004). The acceptor is recognized by the residues located in the NMP-binding domain, the β-hairpin and the LID domain.

4.2.2 Quaternary structure

Uridine cytidine kinase is a tetramer with 222 fold symmetry (Fig. 12B). There are two main types of subunit-subunit interaction. One is formed between helix G and strand 5 of the neighboring residues (Fig. 12C). The strands are aligned in an antiparallel way so that the two five-stranded β-sheets of bordering subunits form one ten-stranded β-sheet within the tetramer. The helices are not aligned to make an antiparallel helix-pair as is the case in thymidine kinases, rather they are tilted relative to each other and the subunit interaction is located at the C-termini of the helices. The second interaction is between the helices (helix E and helix F) in the LID domain. The interacting helices form a four helix bundle attached by hydrophobic and hydrophilic interactions (Fig. 12D).

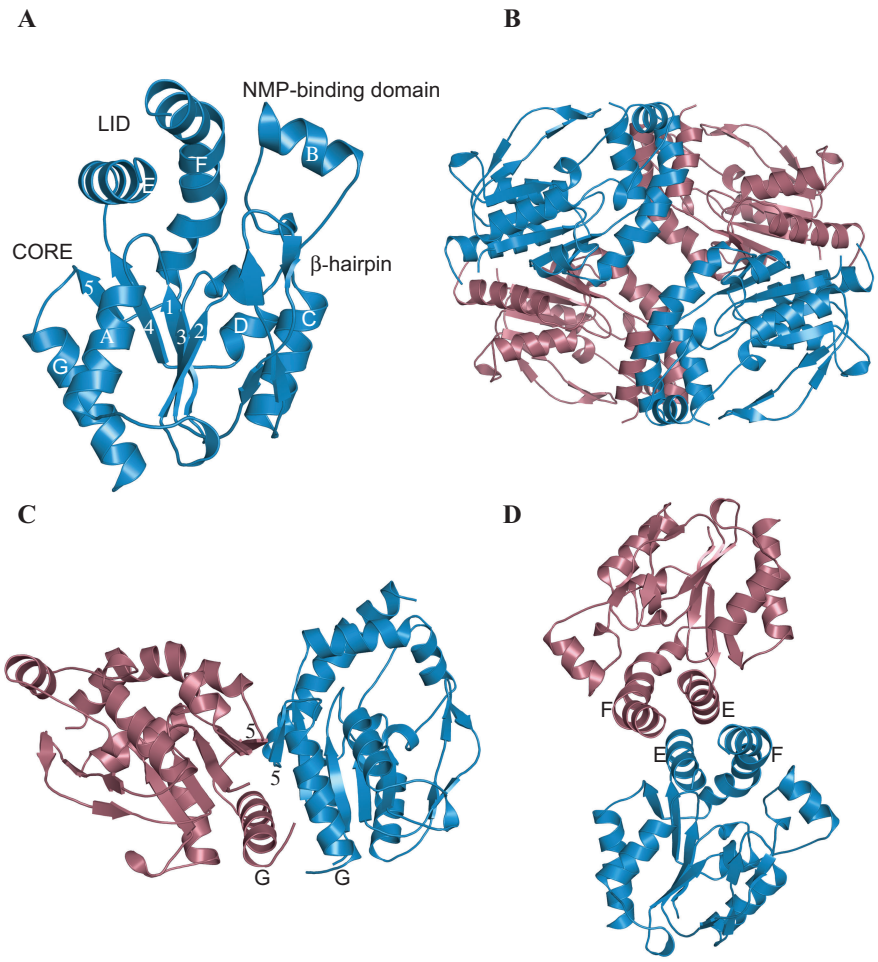


Fig. 12. The structure of UCK1. **A)** The monomer of UCK1. **B)** The tetramer. **C)** The subunit- subunit interaction formed between strand 5 and helix G of each subunit. **D)** The four helix bundle subunit interaction.

4.2.3 Ligand binding

Although UCK1 was crystallized in presence of CMP there is only minor electron density corresponding to the ribose and phosphate moieties of the ligand. The density corresponding to the ligand is weaker in the crystals that were allowed to stay longer in the mother liquor before harvesting. Comparable observation was made by Appleby *et al.* (Appleby et al., 2005) where they noticed decreased electron density for ADP in UCK2 crystals in complex with ADP and CMP that were kept a long time in mother liquor prior harvesting. In the ADP-soaked complex of UCK1, the CMP is fully absent.

ADP binds between the LID and the CORE domains at the strand-5-edge of the β-sheet. Compared to the phosphates, which are buried under the residues from helix E of the LID domain, the base and ribose are exposed. The ADP adenine ring makes stacking interactions with Arg168, and the exocyclic amino group (N6) of the adenine is hydrogen-bonded to the backbone carbonyl oxygen of Asp215. In UCK1 the adenine ring is bound in *syn* conformation, which is in contrast to how the adenine ring of ATP is bound in UCK2. The phosphates of the ADP form more interactions with the enzyme than the base and the ribose. Hydrogen bonds are formed between the phosphates and residues within the P-loop. The interactions are mainly with main-chain atoms, but Lys36 and Ser37 interact with their side chains.

The internal flexibility of UCK1 is dependent on whether the enzyme is in ligand-free or ligand-bound state. High temperature factors of the residues situated in the acceptor binding domains are a result of a partly occupied substrate pocket. The work of Suzuki *et al.* (2004) shows that these regions become stabilized upon substrate binding. In addition to the β-hairpin, the NMP-binding domain and the LID domain, the B values are high for the residues situated on the loop between helix A and strand 2 and the loop connecting strand 5 with helix G. The last loop becomes less flexible when the phosphate donor binds to the enzyme. This is reflected by lower B factors for the residues in this region in the ADP complex compared with the ligand-free UCK1.

4.3 Comparison with other enzymes

4.3.1 Amino acid comparison

The amino acid sequences of human UCK1 and UCK2 share 72% identity (Fig. 13). As described earlier, uridine cytidine kinase is a very specific enzyme when it comes to the choice of the phosphate accepting substrate, and only uridine and cytidine are efficiently phosphorylated. On the basis of the structure of human UCK2 several residues were proposed to play a role as phosphate acceptor determinants (Suzuki*, et al.*, 2004). A tyrosine, a histidine and an arginine (Tyr115, His120 and Arg178 using the human UCK1 numbering) bind the base of the acceptor. Depending on whether the base is a cytosine or a uracil, the tyrosine and the histidine swap between acting as hydrogen bond donor and acceptor. The pocket for the base is narrow and could not accommodate any purine bases. The preference for ribonucleoside over deoxyribonucleoside is assigned to an aspartic acid and an arginine (Asp87 and Arg169 in UCK1) which hydrogen bond with the 2'-and 3'-hydroxyl groups of the ribose moiety. The catalytic base has been assigned to Asp65. These residues are conserved in UCK1 and UCK2.

