Structure-function studies of human cytosolic thymidine kinase

Berenstein, Dvora

Publication date:
2004

Structure-function studies

of human cytosolic thymidine kinase

Ph.D. Thesis

Dvora Berenstein

Department of Life Sciences and Chemistry
Roskilde University
Denmark
2004

To my father and my daughter

The road to wisdom?

Well, it's plain and simple to express:

Err and err and err again

But less and less and less.

Piet Hein

Contents   Page

PART II   EXPERIMENTAL WORK

Chapter 6. The role of amino acid 106 in the catalytic....................114 mechanism and oligomerization pattern of human thymidine kinase 1

Publications
Paper I: 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). J. Biol. Chem. 275, 32187-32192.

Paper II: Frederiksen, H., Berenstein, D. and Munch-Petersen, B. (2004). Effect of valine 106 on structure-function relation of cytosolic human thymidine kinase. Kinetic properties and oligomerization pattern of nine substitution mutants of V106. Eur. J. Biochem. 271, 2248-2256.

Preface

This Ph.D. study has been performed in the laboratory and under supervision of Professor Birgitte Munch-Petersen. The purpose of my work was to investigate the effect of amino acid 106 on structure-function relation of human thymidine kinase 1 (TK1) and to uncover the role of the C-terminal domain in this enzyme.

My thesis is divided in two parts - a general theoretical part and an experimental part which consists of three separate chapters, each with its own introduction and discussion.

The theoretical part introduces the present state of knowledge about the human thymidine kinase 1 and its role in the synthesis of DNA precursors.

The first experimental chapter describes the role of residue 106 for the catalytic and oligomerization properties of human TK1. The results presented in this chapter were obtained in collaboration with Hanne Frederiksen and Jacob Fromberg Christensen that in the years 1998-2000 were Master Students in our laboratory. These results have already been published and the papers (Paper I and II) are enclosed at the end of this thesis. I wish to thank both Hanne and Jacob for our co-operation in the laboratory and during the publication process.

The second experimental chapter describes cloning, expression and characterization of the putative interface domain of human TK1. Some of the results from this chapter are published in Paper II.

The third experimental chapter relates to the role of the last 40 C-terminal residues of TK1. A manuscript is under preparation in collaboration with Post Doc. Chunying Zhu.

At the end of this thesis a final discussion and summary are presented.

The work was started in 1995, but was interrupted by periods with concentrated teaching and my work with the "Genetic Organisation of the *Vibrio harveyi dnaA* Gene Region and Analysis of the Function of the *V. harveyi* DnaA Protein in *Escherichia coli*" published in Journal of Bacteriology 184, 2533-2538 (2002) by D. Berenstein, K. Olesen, C. Speck and O. Skovgaard.

Dvora Berenstein
Copenhagen
October, 2004

General introduction

Thymidine kinase (TK) is a key enzyme in the salvage pathway of nucleotide metabolism. The salvage pathway reuses free deoxyribonucleosides originating from DNA breakdown and recycles them back to DNA precursor deoxyribonucleoside triphosphates (dNTPs). TK catalyzes the initial step of the salvage pathway - transfer of γ-phosphoryl group from ATP to thymidine thereby forming thymidine monophosphate, TMP. Intracellular TMP is quickly phosphorylated to TDP and further to TTP. As a feedback inhibitor of thymidine kinase and an allosteric effector of ribonucleotide reductase, TTP regulates both the thymidine salvage pathway and *de novo* supply of purines and pyrimidines for DNA synthesis and repair. Two genetically distinct thymidine kinase isoenzymes are known in mammalian cells – TK1 and TK2. TK1 is cytoplasmic and cell cycle regulated and its activity is associated with the dividing cell. TK2 is localized to mitochondria and constitutively expressed. Many viruses carry their own thymidine kinase, and in malignant cells thymidine kinases with altered enzymatic properties have been described. Thymidine kinases phosphorylate several clinically important nucleoside analogues with antibacterial, antiviral and anti-leukaemic activity. Knowledge about tissue distribution, regulation, enzymatic properties and three-dimensional structure of thymidine kinases from normal cells and virus-infected or cancer cells is necessary in order to perform rational design and rational use of drugs for selective treatment of the affected cells. A classical example of a selective nucleoside analogue is acyclic guanidine – acyclovir (trade name Zovir), that has a potent anti-herpes virus activity. Acyclovir is phosphorylated by virus encoded thymidine kinase, but not by thymidine kinases encoded by the host. Accordingly, acyclovir phosphate is solely incorporated into DNA and inhibits DNA replication of virus infected cells.

TK1 has up to now been refractory to crystallization and no structural data have been reported for human TK1, nor for any other poxviral or cellular thymidine kinases to which it shows close homology. The investigation described in this thesis was undertaken in order to advance the understanding of structure-function relation of human TK1 with special focus on 1) the importance of amino acid 106 for oligomerization and kinetic properties and 2) the role of the C-terminal of TK1.

Aims of the present investigation

To elucidate the structure-function relation of human TK1 by:

o Identification of the naturally occurring amino acid at position 106 in human TK1 and examination of the significance of valine or methionine at this position by characterization of recombinant TK1 with Val-106 (V106WT) and with Met-106 (V106M) with respect to enzymatic and oligomerization properties. This study was started due to divergence between the TK1 amino acid sequence obtained from DNA sequencing in our laboratory and the TK1 amino acid sequence deposited in the NCBI GenBank database.

o Further investigation of the role of size, conformation and polarity of amino acid 106 for the function and structure of human TK1 by examination of nine recombinant mutant enzymes created by site-directed mutagenesis at amino acid site 106.

o Cloning, expression and characterization of the putative interface fragment of TK1.

o Examination of the role of the last 40 C-terminal residues of TK1 for *in vitro* kinetic and oligomerization properties of the enzyme. The C-terminal is known to have a regulatory role *in vivo* for cell cycle-regulated expression of TK1.

PART I

INTRODUCTION

Chapter 1

Synthesis of DNA precursors: *de novo* versus salvage pathways

All living cells contain DNA as the repository of the genetic information. The synthesis and repair of DNA requires continuous and balanced supply of DNA building blocks – the four deoxyribonucleoside triphosphates, dNTPs (Fig. 1.1).
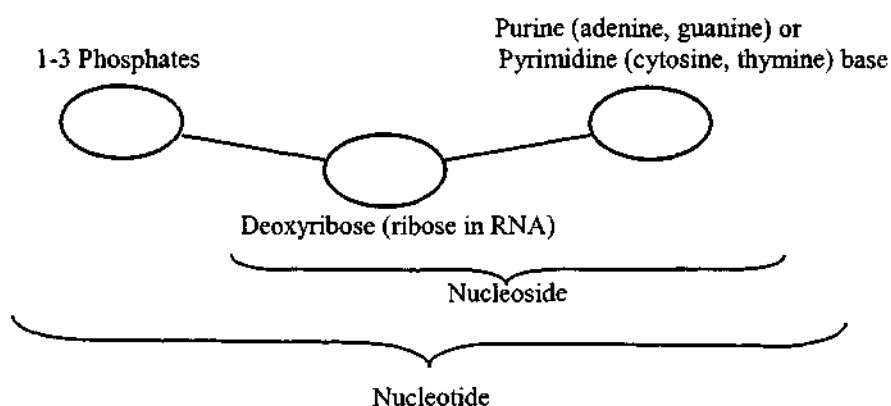
Figure 1. 1. A deoxynucleoside consists of a base and sugar deoxyribose. A nucleotide is a nucleoside with at least one phosphate group attached to it.

Mammalian cells have two pathways for the synthesis of deoxyribonucleotides (Fig. 1 2): *de novo* pathway, by which most of the deoxyribonucleotides arise, and salvage pathway.
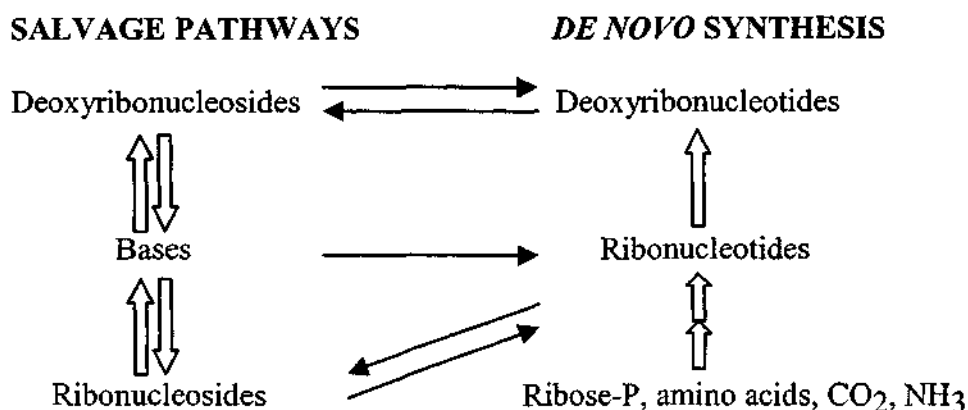
Figure 1.2. Salvage and *de novo* pathways of nucleotide metabolism. *De novo* synthesis includes pyrimidine and purine ribonucleotide synthesis and reduction of ribonucleotides. Salvage pathways indicate utilization of bases and nucleosides from extracellular precursors and from degraded nucleic acids. (After Kornberg & Baker, 1992)

In the *de novo* pathway the purine ring is build on a ribose-5-phosphate moiety (PRPP) by a sequence of reactions with specific amino acids (glycine, aspartate, glutamine) and CO_2. The pyrimidine ring in the *de novo* pathway is synthesized first, from carbamoyl phosphate and aspartate to orotate, which is then linked to ribose-5-phosphate. The end products of the *de novo* pathway are monophosphates of ribonucleosides (NMPs), also called ribonucleotides, and they have not occurred as free bases or ribonucleosides. The ribonucleoside monophosphates are phosphorylated to ribonucleoside diphosphates (NDPs), that in turn are substrates for an enzyme ribonucleotide reductase. Ribonucleotide reductase converts NDPs to dNDPs by reduction of ribose to deoxyribose. The (deoxy) nucleoside diphosphates get phosphorylated to triphosphates (NTPs and dNTPs) by a ubiquitous and non-specific (neither for the base, nor for the pentose sugar) nucleoside diphosphate kinase. The salvage pathway, in contrast to the *de novo* pathway, reuses free bases and deoxyribonucleosides released from damaged DNA (internal sources), or from dead cells and food digestion (external sources), and recycles them back to deoxyribonucleotides. Bases and deoxyribonucleosides, unlike deoxyribonucleotides, can freely pass the cell membrane. Actively dividing cells get nearly all DNA precursors from *de* novo pathway, but in resting cells, *de novo* nucleotide synthesis is absent, and cells are dependent on salvage enzymes for DNA repair and synthesis of mitochondrial DNA (Arnér & Eriksson, 1995). The relative contribution of *de novo* and salvage pathways to the pyrimidine nucleotide pools is different for different tissues and organs of an organism. *De novo* synthesis for example, makes the larger contribution to the intestine uracil nucleotide pool, *de novo* and salvage synthesis make roughly equal contribution to the liver pool, and salvage synthesis makes the larger contribution to the kidney pool (Zaharevitz *et al.*, 1992). Salvage pathway also provides majority of pyrimidine deoxynucleotides in S-phase thymocytes (Cohen *et al.*, 1983) and of deoxyliponucleotides in fibroblasts (Spyrou & Reichard, 1987) and lymphocytes (Spasokoukotskaja *et al.*, 1991). Some parasitic Protozoa like *Giardia lamblia* and *Trichomonas vaginalis* lack *de novo* pathway for synthesis of dNTPs and have to salvage DNA precursors from the host cell (Berens *et al.*, 1995). A strong indication for the importance of the pyrimidine nucleotide salvage pathway *in vivo* are the changes caused by thymidine kinase deficiency in $Tk^{-/-}$ knockout mice: kidney pathology and abnormal immune system resulting in shortened life spans

(less than one year) when compared with the wild type or heterozygous siblings (Dobrovolsky *et al.*, 2003).

Role of ribonucleotide reductase in the *de novo* pathway

Ribonucleotide reductase catalyses the committed step in the *de novo* synthesis of DNA precursors - a direct transformation of ribonucleotide diphosphates to corresponding deoxyribonucleotides - and is the central control element in this pathway. The enzyme has three remarkable features: it is cell cycle regulated, it involves free radicals to transform its substrate, and it is under allosteric control by both ribonucleotides and deoxyribonucleotides.

Cell cycle regulation

In accordance to its role in DNA synthesis, the activity of ribonucleotide reductase attains its maximum during S-phase and the highest reduction of ribonucleotides was reported in rapidly growing malignant cells (Elford *et al.*, 1970). The two subunits are differentially regulated during the cell cycle and the variation in holoenzyme activity is regulated by *de novo* synthesis and breakdown of the small R2 subunit with a half-life of 3 hours (Eriksson *et al.*, 1984). The concentration of R2 protein increased 3 – 7- fold with the onset of S-phase (Eriksson *et al.*, 1984), while the level of protein R1 was found to be constant, with a half-life of at least 15 hours throughout the cell cycle (Engström *et al.*, 1985). The mitotic proteolysis of the R2 protein has been recently demonstrated to occur via a ubiquitin-proteasome pathway and to depend on a KEN box sequence recognized by an ubiquitin protein ligase - the anaphase promoting complex (APC) and its activator protein Cdh-1 (Chabes *et al.*, 2003). The immunocytochemical studies localized ribonucleotide reductase to the cytoplasm, and showed, that the R2 protein was present only in S-phase cells, while the R1 protein was present in both S- and G1 phase of actively dividing cells, but not in terminally differentiated cells that stopped synthesis of DNA (Engström *et al.*, 1984; Engström & Rozell, 1988; Mann *et al.*, 1988). Recently, the localization of mammalian ribonucleotide reductase, especially after DNA damage, became questionable. A DNA damage-inducible homologue of the R2 subunit has been identified by screening a colon-cancer cell line for p53-inducible differentially expressed genes (Tanaka *et al.*, 2000;

Lozano & Elledge, 2000). The p53R2 protein was reported to translocate to the nucleus where it supplied the precursors for DNA synthesis and was directly involved in the process of DNA repair within the damaged resting cells (Tanaka *et al.*, 2000).

Structural and reaction mechanism aspects

Mammalian ribonucleotide reductase is a tetrameric protein composed of two different homodimers: the large subunit R1 and the small subunit R2. Both subunits are important for the catalytic activity of ribonucleotide reductase. The small R2 subunit contains an oxygen-bridged diferric iron center that forms and stabilizes the tyrosine free radical required for catalysis (Fontecave *et al.*, 1992; Fontecave 1998). The R1 subunit provides the catalytic site that contains redox-active dithiols and the binding site for the NDP substrates. During catalysis the radical function from the tyrosine of the R2 gets transferred through a cysteine residue in the catalytic site of R1 to the NDP substrate, and generates a reduced nucleotide. The R1 subunit also contains two separate allosteric specificity sites that bind ribonucleotide and deoxyribonucleotide effectors: the activity site and the specificity site (Fig. 1.3A).

Allosteric regulation

The activity site of the R1 subunit regulates the overall catalytic activity of the enzyme by binding either ATP (the positive effector), or dATP (the negative effector), and the substrate specificity site binds ATP, dATP, dTTP and dGTP, and determines which one of the four nucleotides will be reduced (Fig. 1.3A). Evidence for the independent activity and specificity sites in mouse ribonucleotide reductase was provided by characterization of separate mutations of the R1 protein with altered allosteric regulation (Eriksson *et al.*, 1981).

The rationale of ribonucleotide reductase regulation is as follows: the presence of ATP signals a high level of biosynthesis and abundance of ribonucleotides, whereas accumulation of dATP means abundance of deoxyribonucleotides and renders the enzyme inactive. ATP, as well as dATP in low concentrations bound to the substrate specificity site (Eriksson *et al.*, 1981), favor reduction of pyrimidines (Fig. 1.3B). dTTP bound at the specificity site, inhibits reduction of the pyrimidines CDP and UDP, and activates reduction of the purine GDP. Binding of dGTP stimulates ADP reduction and inhibits

reduction of GDP, CDP and UDP. Of the four dNTPs, only dCTP lacks properties of an allosteric effector for ribonucleotide reductase. (On the other hand, dCTP acts allosterically on dCMP deaminase and deoxycytidine kinase). Binding of any of the allosteric effectors was found to increase the interaction between R1 and R2 subunits (Ingemarson & Thelander, 1996).
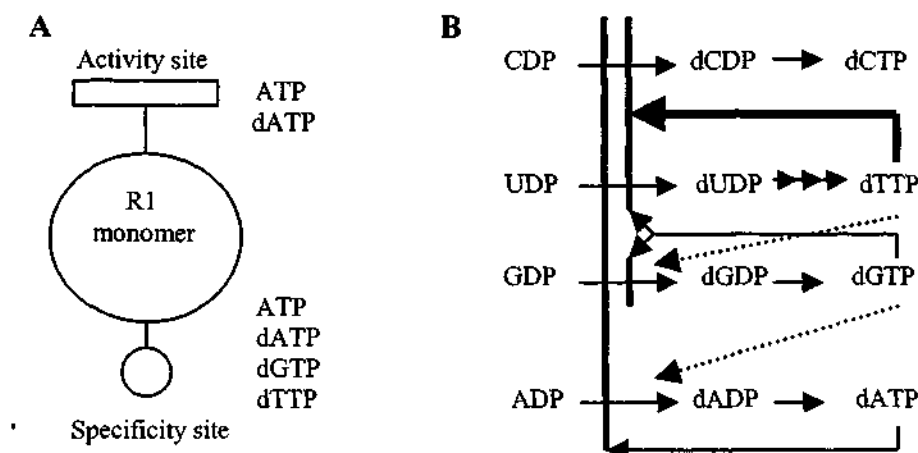
Figure 1.3. A. A simplified model for the allosteric regulation of the mammalian ribonucleotide reductase (after Jordan & Reichard, 1998). B. Scheme of the regulation of deoxyribonucleotide synthesis. The broken arrows stand for positive effects, the solid arrows and bars stand for negative effects. Modified after Thelander & Reichard (1979).

As can be seen, the allosteric sites communicate with the catalytic site regarding choice of the substrate and the activity of the enzyme. When ribonucleotide reductase reactions were monitored in the presence of four substrates simultaneously, also the binding of NDP substrates at the catalytic site influenced dNTP binding at the specificity site (Chimploy & Mathews, 2001). Interactions between the two allosteric effector-binding sites, i.e. the activity and specificity site of the R1 subunit, were also reported (Reichard et al., 2000). It is probable, that these interplays between the two allosteric sites and the catalytic site ensure that the four deoxynucleotides are synthesized according to the need of the cell, and that disturbances in the dNTPs pool size are prevented (Thelander and Reichard, 1979; Jordan and Reichard, 1998).