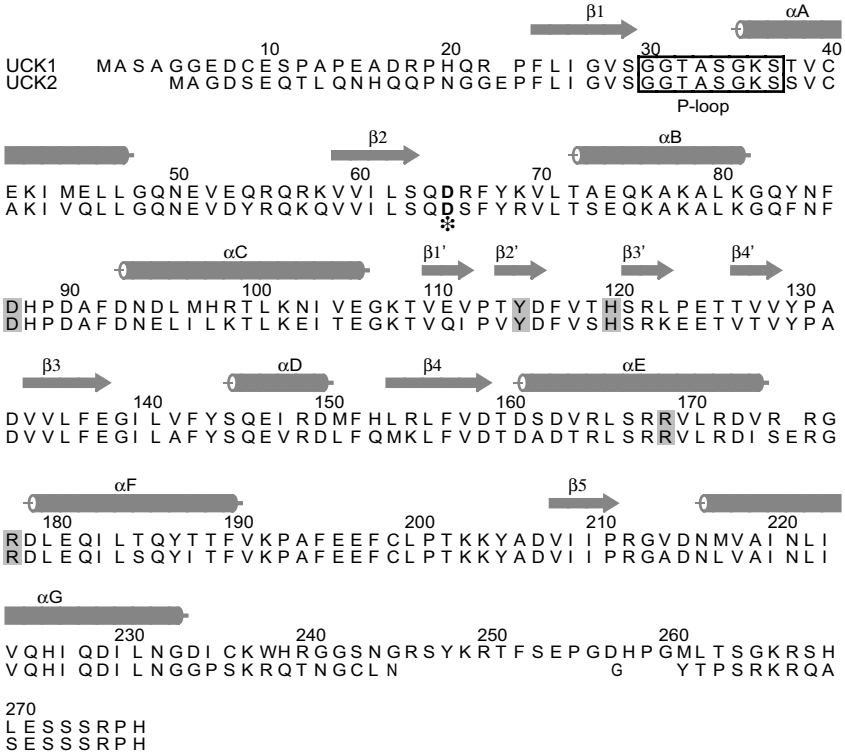


Fig. 13. Sequence alignment of human UCK1 (Q9HA47) and UCK2 (Q9BZX2). The secondary structure elements and the sequence numbering correspond to UCK1. Helix αB corresponds to the NMP-binding domain, the strands indicated with an apostrophe make up the β-hairpin motif, and helices αE and αF make up the LID domain. The P-loop motif is boxed, the catalytic base is indicated with a star, and the residues involved in acceptor binding are shown against gray background. The alignment was made using ClustalW (Chenna *et al.*, 2003) and Alscript (Barton 1993).

4.3.2 Structural comparison

4.3.2.1 UCK1 compared with UCK2

The structures of human UCK1 and UCK2 are almost identical. This is not surprising when considering the high sequence identity between the two enzymes. It has been reported that UCK2 forms two types of tetramers in the crystals (Suzuki*, et al.*, 2004). The ligand-free and the product-bound (ADP- and CMP-bound) UCK2 form a symmetric tetramer with 222 fold symmetry. On the other

hand, the tetramers of the inhibited (UTP- or CTP-bound) and substrate bound (Cyd-bound) UCK2 are distorted and less symmetrical. The biological relevance, if any, of the asymmetric tetramer is not known. UCK1 assembles into a symmetric tetramer similar to the one observed for the apo and product bound UCK2.

Two main conformations of the UCK2 monomer have been described (Suzuki*, et al.*, 2004). Depending on whether the phosphate acceptor site is occupied or empty the enzyme adopts a closed or an open conformation, respectively. The structure of UCK1 described in this thesis shows the enzyme in a semi-closed form (Fig. 14). The LID, the NMP-binding domain and the top of the β-hairpin are not as close to each other as in the UCK2 closed conformation, and they are not as separated as in the open conformation. Suzuki *et al.* proposed that the phosphate acceptor might bind in an ordered manner, where the ribose moiety binds first and induces formation of the base binding pocket. The semi-closed conformation of the UCK1 may be a representation of the sequential binding of the substrate such that the sugar moiety is bound but the base binding pocket is not yet formed, and therefore the cytosine and the ligating residues are flexible.

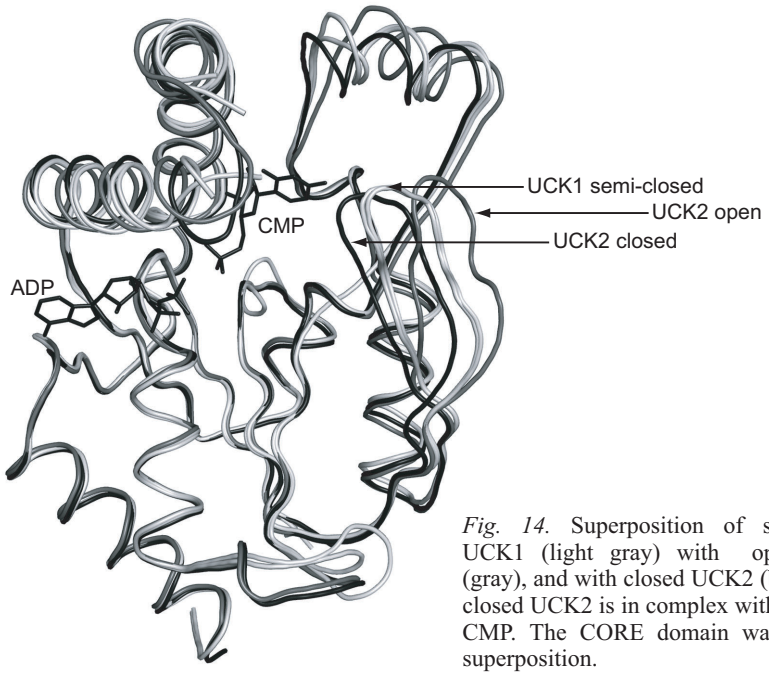


Fig. 14. Superposition of semi-closed UCK1 (light gray) with open UCK2 (gray), and with closed UCK2 (black). The closed UCK2 is in complex with ADP and CMP. The CORE domain was used for superposition.

4.3.2.2 Comparison of a ribonucleoside kinase with a deoxyribonucleoside kinase

Deoxyribonucleoside kinase (with exception of TK1-like enzymes) and ribonucleoside kinases are members of the NMPK fold family. Human dGK and UCK1 may be superimposed with an r.m.s.d. of 1.9 Å for 94 Cα atoms. The

CORE and LID domains are well conserved in these enzymes, but the β-hairpin and the NMP-binding domains of UCK are in dGK substituted by two helix-rich domains (Fig. 15). A further difference between the NKs and dNKs is the acceptor binding. UCK binds the base of the acceptor with residues situated on the β-hairpin, the NMP-binding domain and the LID-domain, whereas in dGK the substrate is bound by residues situated on the helical domain, while the arginine rich LID domain binds the phosphates of the donor or feedback inhibitor (Johansson*, et al.*, 2001). Unlike the tetrameric UCK, dGK is a dimer with the dimer interface formed between two long α-helices which correspond to the β-hairpin in UCK.