A new, and controversial, model for allosteric control of mammalian ribonucleotide reductase activity, based on molecular mass, ligand binding and enzyme kinetics studies has been presented recently. In this model, the most active form of the enzyme is depicted as an $R1_6R2_6$ heterohexamer, and not as typically in the literature, as an $R1_2R2_2$ heterodimer. Formation of the hexamer is proposed to be driven by ATP binding to a third

allosteric site on the R1 subunit (called hexamerization site), while dATP binding to the activity site induces formation of inactive $R1_4$ tetramer, and nucleotide binding to the specificity site of the R1 subunit drives formation of an active $R1_2R2_2$ dimer (Kashlan *et al.*, 2002; Kashlan & Cooperman, 2003). By estimation of the distribution of R1 among dimer, tetramer and hexamer forms *in vivo*, the authors suggest that $R1_6$ may predominate in the cytoplasm of mammalian cells, while $R1_2$ may be important for the nuclear activity of ribonucleotide reductase (Kashlan & Cooperman, 2003).

Biosynthesis of thymidylate

A thymine ribonucleotide as a precursor for (deoxy)thymidine does not exist, therefore the use of the prefix "deoxy" is optional (Kornberg & Baker, 1992). Most of dTMP in mammalian cells is derived from methylation of dUMP by the enzyme thymidylate synthase. dUMP in turn is supplied by deamination of dCMP or by action of a dUTPase on dUTP. Over 80% of dUMP originates from deamination of dCMP, which is the major pathway for dTMP synthesis (Nicander & Reichard, 1985). The *de novo* reactions responsible for synthesis of dTMP are shown in Fig. 1.4:

dUTP

↑

dUDP   dUTPase

↑

dCMP ——————▶ dUMP ——————▶ dTMP
   dCMP deaminase   thymidylate synthase
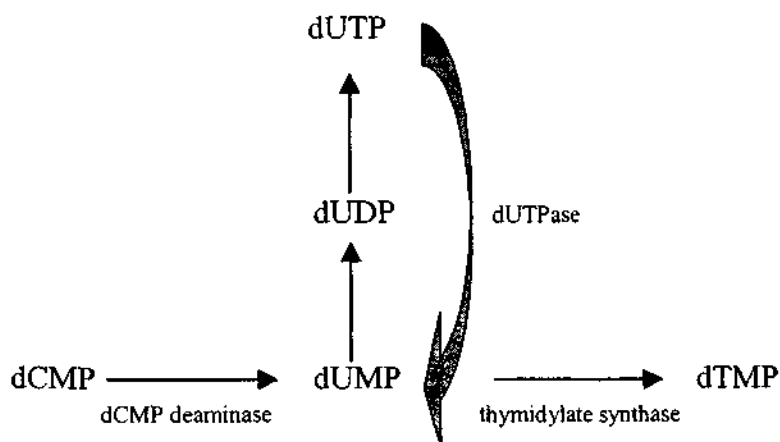
Figure 1.4. Synthesis of dTMP in mammalian cells by *de novo* pathway. The mammalian dUTPase is so effective that dUTP could not be detected in cell extracts (Reichard, 1988).

Salvage pathways

The salvage enzymes save the cell resources of energy, carbon and nitrogen by utilization of preformed bases and nucleosides. Bases can be converted into nucleosides, directly converted into ribonucleotides, or interconverted as for example cytosine to uracil. Also, enzymes engaged in transport of nucleosides across the cell membrane are classified as salvage enzymes. All animal cells possess nucleoside and nucleobase carrier proteins with wide substrate specificity, that facilitate entrance of nucleosides and bases into the cell either by non-concentrative diffusion system or by active transport (Plagemann *et al.*, 1988; Belt *et al.*, 1993; Cass *et al.*, 1999). Some specialized cells in the body, like erythrocytes or brain cells, are deficient in *de novo* synthesis and depend on nucleosides synthesised in for example liver, to be taken up from extracellular fluid and salvaged. The next step in the salvage pathway, phosphorylation of nucleosides to monophosphates, is considered as the key regulatory step, since the phosphorylated nucleotides are trapped intracellularly due to their negative charge (Arnér & Eriksson, 1995). The essential salvage pathways are shown in Fig. 1.5.
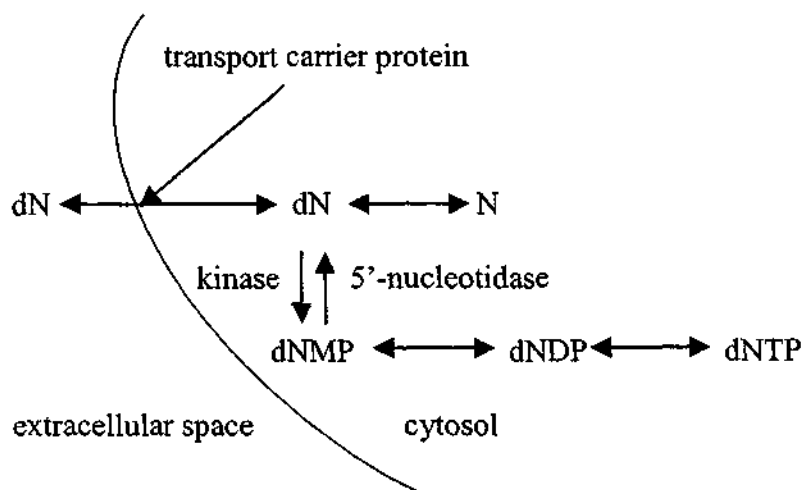
Figure 1.5. Salvage pathways for uptake and phosphorylation of nucleosides. dN is a deoxyribonucleoside, N is a base; dN and N are interconverted by nucleoside phosphorylases (the predominant direction is degradative). Nucleoside monophosphate kinases and nucleoside diphosphate kinases phosphorylate nucleotides of both the *de novo* and salvage pathways.

The four mammalian deoxyribonucleoside kinases - TK1, TK2, dCK and dGK, will be described in the next section. In this section 5'-nucleotidases and deoxyribonucleoside phosphorylases will be described (se Fig. 1.5).

5'-Nucleotidases

Reactions catalysed by deoxyribonucleoside kinases are irreversible, but their action is opposed by 5'-nucleotidases. A kinase and a 5'-nucleotidase catalyse interconversion between deoxyribonucleosides and their corresponding 5'-phosphates – these two enzymes participate in a "substrate cycle", where opposing reactions balance the concentration of the dNTPs by affecting flow of deoxyribonucleosides into the cell and phosphorylating them when dNTPs are in short supply, or by degrading surplus of dNTPs and sending them out of the cell as nucleosides (Bianchi *et al.*, 1986; Höglund *et al.*, 1988; Bianchi *et al.*, 1994).

Several types of 5'-nucleotidases have been described (reviewed by Arnér & Eriksson, 1995; and by Bianchi & Spychala, 2003). The 5'-nucleotidases differ in protein structure, substrate specificity and subcellular localization, some are ubiquitous, others have tissue-specific expression, and their nomenclature is quite confusing. For example, the cytosolic purine 5'-nucleotidase, also called high K_m nucleotidase, has been shown to participate in the regulation of the pools of ribonucleotides IMP and GTP, but not in the regulation of deoxyribonucleotides (Gazziola *et al.*, 2001), while the cytosolic 5'(3')-deoxynucleotidase (dNT-1) was shown to be involved in substrate cycles of pyrimidine ribo- and deoxyribonucleotides (Rampazzo *et al.*, 2000a; Gazziola *et al.*, 2001). A recently identified mitochondrial enzyme, dNT-2, related to dNT-1, was suggested to participate together with thymidine kinase 2 in a substrate cycle that protects mitochondrial DNA polymerase from an excess of dTTP (Rampazzo *et al.*, 2000b; Gallinaro *et al.*, 2002; Rampazzo *et al.*, 2004). The gene for dNT-2 was localized in a region of chromosome 17 which is affected in most cases of Smith-Magenis syndrome (a multiple congenital anomalies syndrome with moderate mental retardation), and the loss of dNT-2 was proposed to be involved in etiology of this disease (Rampazzo *et al.*, 2000b). Nucleotidases are also of clinical interest, because they may alter the therapeutic efficacy of certain nucleoside analogues (Hunsucker *et al.*, 2001).

Cells that are naturally low in nucleotidase activity, as for example T lymphocytes, are especially sensitive to accumulation of dNTPs, if any other enzyme that removes nucleosides (deaminases, phosphorylases) is inhibited or missing (Carson *et al.*, 1981). For instance, adenosine deaminase (ADA) deficiency leads to a severe combined immunodeficiency disorder. ADA catalyses the deamination of adenosine and

deoxyadenosine, and it is known that ADA deficient patients have a 1000-fold increased dATP level in their erythrocytes (Coleman *et al.*, 1978). Since dATP is a powerful inhibitor of ribonucleotide reductase, the levels of other three nucleotides are greatly reduced, DNA replication and repair inhibited, resulting in DNA double strand breaks and cell death (Yoshioka *et al.*, 1987).

Deoxyribonucleoside phosphorylases

Two deoxyribonucleoside phosphorylases are known: purine nucleoside phosphorylase (PNP) and pyrimidine nucleoside phosphorylase (Kornberg & Baker, 1992; Pugmire & Ealick, 2002). The purine phosphorylase catalyses the phosphorolytic cleavage of (deoxy) inosine and (deoxy)guanosine, but not (deoxy)adenosine, which needs first to be deaminated to (deoxy)inosine (Kornberg & Baker, 1992). The mammalian pyrimidine phosphorylase is specific for thymidine, and is also called thymidine phosphorylase, or TP (Pugmire & Ealick, 2002). Similarly to the 5'-nucleotidases, the phosphorylases function against accumulation of to big pools of (deoxy)ribonucletides.