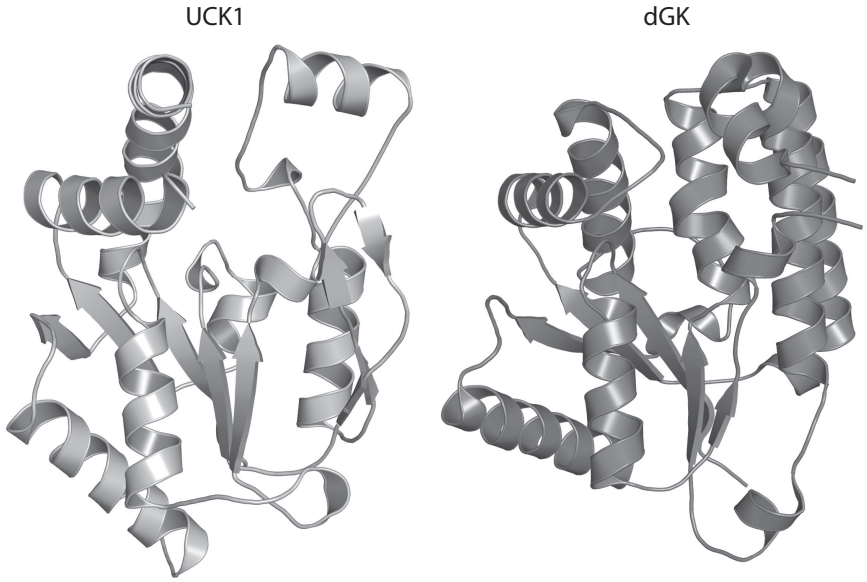


Fig. 15. Cartoon representation of UCK1 and dGK monomers.

What are the structural bases for UCK being a strict ribonucleoside kinase and dGK and TK for having a strong preference for deoxyribonucleosides? Both in dNKs and in TK1-like enzymes, the 3'-position is hydrogen-bonded to the enzyme. In dNKs there are two conserved amino acids, a tyrosine and a glutamic acid, that form hydrogen bonds between the side chains and the 3'-hydroxyl group. In the TK1-like enzymes the interaction is with the main chain atoms of a glycine situated on the lasso loop and a aspartic acid situated just at the beginning of the P-β-hairpin (Johansson*, et al.*, 2001; Sabini*, et al.*, 2003; Welin *et al.*, 2004). The 2'-position is in both enzyme families crowded with hydrophobic residues and a 2'-hydroxyl group would cause steric interference and hinder proper binding. In NKs however, there is an aspartic acid and an asparagine that form hydrogen bonds with 2'- and 3'-hydroxyl groups (Schumacher *et al.*, 2000; Suzuki*, et al.*, 2004). According to Suzuki *et al.* the binding tothe sugar moiety is crucial for the formation of the base recognition site (Suzuki*, et al.*, 2004).

5. Conclusions and future perspectives

5.1 Thymidine kinase

The initial goal of this doctoral work was to determine the 3D-strucutre of TK from the opportunistic pathogen *Ureaplasma urealyticum*. *U. urealyticum* lacks the ability of *de novo* synthesis of dNTPs, and has to rely on its salvage enzymes, *Uu*-TK and dAK, for the synthesis of DNA precursors. Therefore, *Uu*-TK was evaluated as a potential prodrug activator for future antibacterial therapies. Moreover, to learn more about TKs, the structures of TKs of two bacillus species were determined.

Based on the results described in this thesis we can divide dNKs into two structurally separate groups, where TK1-like enzymes create a group of their own. The only structural and sequential similarity between TK1-like enzymes and other dNKs is the P-loop situated on a junction between a β-strand and an α-helix. The active site of TK1-like enzymes is situated in a hydrophobic cleft between a rigid α/β domain and a flexible lasso domain. No major differences in the active sites of the bacterial TKs and hTK1 have been observed, which limits the future design of selective antibacterial agents. However, the structures described in this thesis provide interesting information regarding conformational changes during the time of reaction. The structures of the TKs where neither acceptor nor feedback inhibitor is bound show an open active site with a flexible lasso-domain. The lasso becomes stabilized upon substrate binding and closes down over the active site. A highly intriguing finding was the conformational change of the quaternary structure induced by the binding of dTTP at the phosphate donor site. A less tight tetramer is formed in the presence of a phosphate donor than when the donor site is unoccupied.

A glutamic acid has been proposed to act as a catalytic base, but the complete picture of the reaction mechanism is still to be elucidated. For this purpose a productive complex is desirable. Complexes such as: TK with both substrates (a non-hydrolyzable ATP analog and dT) and Mg-ion, TK with both products (ADP and dTMP), bisubstrate analogs such as AP_4dT or AP_5dT with Mg-ion, and TK with Mg-ion, ADP, dT and suitable transition state mimicking compounds *e.g.* AlF_4^-, BeF_3^-. Moreover, the oligomeric structure of TK1-like enzymes needs further investigation. Hopefully, mutational studies, and native molecular mass determinations with ultracentrifugation and dynamic light scattering analysis, will clarify the matter of the oligomeric state of TKs.

5.2 Uridine-cytidine kinase 1

UCK1 is sequentially as well as structurally highly similar to UCK2. These enzymes have a tissue specific expression pattern and show different enzymatic properties. Structure determination of the UCK1 enzyme was conducted with the expectation to provide some possible explanations for those differences. In order to study the active site, the UCK1 was co-crystallized with CMP and soaked with

the second product ADP. While the structures of UCK2 show two major conformations of the enzyme (one open and one closed), the UCK1 was trapped in a semi-closed conformation. This form may support the hypothesis of a sequential substrate binding mode. The ADP binds in a similar way as to UCK2 with the exception for the adenine base which is in *syn* conformation in UCK1. There are not many interactions between the ring and the sugar of the ADP and the enzyme, which confirms the low specificity for the nucleotide as a phosphate donor. Unfortunately, the CMP did not bind strongly enough to stabilize the acceptor biding site. Many of the residues, which in UCK2 were identified as acceptor binding, are flexible in UCK1. More complex studies are therefore needed in order to explain the reasons for the diverse substrate specificities of UCK1 and UCK2.

6. References

Anand, K., Pal, D. & Hilgenfeld, R. 2002. An overview on 2-methyl-2,4-pentanediol in crystallization and in crystals of biological macromolecules. *Acta Crystallogr D Biol Crystallogr 58*, 1722-8.

Andersen, R.B. & Neuhard, J. 2001. Deoxynucleoside kinases encoded by the yaaG and yaaF genes of *Bacillus subtilis*. Substrate specificity and kinetic analysis of deoxyguanosine kinase with UTP as the preferred phosphate donor. *J Biol Chem 276*, 5518-24.

Appleby, T.C., Larson, G., Cheney, I.W., Walker, H., Wu, J.Z., Zhong, W., Hong, Z. & Yao, N. 2005. Structure of human uridine-cytidine kinase 2 determined by SIRAS using a rotating-anode X-ray generator and a single samarium derivative. *Acta Crystallogr D Biol Crystallogr 61*, 278-84.

Arnér, E.S. & Eriksson, S. 1995. Mammalian deoxyribonucleoside kinases. *Pharmacol Ther 67*, 155-86.

Baillie, L. & Read, T.D. 2001. *Bacillus anthracis*, a bug with attitude! *Curr Opin Microbiol 4*, 78-81.

Bauer, C.B., Fonseca, M.V., Holden, H.M., Thoden, J.B., Thompson, T.B., Escalante-Semerena, J.C. & Rayment, I. 2001. Three-dimensional structure of ATP:corrinoid adenosyltransferase from *Salmonella typhimurium* in its free state, complexed with MgATP, or complexed with hydroxycobalamin and MgATP. *Biochemistry 40*, 361-74.

Barton, G.J. 1993. ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng 6*, 37-40.

Bello, L.J. 1974. Regulation of thymidine kinase synthesis in human cells. *Exp Cell Res 89*, 263-74.

Berenstein, D., Christensen, J.F., Kristensen, T., Hofbauer, R. & Munch-Petersen, B. 2000. Valine, not methionine, is amino acid 106 in human cytosolic thymidine kinase (TK1). Impact on oligomerization, stability, and kinetic properties. *J Biol Chem 275*, 32187-92.

Bergfors, T. 2003. Seeds to crystals. *J Struct Biol 142*, 66-76.

Berry, M.B. & Phillips, G.N., Jr. 1998. Crystal structures of *Bacillus stearothermophilus* adenylate kinase with bound Ap5A, Mg2+ Ap5A, and Mn2+ Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg2+ and Mn2+. *Proteins 32*, 276-88.

Birringer, M.S., Claus, M.T., Folkers, G., Kloer, D.P., Schulz, G.E. & Scapozza, L. 2005. Structure of a type II thymidine kinase with bound dTTP. *FEBS Lett 579*, 1376-82.

Birringer, M.S., Perozzo, R., Kut, E., Stillhart, C., Surber, W., Scapozza, L. & Folkers, G. 2006. High-level expression and purification of human thymidine kinase 1: Quaternary structure, stability, and kinetics. *Protein Expr Purif 47*, 506-15.

Black, M.E. & Hruby, D.E. 1990a. Identification of the ATP-binding domain of vaccinia virus thymidine kinase. *J Biol Chem 265*, 17584-92.

Black, M.E. & Hruby, D.E. 1990b. Quaternary structure of vaccinia virus thymidine kinase. *Biochem Biophys Res Commun 169*, 1080-6.

Black, M.E. & Hruby, D.E. 1992. Site-directed mutagenesis of a conserved domain in vaccinia virus thymidine kinase. Evidence for a potential role in magnesium binding. *J Biol Chem 267*, 6801-6.

Bricogne, G., Vonrhein, C., Flensburg, C., Schiltz, M. & Paciorek, W. 2003. Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. *Acta Crystallogr D Biol Crystallogr 59*, 2023-30.

Briozzo, P., Evrin, C., Meyer, P., Assairi, L., Joly, N., Barzu, O. & Gilles, A.M. 2005. Structure of *Escherichia coli* UMP kinase differs from that of other nucleoside monophosphate kinases and sheds new light on enzyme regulation. *J Biol Chem 280*, 25533-40.

Brown, D.G., Visse, R., Sandhu, G., Davies, A., Rizkallah, P.J., Melitz, C., Summers, W.C. & Sanderson, M.R. 1995. Crystal structures of the thymidine kinase from herpes simplex virus type-1 in complex with deoxythymidine and ganciclovir. *Nat Struct Biol 2*, 876-81.

Carnrot, C., Wehelie, R., Eriksson, S., Bolske, G. & Wang, L. 2003. Molecular characterization of thymidine kinase from *Ureaplasma urealyticum*: nucleoside analogues as potent inhibitors of mycoplasma growth. *Mol Microbiol 50*, 771-80.

Carnrot, C., Vogel, S.R., Byun, Y., Wang, L., Tjarks, W., Eriksson, S. & Phipps, A.J. 2006. Evaluation of *Bacillus anthracis* thymidine kinase as a potential target for the development of antibacterial nucleoside analogs. *Biol Chem 387*, 1575-81.

Chabes, A. & Thelander, L. 2000. Controlled protein degradation regulates ribonucleotide reductase activity in proliferating mammalian cells during the normal cell cycle and in response to DNA damage and replication blocks. *J Biol Chem 275*, 17747-53.

Chabes, A.L., Pfleger, C.M., Kirschner, M.W. & Thelander, L. 2003. Mouse ribonucleotide reductase R2 protein: a new target for anaphase-promoting complex-Cdh1-mediated proteolysis. *Proc Natl Acad Sci U S A 100*, 3925-9.

Champness, J.N., Bennett, M.S., Wien, F., Visse, R., Summers, W.C., Herdewijn, P., de Clerq, E., Ostrowski, T., Jarvest, R.L. & Sanderson, M.R. 1998. Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands. *Proteins 32*, 350-61.

Chen, M.S., Summers, W.P., Walker, J., Summers, W.C. & Prusoff, W.H. 1979. Characterization of pyrimidine deoxyribonucleoside kinase (thymidine kinase) and thymidylate kinase as a multifunctional enzyme in cells transformed by herpes simplex virus type 1 and in cells infected with mutant strains of herpes simplex virus. *J Virol 30*, 942-5.

Chen, M.S., Walker, J. & Prusoff, W.H. 1979. Kinetic studies of herpes simplex virus type 1-encoded thymidine and thymidylate kinase, a multifunctional enzyme. *J Biol Chem 254*, 10747-53.

Cheng, N., Payne, R.C. & Traut, T.W. 1986. Regulation of uridine kinase. Evidence for a regulatory site. *J Biol Chem 261*, 13006-12.

Barton, G.J. 1993. ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng 6*, 37-40.

Chenna, R., Sugawara, Hideaki, Koike,Tadashi, Lopez, Rodrigo, Gibson, Toby J, Higgins, Desmond G, Thompson, Julie D 2003. Multiple sequence alignment with the Clustal series of programs. *Nucleic Acids Res 31*, 3497-500

Cho, Y., Gorina, S., Jeffrey, P.D. & Pavletich, N.P. 1994. Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science 265*, 346-55.

Coppock, D.L. & Pardee, A.B. 1987. Control of thymidine kinase mRNA during the cell cycle. *Mol Cell Biol 7*, 2925-32.

De Clercq, E. 2001. Antiviral drugs: current state of the art. *J Clin Virol 22*, 73-89.

de La Fortelle, E. & Irwin, J. 1997. SHARP: A Maximum-Likelihood Heavy-Atom Parameter Refinement Program for the MIR and MAD Methods. *Methods Enzymol 276*, 472-494.

Dixon, T.C., Meselson, M., Guillemin, J. & Hanna, P.C. 1999. Anthrax. *N Engl J Med 341*, 815-26.