Purine nucleoside phosphorylase (PNP) deficiency is associated with a severe defect of T-lymphocyte function, that results from more than 10-fold accumulation of dGTP in T cells (Cohen *et al.*, 1978). dGTP, in turn, is an allosteric regulator of ribonucleotide reductase and the result is a genotoxic pool imbalance. Arpaia *et al.* (2000) have proposed that the severe combined immune deficiency in PNP knockout mice is due to selective accumulation of dGTP in the mitochondria of T cells, leading to defective mitochondrial DNA repair and T cell apoptosis.

Recently, a severe mitochondrial dysfunction, mitochondrial neurogastrointestinal encephalomyopathy (MNGIE), has been associated with mutation in the gene specifying thymidine phosphorylase. TP activity in leukocytes from MNGIE patients was less than 5% of controls, and the normally undetectable plasma levels of thymidine and deoxyuridine were dramatically increased in all patients (Nishino *et al.*, 1999; Spinazzola *et al.*, 2002; Marti *et al.*, 2003). The increased levels of dThd and dUrd have been associated with site-specific mitochondrial DNA point mutations observed in MNGIE patients with TP deficiency (Marti *et al.*, 2003; Nishigaki *et al.*, 2003), and with multiple deletions and depletion of mtDNA in skeletal muscle (Hirano *et al.*, 1994). Multiple

deletions in mtDNA were also found in HeLa cells cultured in medium supplemented with 50 μM thymidine for 8 months, which supported the hypothesis that dNTP pool imbalance is responsible for mtDNA mutagenesis (Song *et al.*, 2003). On the other hand, increased levels of TP have been reported in many tumours, and are considered to confer poor prognosis (Ackland & Peters, 1999). Thus, too little or too much enzyme activity may lead to disease.

Human deoxyribonucleoside kinases

Basic properties

The four human deoxyribonucleoside kinases have been cloned and expressed in bacteria (TK1 by Bradshaw & Deininger, 1984; TK2 by Johansson & Karlsson, 1997a and by Wang *et al.*, 1999a; dCK by Chottiner *et al.*, 1991, and dGK by Johansson & Karlsson, 1996 and by Wang *et al.*, 1996). The comparison of the predicted amino acid sequences showed, that deoxycytidine kinase - dCK, deoxyguanosine kinase – dGK, and the mitochondrial thymidine kinase - TK2, are closely related , while the cytosolic thymidine kinase - TK1, cannot be included in this enzyme family (Eriksson *et al.*, 2002). Alignment of the dGK amino acid sequence with the dCK, and the Herpes virus thymidine kinase amino acid sequences suggested a close evolutionary relationship (Wang *et al.*, 1996; Johansson & Karlsson, 1996; Eriksson *et al.*, 2002). As early as in 1991, it was suggested that higher and lower vertebrate herpesvirus TKs have evolved from a captured cellular dCK gene (Harrison *et al.*, 1991).

The functional similarities and differences within deoxynucleoside kinases are substantiated by the fact that dCK, dGK and TK2, but not TK1, have relaxed stereospecificities against their nucleoside substrates: in contrast to TK1 they do not distinguish between D- and L-isomers of β-nucleosides (Wang *et al.*, 1999b). Each kinase phosphorylates several deoxyribonucleoside substrates, with the name referring to the preferred one. The final products of deoxyribonucleoside phosphorylation, the corresponding dNTPs, regulate enzyme activity by feedback inhibition – the best inhibitor is the triphosphate of the preferred substrate (Arnér & Eriksson, 1995).

Deoxyribonucleoside kinases play an essential role in activation of anticancer and antiviral nucleoside analogues – the drugs must be converted to the corresponding triphosphates in order to interfere with viral or cellular DNA replication and repair, and eventually lead to apoptosis of malignant cells (Leoni *et al.*, 1998). Because the first phosphorylation step, catalysed by deoxynucleoside kinases, is considered to be rate limiting, it is important to know the substrate specificity and tissue distribution of the particular enzyme to achieve successful therapy (Eriksson *et al.*, 1994).

Thymidine kinases – TK1 and TK2

TK1 and TK2 (EC 2.7.1.21) are two different proteins and they certainly deserve a separate EC number. They are coded by different genes, have different substrate specificity and different substrate kinetics, differ in phosphate donor specificity and are located in different cellular compartments. Besides, TK1 is strictly cell cycle regulated and is absent from non-dividing cells, while TK2 is constitutively expressed.

Cytosolic thymidine kinase - TK1

The cytosolic thymidine kinase, TK1, (EC 2.7.1.21.) is probably the most studied among the salvage pathway enzymes due to the fact that its expression is cell cycle regulated and cancer tissues show higher activity, and because it can activate some clinically important analogues (Arnér & Eriksson, 1995). TK1 is widely distributed in human tissues, but being S-phase correlated, it is only detected in proliferating cells, either normally growing or malignant (Munch-Petersen & Tyrsted, 1977; Tyrsted & Munch-Petersen 1977; Arnér *et al.*, 1992; Arnér & Eriksson, 1995). The human TK1 gene has been localized to chromosome 17 (Miller *et al.*, 1971), more precisely to its long arm (McDougall *et al.*, 1973), at the position 17q23.2-q25.3 (Solomon *et al.*, 1993).

The first human TK1 cDNA clone, derived from SV40 transformed human fibroblasts, was characterized by Bradshaw and Deininger (1984). The cDNA was 1421 bases long (exclusive the polyA tail and the polyG tail), the open reading frame had 702 bases pairs and coded for a protein of 234 residues with a calculated molecular size of 25.5 kDa. The complete human TK1 gene, that originated from HeLa cells, was cloned by Bradshaw

(1983), and sequenced and characterized by Flemington *et al.* (1987). The genomic TK1 was 12.9 kb long, was composed of 7 exons and had 13 repetitive *Alu* sequences within its introns, comprising about 25% of the total sequence of the human *tk* gene (Flemington *et al.*, 1987). This is one of the highest densities of the *Alu* containing sequences in the human genome (Flemington *et al.*, 1987; Slagel *et al.*, 1987). The sizes of the exons were 130, 32, 111, 94, 90, 120 and 851 base pairs for exons 1 through to 7 respectively (Flemington *et al.*, 1987). A detailed description of the human TK1 promoter and of the complicated regulation of TK1 expression on several levels is presented in the Chapter 3: "Regulation of TK1 expression".

The TK1 protein was reluctant to purification due to its low abundance and high instability, and in early investigations addition of ATP has been used to stabilize the enzyme (Lee & Cheng, 1976; Munch-Petersen, 1984). The early observations diverged with regard to the molecular weight of the native TK1 (from 55 to 200 kDa) and its subunit size (24 to 65 kDa), as summarized by Munch-Petersen (1996). The discrepancies were probably due to contaminating proteins and ability of TK1 to oligomerize in the presence of ATP: the size of the active, partly purified TK1 increased from 75 to 175 kDa when eluted from G-200 Sephadex column in the presence of 2 mM ATP (Munch-Petersen, 1984). The first purifications of human TK1 to homogeneity, where the size of the pure, substrate-free enzyme was in agreement with the predicted subunit size from the cloned gene, were reported when TK1 from S-phase HeLa cells was stabilized by digitonin (Sherley & Kelly, 1988a), and TK1 from PHA stimulated human lymphocytes, was stabilized with the detergent CHAPS (Munch-Petersen *et al.*, 1991). The molecular size of native TK1 as found by Sherley & Kelly (1988a) was 96 kDa, indicating a tetramer, as the subunit molecular mass estimated in SDS-PAGE was 24 kDa. The same subunit molecular mass of 24 kDa was found for the pure homogeneous preparation of TK1 from PHA stimulated human lymphocytes (Munch-Petersen *et al.*, 1991). Native TK1 eluted from a Superose 12 column with an estimated molecular mass of 55-56 kDa in the absence and 110-120 kDa in the presence of 2 mM ATP (Munch-Petersen *et al.*, 1991; Munch-Petersen *et al.*, 1993). Use of Superdex 200, a different gel-matrix, for the same enzyme preparation gave higher molecular size estimates: 70 kDa and 150 kDa without and with ATP respectively (Jensen & Munch-Petersen, 1995; Munch-Petersen *et al.*, 1995a), in agreement with the above mentioned earlier results (Munch-Petersen, 1984). Discrepancies

in the estimation of molecular size due to the choice of chromatographic material have also been reported for the purified mouse TK2 and the related deoxynucleoside kinase from *Drosophila melanogaster*, and are mentioned later.