Ducruix, A. & Giegé, R. 1992. *Crysrallization of Nucleic Acids and Proteins: A Practicle Approach*. Oxford Univ. Press, London. pp.

El Omari, K., Solaroli, N., Karlsson, A., Balzarini, J. & Stammers, D.K. 2006. Structure of vaccinia virus thymidine kinase in complex with dTTP: insights for drug design. *BMC Struct Biol 6*, 22.

Eriksson, S., Kierdaszuk, B., Munch-Petersen, B., Öberg, B. & Johansson, N.G. 1991. Comparison of the substrate specificities of human thymidine kinase 1 and 2 and deoxycytidine kinase toward antiviral and cytostatic nucleoside analogs. *Biochem Biophys Res Commun 176*, 586-92.

Eriksson, S., Munch-Petersen, B., Johansson, K. & Eklund, H. 2002. Structure and function of cellular deoxyribonucleoside kinases. *Cell Mol Life Sci 59*, 1327-46.

Galmarini, C.M., Mackey, J.R. & Dumontet, C. 2002. Nucleoside analogues and nucleobases in cancer treatment. *Lancet Oncol 3*, 415-24.

Gentry, G.A. 1992. Viral thymidine kinases and their relatives. *Pharmacol Ther 54*, 319-55.

Gilles, A.M., Glaser, P., Perrier, V., Meier, A., Longin, R., Sebald, M., Maignan, L., Pistotnik, E. & Barzu, O. 1994. Zinc, a structural component of adenylate kinases from gram-positive bacteria. *J Bacteriol 176*, 520-3.

Gilles, A.M., Presecan, E., Vonica, A. & Lascu, I. 1991. Nucleoside diphosphate kinase from human erythrocytes. Structural characterization of the two polypeptide chains responsible for heterogeneity of the hexameric enzyme. *J Biol Chem 266*, 8784-9.

Glass, J.I., Lefkowitz, E.J., Glass, J.S., Heiner, C.R., Chen, E.Y. & Cassell, G.H. 2000. The complete sequence of the mucosal pathogen *Ureaplasma urealyticum*. *Nature 407*, 757-62.

Gronowitz, J.S., Kallander, F.R., Diderholm, H., Hagberg, H. & Pettersson, U. 1984. Application of an in vitro assay for serum thymidine kinase: results on viral disease and malignancies in humans. *Int J Cancer 33*, 5-12.

Guittet, O., Håkansson, P., Voevodskaya, N., Fridd, S., Graslund, A., Arakawa, H., Nakamura, Y. & Thelander, L. 2001. Mammalian p53R2 protein forms an active ribonucleotide reductase in vitro with the R1 protein, which is expressed both in resting cells in response to DNA damage and in proliferating cells. *J Biol Chem 276*, 40647-51.

Härd, T., Kellenbach, E., Boelens, R., Maler, B.A., Dahlman, K., Freedman, L.P., Carlstedt-Duke, J., Yamamoto, K.R., Gustafsson, J.Å. & Kaptein, R. 1990.

Solution structure of the glucocorticoid receptor DNA-binding domain. *Science 249*, 157-60.
Harlow, K.W., Nygaard, P. & Hove-Jensen, B. 1995. Cloning and characterization of the gsk gene encoding guanosine kinase of *Escherichia coli*. *J Bacteriol 177*, 2236-40.
He, Q., Zhang, P., Zou, L., Li, H., Wang, X., Zhou, S., Fornander, T. & Skog, S. 2005. Concentration of thymidine kinase 1 in serum (S-TK1) is a more sensitive proliferation marker in human solid tumors than its activity. *Oncol Rep 14*, 1013-9.
He, Q., Zou, L., Zhang, P.A., Lui, J.X., Skog, S. & Fornander, T. 2000. The clinical significance of thymidine kinase 1 measurement in serum of breast cancer patients using anti-TK1 antibody. *Int J Biol Markers 15*, 139-46.
Heidelberger, C., Chaudhuri, N.K., Danneberg, P., Mooren, D., Griesbach, L., Duschinsky, R., Schnitzer, R.J., Pleven, E. & Scheiner, J. 1957. Fluorinated pyrimidines, a new class of tumour-inhibitory compounds. *Nature 179*, 663-6.
Herzfeld, A. & Raper, S.M. 1979. Uridine kinase activities in developing, adult and neoplastic rat tissues. *Biochem J 182*, 771-8.
Holm, L. & Sander, C. 1993. Protein structure comparison by alignment of distance matrices. *J Mol Biol 233*, 123-38.
Hruby, D.E. & Ball, L.A. 1982. Mapping and identification of the vaccinia virus thymidine kinase gene. *J Virol 43*, 403-9.
Huang, S.H., Tang, A., Drisco, B., Zhang, S.Q., Seeger, R., Li, C. & Jong, A. 1994. Human dTMP kinase: gene expression and enzymatic activity coinciding with cell cycle progression and cell growth. *DNA Cell Biol 13*, 461-71.
Ichikawa, W. 2006. Prediction of clinical outcome of fluoropyrimidine-based chemotherapy for gastric cancer patients, in terms of the 5-fluorouracil metabolic pathway. *Gastric Cancer 9*, 145-55.
Ives, D.H. & Ikeda, S. 1998. Life on the salvage path: the deoxynucleoside kinase of *Lactobacillus acidophilus* R-26. *Prog Nucleic Acid Res Mol Biol 59*, 205-55.
Ives, D.H., Morse, P.A., Jr. & Potter, V.R. 1963. Feedback inhibition of thymidine kinase by thymidine triphosphate. *J Biol Chem 238*, 1467-74.
Janin, J., Dumas, C., Morera, S., Xu, Y., Meyer, P., Chiadmi, M. & Cherfils, J. 2000. Three-dimensional structure of nucleoside diphosphate kinase. *J Bioenerg Biomembr 32*, 215-25.
Jensen, G.B., Hansen, B.M., Eilenberg, J. & Mahillon, J. 2003. The hidden lifestyles of *Bacillus cereus* and relatives. *Environ Microbiol 5*, 631-40.
Johansson, K., Ramaswamy, S., Ljungcrantz, C., Knecht, W., Piskur, J., Munch-Petersen, B., Eriksson, S. & Eklund, H. 2001. Structural basis for substrate specificities of cellular deoxyribonucleoside kinases. *Nat Struct Biol 8*, 616-20.
Kang, G.J., Cooney, D.A., Moyer, J.D., Kelley, J.A., Kim, H.Y., Marquez, V.E. & Johns, D.G. 1989. Cyclopentenylcytosine triphosphate. Formation and inhibition of CTP synthetase. *J Biol Chem 264*, 713-8.
Karlström, H.O. 1970. Inability of *Escherichia coli* B to incorporate added deoxycytidine, deoxyandenosine, and deoxyguanosine into DNA. *Eur J Biochem 17*, 68-71.
Katahira, R. & Ashihara, H. 2006. Profiles of purine biosynthesis, salvage and degradation in disks of potato (*Solanum tuberosum* L.) tubers. *Planta 225*, 115-26.