When pure, substrate free TK1 was isolated, and stabilized in the presence of CHAPS, a new ATP effect on enzyme activity was observed (Munch-Petersen *et al.*, 1993; Munch-Petersen *et al.*, 1995a). Incubation of TK1 or storage with 2 mM ATP induced a reversible, enzyme concentration-dependent transition from a form with low substrate affinity ($K_m \cong$ 12 μM) and a non-hyperbolic biphasic kinetics to a form with high substrate affinity ($K_m \cong$ 0.7 μM) and hyperbolic Michaelis-Menten kinetics (Munch-Petersen *et al.*, 1993). Similar shifts could not be induced by thymidine. Because the effect of ATP on the size of TK1 was reversible as well, it was reasoned that the tetramer was associated with the high-affinity form, and the dimer with the low-affinity form (Munch-Petersen *et al.*, 1993). Surprisingly, the recombinant TK1 expressed in *E. coli* from a plasmid that contained the amino acid coding sequence from the TK1 clone of Bradshaw and Deininger (1984), showed no ATP effect and behaved as a permanent tetramer with high affinity to thymidine ($K_m \cong$ 0.4 μM) irrespective of the preexposure to ATP (Jensen & Munch-Petersen, 1995; Munch-Petersen *et al.*, 1995a). At that time, the divergent properties of the recombinant TK1 were explained by absence of post-translational modification of TK1 when expressed in *E. coli* (Jensen, 1994; Munch-Petersen *et al.*, 1995a). When the amino acid coding region of TK1 from normal human lymphocytes was cloned into the pGEX-2T vector (described in Paper I, Berenstein *et al.*, 2000), it turned out that the recombinant lymphocyte TK1 (rLy-TK1) behaved like the endogenous enzyme regarding the ATP transition effect on oligomerization and thymidine affinity. Explanation of these conflicting results came with sequencing of the lymphocyte TK1 cDNA which showed presence of GTG codon for valine at amino acid position 106 and not the ATG codon for methionine which is present at this position in the cDNA clone of Bradshaw and Deininger (1994). The first sequencing reactions were run by Tina Kristensen, at that time a Ph.D. student (Kristensen, 1996). The project was continued by me, Jacob Fromberg Christensen and Hanne Frederiksen, at that time M. Sc. students (Frederiksen, 1999; Christensen, 2000). The results are presented in Chapter 6 "The role of amino acid 106 in the catalytic mechanism and oligomerization pattern of human thymidine kinase 1"and have already been published (Paper I and II; Berenstein *et al.*, 2000; Frederiksen *et al.*, 2004).

The substrate specificity of TK1 is more restricted than of the other salvage enzymes. TK1 accepts only thymidine and deoxyuridine, but it has the highest specificity constant k_{cat}/K_m with its primary substrate, dThd, which is equal to $8 \times 10^6 \, M^{-1} sec^{-1}$, about 40 times higher than that of dCK with dCyd (Munch-Petersen, personal communication). The preferred phosphate donors are ATP and dATP, while dTTP is the feedback inhibitor (Munch-Petersen, 1984; Munch-Petersen et al., 1995b). Analogues of dThd and dCyd with minor substitutions at the 5-position, e.g. FdUrd, 5-fluorodeoxyuridine, or at the 3'-position of the sugar, e.g. the anti-HIV compounds FLT, 3'-fluoro-2',3'-dideoxythymidine and AZT, 3'-azido-2',3'-dideoxythymidine are also allowed (Munch-Petersen et al., 1991; Amér & Eriksson, 1995). Bulkier substitutions such as 5-propenyl in 5-propenyldUrd, and arabinofuranosyl- or acyclic sugars are not accepted (Eriksson et al., 1991; Amér & Eriksson, 1995). AZT was an efficient substrate, with K_m values in the same range as that obtained with thymidine (Munch-Petersen et al., 1995b). AZT substrate kinetics displayed a hyperbolic reaction mechanism, and it was proposed that the TK1 form reacting with AZT was the tetrameric, high-affinity form (Munch-Petersen et al., 1995b).

Several studies have shown that activity of TK1 fluctuates with the cell cycle and is highly correlated to DNA synthesis (Bello, 1974; Munch-Petersen & Tyrsted, 1977). TK1 has therefore been suggested as a tumour marker/indicator of malignancy (Hallek, 1992; Romain et al., 2000; Mao et al., 2002; Mizutani et al., 2003) and as a predictor of tumour recurrence (O'Neill et al., 1995).

The three-dimensional structure of TK1 is not known, nevertheless, in Chapter 4: "Structure of the human TK1", I will approach to describe its putative 3D structure based on the prediction of the secondary structure and comparison to the available 3D structures of the multisubstrate deoxynucleoside kinase from Drosophila melanogaster, dNK, the human deoxyguanosine kinase, dGK (Johansson et al., 2001) and the Herpes simplex virus thymidine kinase (Wild et al., 1995; Wild et al., 1997).

Mitochondrial thymidine kinase - TK2

Mitochondrial thymidine kinase, TK2, shares up to now its EC number with TK1 (EC 2.7.1.21). TK2 is present in most tissues, in correlation to the number of mitochondria and independent of the growth state of the cell (Amér & Eriksson, 1995). When compared to the level of TK1 in proliferating cells, the level of TK2 is low, but TK2 becomes

significant in resting and terminally differentiated cells, such as muscle or brain and unstimulated lymphocytes, where TK1 activity is barely detectable (Munch-Petersen & Tyrsted, 1977; Eriksson & Wang, 1997). Presence of thymidine kinase activity in mitochondria was detected in mouse TK$^-$ fibroblasts, deficient in the cytoplasmic TK activity and resistant to BrdUrd (5-bromodeoxyuridine). The TK$^-$ cells incorporated thymidine or BrdUrd into mitochondrial, but not nuclear DNA (Attardi & Attardi, 1972; Berk & Clayton, 1973). The human gene for TK2 has been located to chromosome 16 of the nucleus using mouse/human hybrid cells (Willecke et al., 1977), and later mapped to region 16q22 (Johansson & Karlsson, 1997a).

Initial attempts to clone the human TK2 gene resulted in incomplete human TK2 cDNA which did not include a mitochondrial translocation signal (Johansson & Karlsson, 1997a; Wang et al., 1999a). The cloning of full-length human TK2 cDNA containing the 5'-region with the mitochondrial leader sequence was reported by Wang et al. (2003). The human TK2 gene was shown to occupy a 45 kb DNA fragment and to consist of 10 exons, from 32 bp to 1304 bp, and 9 introns, from 0.572 kb up to 11.1 kb (Wang et al., 2003). The mitochondrial leader sequence was coded by exon 1 (Wang et al., 2003). The size of the TK2 polypeptide deduced from its coding sequence was about 30 kDa (Wang et al., 1999a; Wang et al., 2003), in agreement with the subunit size of the purified native TK2 determined by SDS electrophoresis (Munch-Petersen et al., 1991). As to the molecular size of native TK2, diverging observations have been reported. Native TK2 purified from leukemic spleen eluted as a monomer of 30 kDa from Superose 12 column (Munch-Petersen et al., 1991). Also recombinant human TK2 subjected to gel filtration on a Superdex 200 column eluted as a peak of 30-40 kDa, indicating a monomer (Wang et al., 2003). But partially purified lymphocyte TK2 (DEAE fractions) eluted from Superdex 200 before bovine serum albumine, with an apparent size of 70 kDa, indicating a dimer (Munch-Petersen, 1984; Munch-Petersen, 1996). Similar discrepancies have been reported for the purified mouse TK2 and the related deoxynucleoside kinase from Drosophila melanogaster (Dm-dNK): both enzymes eluted as a dimer when judged by Superdex 200 chromatography, but as a monomer when judged by Superose 12 (Wang & Eriksson, 2000; Munch-Petersen, et al., 1998; Knecht et al., 2002a). Because Dm-dNK, dGK and dCK are dimeric proteins in crystals (Johansson et al., 2001; Sabini et al., 2003), and dCK and dGK were also shown to be dimers by gel filtration (Datta et al., 1989a; Wang et al., 1993), it is

now believed that TK2 and *Dm*-dNK are dimers. The artefactual result on Superose 12 might be explained with interaction of the enzymes with the hydrophobic matrix of the gel filtration column (Knecht *et al.*, 2002a; Munch-Petersen, personal communication).

TK2 has broader substrate specificity than TK1, it phosphorylates all three pyrimidine nucleosides - deoxythymidine, deoxyuridine, deoxycytidine and their analogues (Munch-Petersen *et al.*, 1991). TK2 shows negative cooperative reaction mechanism with thymidine, as the affinity of TK2 to the substrate decreases with increasing substrate concentration. Two apparent K_m values for thymidine were calculated from biphasic Hofstee plot at low and at high substrate concentration: 0.3 μM at substrate concentration below 8 μM, and 16 μM at substrate concentrations above 8-10 μM (Munch-Petersen *et al.*, 1991). It may be advantageous for resting cells to have an enzyme with negative cooperativity: at the physiological extracellular thymidine concentration ranging from 0.1 μM to 1.2 μM (Holden *et al.*, 1980), TK2 with high affinity, although present in low amounts, will supply the cell with thymidine nucleotide for DNA repair and replication of mitochondrial DNA. On the other hand, decrease of TK2 affinity towards its substrate at higher substrate concentration will protect the cell against unnecessary nucleotide production.

TK2 phosphorylated dCyd and dUrd with Michaelis-Menten kinetics, the K_m values were 36 μM and 6 μM respectively (Munch-Petersen *et al.*, 1991). ATP and CTP both served as phosphate donors, and dTTP was a noncooperative, competitive inhibitor. With T or U as base, TK2 accepts 2'- and 3'- substitutions on sugars, as well as arabinofuranosyl sugars, but with C as base only 2'- substitutions are allowed (Eriksson *et al.*, 1994). The clinically important analogues phosphorylated by TK2 are: FdUrd (5-fluoro-2'-deoxyuridine), where the monophosphate is an inhibitor of thymidylate synhethase; AraT (1-β-D-arabinofuranosyl-thymidine), that is neither phosphorylated by TK1, nor dCK and therefore could be used in a TK2 selective assay in crude cell extracts (Arnér *et* al., 1992); the anti-hepatitis drug FIAU (1-(2'-deoxy-2'fluoro 1-β-D-arabinofuranosyl)-5-iodouracil) (Wang & Eriksson, 1996); and to a limited extend the antiviral analogue AZT (3'-azido-2'3'-dideoxythymidine). The V_{max} for AZT was 5 – 6 % of that with dThd, and the substrate kinetics with AZT gave biphasic Hofstee plots indicating negative cooperativity (Munch-Petersen *et al.*, 1991). The negative cooperativity was not seen with dCyd, dUrd

or FdUrd, but with dThd and AZT, indicating that this kinetic behaviour might be associated with the base and not the sugar.