Kauffman, M.G. & Kelly, T.J. 1991. Cell cycle regulation of thymidine kinase: residues near the carboxyl terminus are essential for the specific degradation of the enzyme at mitosis. *Mol Cell Biol 11*, 2538-46.
Ke, P.Y. & Chang, Z.F. 2004. Mitotic degradation of human thymidine kinase 1 is dependent on the anaphase-promoting complex/cyclosome-CDH1-mediated pathway. *Mol Cell Biol 24*, 514-26.
Kimura, N., Shimada, N., Fukuda, M., Ishijima, Y., Miyazaki, H., Ishii, A., Takagi, Y. & Ishikawa, N. 2000. Regulation of cellular functions by nucleoside diphosphate kinases in mammals. *J Bioenerg Biomembr 32*, 309-15.
Knecht, W., Petersen, G.E., Munch-Petersen, B. & Piskur, J. 2002a. 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*, 529-40.
Knecht, W., Petersen, G.E., Sandrini, M.P., Søndergaard, L., Munch-Petersen, B. & Piskur, J. 2003. Mosquito has a single multisubstrate deoxyribonucleoside kinase characterized by unique substrate specificity. *Nucleic Acids Res 31*, 1665-72.
Knecht, W., Sandrini, M.P., Johansson, K., Eklund, H., Munch-Petersen, B. & Piskur, J. 2002b. A few amino acid substitutions can convert deoxyribonucleoside kinase specificity from pyrimidines to purines. *Embo J 21*, 1873-80.
Koizumi, K., Shimamoto, Y., Azuma, A., Wataya, Y., Matsuda, A., Sasaki, T. & Fukushima, M. 2001. Cloning and expression of uridine/cytidine kinase cDNA from human fibrosarcoma cells. *Int J Mol Med 8*, 273-8.
Kotiranta, A., Lounatmaa, K. & Haapasalo, M. 2000. Epidemiology and pathogenesis of *Bacillus cereus* infections. *Microbes Infect 2*, 189-98.
Kuzin, A.P., Abashidze, M., Forouhar, F., Vorobiev, S. M., Acton, T. B., Ma, L.-C., Xiao, R., Montelione, G. T., Tong, L., Hunt, J. F. 2004. X-Ray Structure of *Clostridium Acetobutylicum* Thymidine Kinase with ADP. *Northeast Structural Genomics Target Car26. PDB code: 1XX6*
Lacombe, M.L., Milon, L., Munier, A., Mehus, J.G. & Lambeth, D.O. 2000. The human Nm23/nucleoside diphosphate kinases. *J Bioenerg Biomembr 32*, 247-58.
Li, H.X., Lei, D.S., Wang, X.Q., Skog, S. & He, Q. 2005. Serum thymidine kinase 1 is a prognostic and monitoring factor in patients with non-small cell lung cancer. *Oncol Rep 13*, 145-9.
Liu, Q.Y. & Summers, W.C. 1988. Site-directed mutagenesis of a nucleotide-binding domain in HSV-1 thymidine kinase: effects on catalytic activity. *Virology 163*, 638-42.
Mikkelsen, N.E., Johansson, K., Karlsson, A., Knecht, W., Andersen, G., Piskur, J., Munch-Petersen, B. & Eklund, H. 2003. Structural basis for feedback inhibition of the deoxyribonucleoside salvage pathway: studies of the Drosophila deoxyribonucleoside kinase. *Biochemistry 42*, 5706-12.
Miller, W.H. & Miller, R.L. 1991. Guanosine kinase from *Trichomonas vaginalis*. *Mol Biochem Parasitol 48*, 39-46.
Møllgaard, H. 1980. Deoxyadenosine/deoxycytidine kinase from Bacillus subtilis. Purification, characterization, and physiological function. *J Biol Chem 255*, 8216-20.

Munch-Petersen, B., Cloos, L., Tyrsted, G. & Eriksson, S. 1991. Diverging substrate specificity of pure human thymidine kinases 1 and 2 against antiviral dideoxynucleosides. *J Biol Chem 266*, 9032-8.

Munch-Petersen, B., Piskur, J. & Søndergaard, L. 1998. Four deoxynucleoside kinase activities from *Drosophila melanogaster* are contained within a single monomeric enzyme, a new multifunctional deoxynucleoside kinase. *J Biol Chem 273*, 3926-31.

Munch-Petersen, B., Tyrsted, G. & Cloos, L. 1993. Reversible ATP-dependent transition between two forms of human cytosolic thymidine kinase with different enzymatic properties. *J Biol Chem 268*, 15621-5.

Murata, D., Endo, Y., Obata, T., Sakamoto, K., Syouji, Y., Kadohira, M., Matsuda, A. & Sasaki, T. 2004. A crucial role of uridine/cytidine kinase 2 in antitumor activity of 3'-ethynyl nucleosides. *Drug Metab Dispos 32*, 1178-82.

Niculescu-Duvaz, I. & Springer, C.J. 2005. Introduction to the background, principles, and state of the art in suicide gene therapy. *Mol Biotechnol 30*, 71-88.

Nordlund, P. & Reichard, P. 2006. Ribonucleotide reductases. *Annu Rev Biochem 75*, 681-706.

Okinaka, R., Cloud, K., Hampton, O., Hoffmaster, A., Hill, K., Keim, P., Koehler, T., Lamke, G., Kumano, S., Manter, D., Martinez, Y., Ricke, D., Svensson, R. & Jackson, P. 1999. Sequence, assembly and analysis of pX01 and pX02. *J Appl Microbiol 87*, 261-2.

Painter, J. & Merritt, E.A. 2006. Optimal description of a protein structure in terms of multiple groups undergoing TLS motion. *Acta Crystallogr D Biol Crystallogr 62*, 439-50.

Plagemann, P.G., Wohlhueter, R.M. & Woffendin, C. 1988. Nucleoside and nucleobase transport in animal cells. *Biochim Biophys Acta 947*, 405-43.

Pollack, J.D., Myers, M.A., Dandekar, T. & Herrmann, R. 2002. Suspected utility of enzymes with multiple activities in the small genome Mycoplasma species: the replacement of the missing "household" nucleoside diphosphate kinase gene and activity by glycolytic kinases. *Omics 6*, 247-58.

Rasko, D.A., Altherr, M.R., Han, C.S. & Ravel, J. 2005. Genomics of the *Bacillus cereus* group of organisms. *FEMS Microbiol Rev 29*, 303-29.

Reichard, P. 1988. Interactions between deoxyribonucleotide and DNA synthesis. *Annu Rev Biochem 57*, 349-74.

Reinstein, J., Brune, M. & Wittinghofer, A. 1988. Mutations in the nucleotide binding loop of adenylate kinase of *Escherichia coli*. *Biochemistry 27*, 4712-20.

Reinstein, J., Schlichting, I. & Wittinghofer, A. 1990. Structurally and catalytically important residues in the phosphate binding loop of adenylate kinase of *Escherichia coli*. *Biochemistry 29*, 7451-9.