Recently, several mutations in the TK2 gene that reduce enzyme activity to 14-45% have been identified, and found to be associated with myopathic form of mitochondrial DNA depletion syndrome (MDS) in homozygous individuals (Saada *et al.*, 2001; Elpeleg *et al.*, 2002; Mancuso *et al.*, 2002; Carrozzo *et al.*, 2003). MDS is phenotypically heterogeneous, but starts in early childhood and is characterized by tissue-specific reduction in mtDNA copy number. Mutations in the mitochondrial deoxyguanosine kinase have been associated with the hepatocerebral form of MDS and will be described later. Introduction of one of the TK2 gene mutations found in a patient with severe myopathy - H121N, to the recombinant TK2 resulted in an enzyme that lost the negative cooperativity with thymidine as substrate and displayed Michaelis-Menten kinetics (Wang *et al.*, 2003). This finding emphasized the significant physiological role played by the enzymatic reaction mechanism. The indispensability of the deoxyribonucleoside salvage pathways in the pathogenesis of mitochondrial DNA depletion was further underscored by an observation of a markedly decreased dTTP content and a reduced dATP content in non-replicating fibroblasts from two patients with mutations in TK2 (Saada *et al.*, 2003).

Deoxycytidine kinase - dCK

Deoxycytidine kinase (EC 2.7.1.74) activity is found in the cytosol of many cells of the body: the highest activity in lymphoid tissues and the lowest in nerve, liver and muscle tissues (Spasokoukotskaja *et al.*, 1995; Eriksson & Wang, 1997). Human dCK cDNA was cloned in 1991 (Chottiner *et al.*, 1991) and the gene has been mapped to chromosome 4, band q13.3-q.21.1 (Song *et al.*, 1993; Stegmann *et al.*, 1993). The substrate specificity of dCK is broad, and includes purine deoxyribonucleosides and clinically important analogues in addition to deoxycytidine (Eriksson *et al.*, 1994). Probably due to the high expression of dCK in lymphoblasts, nucleoside analogues, that are efficient substrates for dCK, such as CdA (2-chloro-2'-deoxyadenosine), are efficient anti-leukemic drugs (Arnér & Eriksson, 1995). With C as the base, dCK accepts analogues with acyclic sugars and analogues with both 2'- and 3'- modifications, as for example the anti-HIV drug ddC (2'-3'-dideoxycytidine). With A or G as the base, dCK is much more restricted on the sugar,

and for example accepts AraA (9-β-D-arabinosyladenine), but not acyclic sugars and 3'-substitutions (Eriksson *et al.*, 1994; Eriksson & Wang, 1997).

dCK does not follow Michaelis-Menten kinetics, but shows negative cooperativity towards deoxycytidine and ATP (Ives & Durham, 1970; Bohman & Eriksson, 1988). dCK phosphorylates its natural nucleosides with different specificities: lowest K_m values for deoxycytidine (0.8 μM), and much higher for deoxypurines (approximately 500 μM) (Bohman & Eriksson, 1988; Datta *et al.*, 1989a,b). Still, with the specificity constant k_{cat}/K_m equal to 2×10^5, 8×10^4, and 6×10^4 (in $M^{-1}sec^{-1}$) for dCyd, dAdo and dGuo respectively (numbers from Eriksson *et al.*, 2002), all the substrates are efficiently phosphorylated. Any endogenous cellular nucleoside triphosphate, except dCTP which is a feedback inhibitor, could function as a phosphate donor for dCK (Shewach *et al.*, 1992; Datta *et al.*, 1989a; White & Capizzi, 1991; Hughes *et al.*, 1997). UTP, or a combination of UTP and another nucleotide, was the preferred phosphate donor, and more efficient than ATP: K_m values for nucleoside substrates decreased when UTP, and not ATP, was used as the phosphate donor (Shewach *et al.*, 1992). Other investigations also indicated UTP to be the preferred intracellular phosphate donor for dCK (White & Capizzi, 1991; Krawiec *et al.*, 1995). In later studies from Shewach's group, K_m values of 1.2μM and 54 μM for UTP and ATP respectively were found, and ATP was shown to be the poorest phosphate donor when compared to UTP, GTP, CTP and dTTP (Hughes *et al.*, 1997). The reaction mechanism with ATP as phosphate donor was bi-bi random, but with UTP it was ordered – binding of UTP prior to deoxycytidine was strongly preferred (Hughes *et al.*, 1997).

A model for dCK action based on stopped-flow fluorescence spectroscopy was developed (Turk *et al.*, 1999): From the fluorescence changes induced by binding of various phosphate donors and deoxycytidine to dCK, the authors conclude that UTP binds to dCK before dCyd and that binding of UTP induces a conformational change in the enzyme enabling efficient phosphoryl transfer. Binding of dCyd to dCK-UTP complex is followed by additional conformational rearrangement and subsequent phosphoryl transfer (Turk *et al.*, 1999). Recently, Krawiec *et al.* (2003) have reported that inorganic tripolyphosphate (PPP$_i$) is a good ATP-competitive donor with deoxycytidin, but not with deoxyadenosine as substrate. The authors conclude that interaction of the base moieties of 5'-NTPs with the enzyme is not essential for donor activity and phosphate transfer reaction (Krawiec *et al.*, 2003).

With the crystal structure of dCK lately solved (Sabini *et al.*, 2003), the molecular nature of its substrate specificity became available, and dCK mutants catalytically superior toward anticancer and antivirus prodrugs are being designed. The protein structure can also be used to develop improved antiviral and anticancer drugs (Sabini *et al.*, 2003).

Human dCK is a homodimers of two 30 kDa subunits (Bohman & Eriksson, 1988; Datta *et al.*, 1989a) coded by an ORF of 780 bp in the cloned cDNA (Chottiner *et al.*, 1991). The promoter sequence of the human dCK gene has one imperfect E2F factor binding site to which E2F factor binds only weakly (Song *et al.*, 1993; Chen *et al.*, 1995; Johansson *et al.*, 2000). This might be a part of the explanation why dCK, despite its functional similarity to thymidine kinase 1, is not cell cycle regulated. Besides the E2F binding site, the dCK promoter contains two GC-boxes and an E-box, the binding sites for transcription factors Sp1 and USF (upstream stimulatory factor), respectively (Chen *et al.*, 1995). Physical interactions between proteins from the Sp and the USF families have been documented and found to elevate or attenuate the dCK promoter activity depending on the concentration and the combination of specific Sp and USF factors (Ge *et al.*, 2003). Such interactions between Sp and USF families of proteins could contribute to tissue-specific pattern of dCK expression (Ge *et al.*, 2003) and to differential expression of dCK during the cell cycle. In fact, the results regarding cell cycle regulation of dCK are conflicting. Mitchell *et al.* (1993), for example, found the same dCK mRNA level in resting and proliferating T lymphocytes, and a constant dCK mRNA level through the cell cycle. Two of four cell lines investigated by Hengstschläger *et al.* (1993) displayed a 10-fold increase in the activity of dCK at the G1/S boundary, but as none of the cell lines demonstrated differences in the level of dCK mRNA, a post-transcriptional mechanism involving stability of dCK protein was proposed to regulate expression of dCK in certain cell types (Hengstschläger *et al.*, 1993). Actually, Spasokoukotskaja *et al.* (1995) have concluded from expression of dCK detected by quantitative immunoblotting and by enzymatic assays, that variability between various cell types in dCK expression is much larger than that between cell cycle phases. On the other hand, increased dCK expression and activity levels have been found in some solid tumours (Spasokoukotskaja *et al.*, 1995). Also, an increase in dCK enzyme activity without any increase in the dCK protein level was reported in lymphocytes treated with inhibitors of DNA synthesis, deoxynucleoside analogues and gamma-irradiation, as well as with protein phosphatase inhibitors (Sasvári-Székely *et al.*,

1998; Spasokoukotskaja *et al.*, 1999; Csapó *et al.*, 2001a; Csapó *et al.*, 2003). The elevated dCK activity was suggested to be due to reversible phosphorylation of the enzyme (Csapó *et al.*, 2001b; Csapó *et al.*, 2003), and assumed to be a mechanism compensating for DNA damage through increased supply of all dNTPs needed for DNA repair (Keszler *et al.*, 2003; Csapó *et al.*, 2003). Free cytosolic calcium was found to be necessary for the activation of dCK and maintaining its active state, not as a cofactor, but presumably through regulation of a putative kinase which phosphorylates and activates dCK (Keszler *et al.*, 2004). Because a correlation was found between the dCK activity and the intensity of dCK signal in native, but not in denaturing, Western blots, the activation process was suggested to be accompanied by a more open enzyme conformation, where the epitope becomes more accessible to the antibody (Keszler *et al.*, 2004). These results indicate that dCK can exist in different conformational states, and by this support the conclusion from the above mentioned fluorescence quenching experiments of Turk *et al.* (1999).

Unexpectedly, overexpressed dCK was found not in the cytosol but almost exclusively in the cell nucleus, and a nuclear localization sequence, KRSCPSFSASSEGIRIKK, was identified in the N-terminal of dCK (Johansson *et al.*, 1997b). Because endogeneous dCK was found only in the cytosol, presence of a cytoplasmic retention mechanism for dCK, that involves another, up to now unknown, protein has been suggested (Hatzis *et al.*, 1998).

Deoxyguanosine kinase - dGK

Deoxyguanosine kinase (EC 2.7.1.113) is a mitochondrial enzyme found in most tissues: brain, muscle, liver, spleen, skin, resting and mitogen-stimulated lymphocytes (Wang *et al.*, 1993). The gene for the human dGK has been cloned (Johansson & Karlsson, 1996; Wang *et al.*, 1996) and localized to the short arm of chromosome 2, 2p13 (Johansson *et al.*, 1996). The level of dGK activity seems to be correlated to the amount of mitochondria found in a cell or tissue (Amér & Eriksson, 1995). The highest activity of dGK, up to 10 – fold higher as compared to other tissues, is found in the brain (Wang *et al.*, 1993). A cDNA from a human brain cDNA library was found to encode a 30 kDa protein (Wang *et al.*, 1996; Johansson & Karlsson, 1996). The N-terminal of the deduced amino acid sequence had similarities to mitochondrial targeting signal peptide, and cleavage during mitochondrial import would result in mature protein of 28 kDa (Wang *et al.*, 1996; Johansson & Karlsson, 1996). Western blotting of mitochondrial and cytosolic fractions

with polyclonal antibody raised against human recombinant deoxyguanosine kinase, and *in situ* immunofluorescence studies with the same antibody, confirmed the localization of the dGK to the mitochondrial matrix (Jüllig & Eriksson, 2000). Surprisingly, dGK was found to relocate to the cytosole together with cytochrome *c* during apoptosis (Jüllig & Eriksson, 2001). As dGK is involved in dATP synthesis, and dATP facilitates formation of the apoptosome complex (Purring-Koch & McLendon, 2000), dGK might assist in amplifying the apoptotic cascade (Jüllig & Eriksson, 2001).

dGK accepts the three naturally occurring purine deoxyribonucleosides and purine analogues as substrates – K_m values for dGuo, dAdo and dIno are respectively 2.5-7.6 μM, 60 μM, and 21 μM (Arnér & Eriksson, 1995), and the k_{cat}/K_m value is 6×10^4 $M^{-1}sec^{-1}$ for dGuo , and 4×10^3 $M^{-1}sec^{-1}$ for dAdo (Eriksson *et al.*, 2002). dCyd can also be accepted as a substrate, but with very limited efficiency (Wang *et al.*, 1996; Sjöberg *et al.*, 1998), the k_{cat}/K_m value is 1×10^3 $M^{-1}sec^{-1}$ (Eriksson *et al.*, 2002). Similarly to dCK, the recombinant dGK can use both ATP and UTP, as well as the inorganic polyphosphate (PPP_i) as phosphate donors (Krawiec *et al.*, 2003), and UTP improved the phosphorylation efficiency in comparison with ATP for all dGK substrates investigated (Zhu *et al.*, 1998; Sjöberg *et al.*, 2001). The end products dGTP and dATP are feed-back inhibitors (Gower *et al.*, 1979; Sjöberg *et al.*, 2001).

As can be seen, the substrate specificity of dGK is partially overlapping with dCK, but dGK is more restrictive. dGK accepts bases with minor 2-or 7-substitutions (for example CdA, *i. e.* 2-chloro-2'-deoxyadenosine), and sugars with 2'-substitutions and arabinofuranosyl (for example AraA and AraG). But neither 3'-substitutions, nor acyclic sugars are allowed (Wang *et al.*, 1993; Sjöberg *et al.*, 1998). The crystal structure of human dGK has been determined (Johansson *et al.*, 2001), and the reason for its restrictive substrate specificity was suggested. Presence of Arg-118 in the active site of dGK provides hydrogen bonds to ring N7 and carbonyl O6 of guanine, while an amino group at 6-position of adenine makes the binding of this base less favourable. Arg-118 is held tightly in place by Asp-147, and the presence of Arg-Asp couple is not suitable for binding dT with a methyl group in this position (Johansson *et al.*, 2001). The human dGK was present as a dimer in the crystals (Johansson *et al.*, 2001), in accordance to the previously determined dimeric structure of native bovine dGK (Wang *et al.*, 1993).

Recently, several mutations in the dGK gene were ascribed to be the cause of the hepatocerebral form of MDS, characterized by liver failure and neurological abnormalities (Mandel *et al.*. 2001; Salviati *et al.*, 2002; Taanman *et al.*, 2002; Mancusso *et al.*, 2003). Four mutations resulted in premature termination of translation and absence of dGK activity in liver homogenates. One patient had greatly decreased dGK activity and was a heterozygote for two missense mutations (R142K and E227K) affecting critical residues of dGK (Salviati *et al.*, 2002). The effect of the mutations on the activity of dGK has been analysed in recombinant enzymes (Wang & Eriksson, 2003). 11-, 18- and 22 amino acids C-terminal truncations led to inactive proteins, R142K had very low activity with dGuo (about 8% of the wild type), and no activity with dAdo, while E227K had unchanged K_m values but very low V_{max} values for all its substrates, resulting in 2% catalytic efficiency of the wild type (Wang & Eriksson, 2003). Even with such low dGK activity, the clinical symptoms were milder than in patients totally devoid of dGK activity, and the disease developed later in life. Experiments with one of the patients dGK-deficient fibroblasts have shown that decline in mtDNA content can be prevented by dGMP and dAMP supplemented to the growth medium (Taanman *et al.*, 2003). The described correlations between activity of dGK, availability of deoxynucleotides and disease underline the crucial role of salvage pathways.

Compartmentalization of deoxynucleoside kinases and dNTPs pools

De novo synthesis of deoxyribonucleotides occurs in the cytoplasm and the dNTPs pass freely into the nucleus (Leeds *et al.*, 1985; Leeds & Mathews, 1987; Wawra, 1988). The dNTP pools of nucleus and cytoplasm are in equilibrium because the nuclear membrane allows diffusion of ions and molecules up to 60 kDa through nuclear pore complexes (Nigg, 1997), and mediates the selective transport of macromolecules bearing the specific nuclear localization signal (Wente, 2000; Rout & Aitchison; 2001).

The mitochondrial membrane, on the contrary, is not permeable to charged molecules such as nucleotides, and several studies concluded that mitochondrial dNTPs are physically separated from cytoplasmic dNTPs (Bogenhagen & Clayton, 1976; Bestwick *et al.*, 1982).

Zhu *et al.* (2000) have shown that nucleoside analogues phosphorylated in the cytosol or nucleus were predominantly incorporated into nuclear DNA, while nucleoside analogues phosphorylated in mitochondria became trapped in this compartment, and were incorporated into mitochondrial DNA. The mitochondrial dGK and TK2 phosphorylate all four deoxyribonucleosides and may contribute most of the deoxyribonucleotides required for mitochondrial DNA replication and repair. A single report describes presence of ribonucleotide reductase activity in extracts of washed mitochondria (Young *et al.*, 1994a), but the report has not been followed up, and the gene for mitochondrial ribonucleotide reductase is not known. On the other hand, the results of Chen & Cheng (1992) indicated that the anti-HIV compound, ddC, was phosphorylated in the cytoplasm by the cytoplasmic dCK, and then, as ddCTP was imported into mitochondria. Also, mitochondria isolated from rat liver were shown to possess the capacity for inward transport of thymidine and 2'-deoxyguanosine (Cass *et al.*, 1999 and references therein). Purification of a carrier protein transporting dCTP from cytoplasm to mitochondria (Bridges *et al.*, 1999) supports the possibility that cytoplasmic dNTPs and phosphorylated nucleoside analogues could be utilized by mitochondria. Another mitochondrial deoxynucleotide carrier protein (DNC) was shown to catalyze transport of all four dNDPs (also NDPs and dNTPs, but with lower affinities) and their dideoxyanalogues from cytoplasm into mitochondria in exchange for ATP or ADP (Dolce *et al.*, 2001). A mutated DNC protein, devoid of transport activity, was associated with a congenital disorder, Amish microcephaly (Rosenberg *et al.*, 2002). Recently, a deoxynucleoside transporter protein has been localized to the mitochondrial membrane, and found to enhance the mitochondrial toxicity of nucleoside drugs (Lai *et al.*, 2004). The earlier view of physically separated cytoplasmic contra mitochondrial nucleotide pools must be revisited, especially in the light of recent evidence that the two compartments are in rapid exchange: *de novo* synthesized thymidine phosphates enter rapidly from cytosol to mitochondria, and thymidine phosphates phosphorylated inside mitochondria move rapidly into the cytosol (Pontarin *et al.*, 2003; Rampazzo *et al.*, 2004).