Robertson, J.A., Stemke, G.W., Davis, J.W., Jr., Harasawa, R., Thirkell, D., Kong, F., Shepard, M.C. & Ford, D.K. 2002. Proposal of Ureaplasma parvum sp. nov. and emended description of *Ureaplasma urealyticum* (Shepard et al. 1974) Robertson et al. 2001. *Int J Syst Evol Microbiol 52*, 587-97.

Ropp, P.A. & Traut, T.W. 1996. Cloning and expression of a cDNA encoding uridine kinase from mouse brain. *Arch Biochem Biophys 336*, 105-12.

Ropp, P.A. & Traut, T.W. 1998. Uridine kinase: altered enzyme with decreased affinities for uridine and CTP. *Arch Biochem Biophys 359*, 63-8.

Sabini, E., Ort, S., Monnerjahn, C., Konrad, M. & Lavie, A. 2003. Structure of human dCK suggests strategies to improve anticancer and antiviral therapy. *Nat Struct Biol 10*, 513-9.
Saito, H. & Tomioka, H. 1984. Thymidine kinase of bacteria: activity of the enzyme in actinomycetes and related organisms. *J Gen Microbiol 130*, 1863-70.
Saito, H., Tomioka, H. & Ohkido, S. 1985. Further studies on thymidine kinase: distribution pattern of the enzyme in bacteria. *J Gen Microbiol 131*, 3091-8.
Sandrini, M.P., Clausen, A.R., Munch-Petersen, B. & Piskur, J. 2006. Thymidine kinase diversity in bacteria. *Nucleosides Nucleotides Nucleic Acids 25*, 1153-8.
Saraste, M., Sibbald, P.R. & Wittinghofer, A. 1990. The P-loop--a common motif in ATP- and GTP-binding proteins. *Trends Biochem Sci 15*, 430-4.
Schneider, T.R. & Sheldrick, G.M. 2002. Substructure solution with SHELXD. *Acta Crystallogr D Biol Crystallogr 58*, 1772-9.
Schumacher, M.A., Scott, D.M., Mathews, II, Ealick, S.E., Roos, D.S., Ullman, B. & Brennan, R.G. 2000. Crystal structures of *Toxoplasma gondii* adenosine kinase reveal a novel catalytic mechanism and prodrug binding. *J Mol Biol 298*, 875-93.
Sherley, J.L. & Kelly, T.J. 1988a. Human cytosolic thymidine kinase. Purification and physical characterization of the enzyme from HeLa cells. *J Biol Chem 263*, 375-82.
Sherley, J.L. & Kelly, T.J. 1988b. Regulation of human thymidine kinase during the cell cycle. *J Biol Chem 263*, 8350-8.
Sigal, I.S., Gibbs, J.B., D'Alonzo, J.S., Temeles, G.L., Wolanski, B.S., Socher, S.H. & Scolnick, E.M. 1986. Mutant ras-encoded proteins with altered nucleotide binding exert dominant biological effects. *Proc Natl Acad Sci U S A 83*, 952-6.
Singh, B., Hao, W., Wu, Z., Eigl, B. & Gupta, R.S. 1996. Cloning and characterization of cDNA for adenosine kinase from mammalian (Chinese hamster, mouse, human and rat) species. High frequency mutants of Chinese hamster ovary cells involve structural alterations in the gene. *Eur J Biochem 241*, 564-71.
Smith, D.G., Russell, W.C., Ingledew, W.J. & Thirkell, D. 1993. Hydrolysis of urea by *Ureaplasma urealyticum* generates a transmembrane potential with resultant ATP synthesis. *J Bacteriol 175*, 3253-8.
Spychala, J., Datta, N.S., Takabayashi, K., Datta, M., Fox, I.H., Gribbin, T. & Mitchell, B.S. 1996. Cloning of human adenosine kinase cDNA: sequence similarity to microbial ribokinases and fructokinases. *Proc Natl Acad Sci U S A 93*, 1232-7.
Story, R.M., Weber, I.T. & Steitz, T.A. 1992. The structure of the *E. coli* recA protein monomer and polymer. *Nature 355*, 318-25.
Stura, E. & Wilson, I. 1990. Analytical and Production Seeding Techniques. *Methods 1*, 38-49.
Suzuki, N.N., Koizumi, K., Fukushima, M., Matsuda, A. & Inagaki, F. 2004. Structural basis for the specificity, catalysis, and regulation of human uridine-cytidine kinase. *Structure 12*, 751-64.
Tabata, S., Tanaka, M., Endo, Y., Obata, T., Matsuda, A. & Sasaki, T. 1997. Anti-tumor mechanisms of 3'-ethynyluridine and 3'-ethynylcytidine as RNA synthesis inhibitors: development and characterization of 3'-ethynyluridine-resistant cells. *Cancer Lett 116*, 225-31.

Tian, G.C., Yan, H.G., Jiang, R.T., Kishi, F., Nakazawa, A. & Tsai, M.D. 1990. Mechanism of adenylate kinase. Are the essential lysines essential? *Biochemistry 29*, 4296-304.
Traut, T.W. 1994. Physiological concentrations of purines and pyrimidines. *Mol Cell Biochem 140*, 1-22.
Waites, K.B., Katz, B. & Schelonka, R.L. 2005. Mycoplasmas and ureaplasmas as neonatal pathogens. *Clin Microbiol Rev 18*, 757-89.
Van Rompay, A.R., Johansson, M. & Karlsson, A. 2000. Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. *Pharmacol Ther 87*, 189-98.
Van Rompay, A.R., Norda, A., Linden, K., Johansson, M. & Karlsson, A. 2001. Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases. *Mol Pharmacol 59*, 1181-6.
Wang, L., Westberg, J., Bolske, G. & Eriksson, S. 2001. Novel deoxynucleoside-phosphorylating enzymes in mycoplasmas: evidence for efficient utilization of deoxynucleosides. *Mol Microbiol 42*, 1065-73.
Welin, M., Kosinska, U., Mikkelsen, N.E., Carnrot, C., Zhu, C., Wang, L., Eriksson, S., Munch-Petersen, B. & Eklund, H. 2004. Structures of thymidine kinase 1 of human and mycoplasmic origin. *Proc Natl Acad Sci U S A 101*, 17970-5.
Verschuur, A.C., Van Gennip, A.H., Leen, R., Muller, E.J., Elzinga, L., Voute, P.A. & Van Kuilenburg, A.B. 2000. Cyclopentenyl cytosine inhibits cytidine triphosphate synthetase in paediatric acute non-lymphocytic leukaemia: a promising target for chemotherapy. *Eur J Cancer 36*, 627-35.
Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. 1995. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett 368*, 289-92.
Wild, K., Bohner, T., Folkers, G. & Schulz, G.E. 1997. The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue. *Protein Sci 6*, 2097-106.
Winn, M.D., Isupov, M.N. & Murshudov, G.N. 2001. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr D Biol Crystallogr 57*, 122-33.
Xu, Y.W., Morera, S., Janin, J. & Cherfils, J. 1997. AlF3 mimics the transition state of protein phosphorylation in the crystal structure of nucleoside diphosphate kinase and MgADP. *Proc Natl Acad Sci U S A 94*, 3579-83.
Yan, H. & Tsai, M.D. 1999. Nucleoside monophosphate kinases: structure, mechanism, and substrate specificity. *Adv Enzymol Relat Areas Mol Biol 73*, 103-34, x.
Zhang, J., Jia, Q., Zou, S., Zhang, P., Zhang, X., Skog, S., Luo, P., Zhang, W. & He, Q. 2006. Thymidine kinase 1: a proliferation marker for determining prognosis and monitoring the surgical outcome of primary bladder carcinoma patients. *Oncol Rep 15*, 455-61.
Zou, L., Zhang, P.G., Zou, S., Li, Y. & He, Q. 2002. The half-life of thymidine kinase 1 in serum measured by ECL dot blot: a potential marker for monitoring the response to surgery of patients with gastric cancer. *Int J Biol Markers 17*, 135-40.

Acknowledgements

First and foremost I would like to thank my supervisors. **Hasse**, thank you for your optimism, patience and good discussions Thank you for letting me work so independently. Even if work sometimes felt frustrating and impossible, it was also very satisfying and I now feel prepared for the next step. Thank you **Fredrik** for your encouragement and for taking the car to the lab in the middle of a Friday night to turn of the cryo-generator.

I would also like to thank TK-collaborators: **Staffan Eriksson** and his crew (**Liya, Rahma, Cicci, Elena, Ashraf**) for interesting kinase-meetings, journal club and teaching us some enzyme kinetics. Special thanks goes to Cicci, for your never ending supply of fresh protein whenever I needed more, and for our small chats about mostly everything. It was enjoyable to collaborate with you. The "crazy" professor **Jure Piskur** and **Michael Sandrini** for taking good care of Martin and me when we visited your lab. Jure, Michael and **Anders**, thank you for good collaboration on the *Bc*-TK project.

My gratitude goes also to **SGC** at Karolinska Institutet for giving me the opportunity to work with you. Thank you all the members of the consortium for creating a nice working atmosphere. In particular I would like to thank the members of **Team 1**, Pär Nordlund, Pål Stenmark, Tomas Nyman, Susanne Flodin, Martin Moche, for sharing your projects with me.

I would like to thank **Ulf Willen** from Malvern Instruments, Uppsala, for your kind help with the DLS experiments on *Ba*-TK and *Bc*-TK.

Thank you: **Vladimir**, my first room-mate, for your short questions. **Andrea** my second room-mate for giving great concerts, being a master of Illustrator, and for sharing knäckebröd and your fascination for ncRNA. Professor and Hinas is a perfect combination. Good luck in Boston. We'll come and visit you. Thank you **Martin** for sharing the struggle with TKs, running the BK Modestos and for making **Lousie** come back to us. **Nisse** for being so enthusiastic about the latest releases of Linux and for being keen on trying them all out on my computer. **Lotta,** thank you for your positive thinking and lied-back stile in combination with enthusiasm for what you are doing. Keep on fishing! **Glareh,** thank you for your contiguous laughter. **Anton** and **Anatoly** for asking questions at the group meetings. **Jessica** for being an excellent master student. **Malin** and **Rosie**, you are simply the best! Thanks for the support, chats, laughs, synchrotron and skiing trips. Malin, you did a fantastic performance on P3, explaining crystallography to the late-night listeners. Thank you Rosie, for e-mailing me before I had even started in the group. Thank you **Jenny** and **Gösta** for your company in Florence, and all the fun in Piza, at Monte Carla and Blobban. The Poles for teaching me some "lab-Polish". **Agata** for always caring about other people. I'm glad you decided to return to Uppsala. Thank you **Wojtek** for our discussions and remember that some projects might be worth working on even with expressions below 7 mg/L culture.

Thank you **Stefan** for taking good care of the department, **Ulla**, **Jerry** and **Margareta** for fine leadership with the biochemistry courses, and **Elleonor** for your help with all administrative related issues. The place would not work without the great efforts and technical support from **Erling** and **Christer**. Thank you Christer for lending your hair-drier (twice).

There would be no structures without synchrotron trips. Many thanks to **Inger** and **Karin** for leading our BAG. Thank you Karin, **Linda**, **Al** and **Sara** for great company during the late nights at the beamlines.

Alwyn, I am thankful for your help with O, especially during model building of the first *Uu*-TK structure.

Thank you **Tex** for a great crystallization course and for always having a piece of advice when it comes to purification and crystallization problems.

Thank you: **Anna J.**, **Anna L**., **Adrian**, **Mark, Patrik**, **Emma**, **Talal** (I miss your fantastic dinners), **Andreas K.**, **Sherry**, **Gerard**, **Marian**, **EvaLena**, all the boys in the **Åqvist** group, all boys and girls working with **Janos** and **David**, **Magnus**, **Pavel** (I hope my lead-prophecy wont come true), **Annette**, **Nisse**, **Maria**, **Mats**, **Jonas**, **Sanju**, **Martin S.**, **Lars**, **Kaspars**, **Wimal**, **Seved**, **Torsten**, **Tom**, **Christofer**, **Nina**, **Alina**, **Daniel**, **Henrik x2** , **Ulrika, Lena**, **Dee, Monika,**...for making BMC a great place to work at.

I would also like to thank a few people outside the lab. My oldest friends **Agata** and **Dorota** for always being there, even though you are so far away.. **Jenny** and **Sofia** for all the fun we've had together. My "Karlstad-friends": **Peder**, **Anna**, världens bästa **Karlsson**, **Goswin** (det ordnade sig, precis som du alltid säger), **Daniel**, **Jocke**, **Johan**, **Rickard**, **Johanna P.**, **B.**, **Mikael** and the girls, my dearest friends **Sara**, **Anna** and **Cilla** (never stop asking questions) for our fabulous study sessions.

Thank you **Johan** and **Lisa** for all the fun here in Uppsala. Skål för Kamutasha!

Maćku to dzięki tobie znalazłam się w Uppsali, więc i ty się nie co przyczyniłes do tej pracy. Zawsze miałeś coś wesołego do opowiedzenia gdy tu studiowałes a ja, tak już od małego, chciałam być tam gdzie starszy brat bo tam zawsze najfajniej. Ponieważ jestem dość pamiętliwa chcę ci przypomnieć że wciąż czekam na te Birkenstocki które mi obiecałas jakieś pięć lat temu ☺.

Kochanej **mamie** i kochanemu **ojcu** za nieustające wsparcie. Bez was nie dałabym sobie rady. Dziekuję wam za wszystko.

Slutligen vill jag tacka **Jonas** för att du flyttade till Uppsala. Ditt stöd har betytt väldigt mycket för mig inte minst under den senaste tiden. Tack för de fina överraskningarna.