The dNTP pool sizes

A knowledge of physiological metabolite concentration is necessary in order to understand the kinetics of enzymes involved. It is also important for pharmacological studies, because analogues compete with the natural compounds *in vivo*. Concentrations of dNTPs have been determined by different methods, with different cells or tissues (most rodent and human), quiescent or growing, and of course with very different results (summarized by Munch-Petersen, 1996). For example, mouse embryo cells were determined to contain 2.5 pmol of dATP and 3.0 pmol dTTP per μg of DNA during DNA synthesis phase, and the pools were calculated to suffice for 2 to 4 minutes of DNA synthesis (Nordenskjöld *et al.*, 1970; Reichard, 1988). dNTP pools in quiescent G0 mouse L cells were approximately 0.2, 0.1, 0.02, and 1.2 (in pmol/μg DNA) for dATP, dTTP, dGTP and dCTP respectively, and for dividing cells 0.5, 2.0, 0.2, and 7 respectively (Skoog *et al.*, 1974). In G1 Chinese hamster ovarian cells pool levels were 2.8, 2.0, 0.2, and 5 pmol/μg DNA, respectively, and their amounts in S-phase cells varied between 4 (dATP) and 7 (dCTP) times those observed in G1 cells (Skoog *et al.*, 1973). In spite of the different results, there was a general correlation that the pools increase with the rate of DNA synthesis, and that the dCTP pool was the highest and the dGTP pool the lowest in the studied rodent cell lines. However, in the phytohemagglutinin (PHA) stimulated human lymphocytes the dTTP pool, and not dCTP pool, was consistently found to be the highest (Munch-Petersen *et al.*, 1973; Tyrsted, 1975; Tyrsted & Munch-Petersen, 1977). The dGTP pool seemed to be the lowest in both the unstimulated and PHA-stimulated human lymphocytes (Tyrsted, 1975). The natural asymmetry in the dNTP pool sizes has been repeatedly observed in cultured cells (Mathews & Ji, 1992), and also in extracts of whole rat embryos (Mole *et al.*, 1998), with dGTP representing 5 – 10% of the sum of the four dNTP pools, and dTTP representing 40 – 60%. Recent estimates of cytosolic dNTP pools in human and mouse cell lines showed that that the two pyrimidine pools were always larger than the two purine pools: dTTP + dCTP constituted 70% of all dNTPs, and dGTP only 7% (Rampazzo *et al.*, 2004). A possible explanation for this asymmetry was suggested by Chimploy *et* al., (2000) based on experiments showing that the activity of mouse ribonucleotide reductase toward GDP reduction was more sensitive to oxygen depletion than reduction of the three

other ribonucleotides. Due to natural oxygen limitation *in vivo*, this could lead to underrepresentation of dGTP in mammalian cells.

Traut (1994a) has compared concentrations of bases, nucleosides and nucleotides from 600 published values for mammalian systems. For deoxynucleotides in dividing cells the following average concentrations ± SD are given: dATP, 24 ± 22; dTTP, 37 ± 30; dGTP, 5.2 ± 4.5; and dCTP, 29 ± 19. The concentration in tumour cells was 6-11 fold over normal cells, and the concentrations in the extracellular fluids were usually lower than corresponding intracellular concentrations (Traut, 1994a).

Recently, the mitochondrial dNTP pools have been estimated to amount between 3 and 7% of the corresponding cytosolic pools: 3% of dTTP and dCTP, 4% of dATP, and 7% of dGTP (Rampazzo *et al.*, 2004).

Significance of dNTPs pool imbalance

The salvage pathway enzymes, together with the *de novo* pathway are responsible for the maintenance of balanced dNTP pools. Deficiencies of salvage enzymes or exogenous addition of deoxyribonucleosides may unbalance the corresponding dNTP pool and result in increased mutagenesis (Weinberg *et al.*, 1981; Meuth, 1989; Wilkinson & McKenna, 1989; Lei *et al.*, 1998a), influence the sensitivity to mutagens (Kunz, 1988; Kunz & Kohalmi, 1991), and on the cellular level induce chromosome rearrangements, breakage and loss, and oncogenic transformation (Kunz *et al*, 1994; Lei *et al.*, 1998b). Thymidine kinase deficient cell lines, that showed significant decrease in dTTP pools, were more sensitive to the mutagenic and lethal effect of alkylating agents than the wild type strain (Wakazono *et al.*, 1996). Similar results were obtained with TK deficient Raji cells (Best *et al.*, 1997), and TK deficient Friend mouse leukaemia cells (McKenna *et al.*, 1991). that showed decreased survival and increased mutant frequency relative to wild type cells following mutagen treatment. In the same investigation, McKenna *et al.* (1991) have found a greater than 6-fold increase in the dCTP : dTTP ratio in TK deficient cells and suggested a role for thymidine kinase in balancing dNTP pools. Increase of intracellular dTTP pool by exogenous addition of thymidine resulted in increased mutation frequency at the

adenine phosphoribosyl trasferase *(aprt)* locus, which was relieved by addition of deoxycytidine to the medium (Goncalves, *et al.*, 1984). More than 97% of mutations occurring at the *aprt* locus in the presence of excess dTTP were explained by the misincorporation of the dTTP in place of dCTP, the mutations were predominantly C to T transitions and G to T transversions (Phear & Meuth, 1989). Interestingly, analysis of the spontaneous mutation spectra at the *(aprt)* gene locus in TK1-deficient Friend murine eryhroleukemia cells has revealed increased mutagenicity, but has not shown mis-incorporation of dCTP during DNA replication (Hyland *et al.*, 2000). The predominant mutations found were GC to AT transitions, and because the TK$^-$ cells may have a reduced activity of the uracil-DNA glycosylase, their occurrence was suggested to result from deamination of the cytosine and the reduced efficiency of repair by uracil-DNA glycosylase of uracil residues that arise following cytosine deamination (Hyland *et al.*, 2000).

Molecular analysis of mutations in the *hprt* (hypoxanthine-guanine phosphoribosyltransferase) gene have shown that unbalanced dNTP pools can affect RNA splicing by causing base substitutions in the splicing consensus sequence and aberrant processing of mRNA (Darè *et al.*, 1995). Pool imbalances induced by addition of thymidine and hydroxyurea to cultured dog cells infected with retroviruses, spleen necrosis- and murine leukaemia virus, were found to increase the retrovirus mutation frequency (Julias & Pathak, 1998). This finding may have implications for treatment of HIV patients with hydroxyurea (Julias & Pathak, 1998).

Regulation of the dNTP supply was also suggested to be a control point in apoptosis (Oliver *et al.*, 1997). Withdrawal of the cytokine interleukin-3 (IL-3) from haemopoietic cells resulted in rapid decrease of dTTP, dGTP and dATP levels, decrease in thymidine kinase activity and in precursor uptake, and was followed by DNA fragmentation and apoptosis. Readdition of IL-3 resulted in a rapid restoration of normal dNTP pools and enhanced TK activity (Oliver *et al.*, 1996). Overexpression of herpes simplex virus-1 thymidine kinase in the haemopoietic cells protected them from apoptosis induced by either inhibitors of dNTP synthesis, or by IL-3 deprivation (Oliver *et al.*, 1997). The results suggested that TK may be important for the regulation of apoptosis in haemopoietic cells. As mentioned above, a role for dGK in amplifying the apoptotic cascade also has been proposed (Jüllig & Eriksson, 2001).

Folate deficiency, which is the most common vitamin deficiency, is also associated with imbalanced dNTP pools (Koury & Horne, 1994). The folate-deficient proerythroblasts underwent apoptosis when cultured in folate-free medium, the apoptosis was prevented by addition of either folic acid or thymidine. While folate concentration in excess had no effect on proliferation of erythroblasts, addition of more than 10 µM thymidine increased apoptosis. Simultaneous addition of deoxycytidine prevented apoptosis and indicated that imbalanced deoxypyrimidine pools were detrimental to erythroblasts (Koury & Horne, 1994).

The severe diseases that result from the lack of the salvage enzymes adenine deaminase, thymidine phosphorylase, and the mitochondrial TK2 and dGK have already been mentioned above. The hypothesis that imbalanced mitochondrial dNTP pools are responsible for mutagenesis of mtDNA has been tested and confirmed by the finding of deletions in mtDNA of HeLa cells growing for 8 months in thymidine-supplemented medium (Song et al., 2003).

Role of thymidine kinase 1 for regulation of dNTP pools – a short summary

TK1 phosphorylates thymidine to thymidylate (dTMP) and deoxyuridine to deoxyuridylate (dUMP) that subsequently gets methylated to thymidylate (Fig. 1.4). Due to the ubiquitous nucleotide kinases, dTMP is quickly phosphorylated to dTTP, and so thymidine kinase is an important enzyme for the synthesis of dTTP. Also, TK1 regulates the size of dTTP pool through "substrate cycles" involving phosphorylation of deoxynucleosides by their corresponding kinases and dephosphorylation by nucleotidases (Fig. 1.5). dTTP, in turn, is an important effector controlling the specificity of ribonucleotide reductase: increase in dTTP pool down-regulates the reduction of pyrimidine ribonucleotides and shifts the specificity towards reduction of purine ribonucleotides (Fig. 1.3). A cell with deficient TK1 activity has thus diminished ability to control its dTTP pool and to balance its supply of all four DNA precursors.

Chapter 2

The mammalian cell cycle

Since the activity of TK1 fluctuates with the cell cycle, I will first review the general mechanisms that control the events of the cell cycle. In the next chapter I will describe the regulatory mechanisms of TK1 expression. Here, I will concentrate on the features of the cell cycle that are relevant for induction of TK1 in mammalian cells. The nomenclature generally used for mammalian proteins will also be used here.

The cell cycle is characterized by two ordered events essential for cellular reproduction: duplication of the genome and cell division. Four distinct phases are recognized in the cell cycle: G1 - growth phase, S phase - the period of DNA synthesis, G2 - gap phase between the S and M phases, and M phase - mitosis, where sister chromatids segregate to opposite poles of the cell. These phases are followed by the physical division of the cytoplasm, the cytokinesis, resulting in two identical daughter cells with 2N DNA content. A G1 cell may continue to grow and proliferate, or, if growth factors or nutrients are insufficient, it may become quiescent and enter a G0 stage. Quiescent (G0) cells are either differentiated to perform some specific function, or, especially in cell culture, they become senescent and eventually die. Some G0 cells, like peripheral blood lymphocytes or liver cells, have the potential to re-enter the cell cycle if appropriately stimulated. Lymphocytes for example become mitotic after exposure to specific antigens. Terminally differentiated cells, like nerve cells or muscle cells, are likewise in non-dividing stage, but they have no potential to re-enter the cell cycle.

The commitment to progress through G1 and to enter the S phase requires accumulation of specific proteins and occurs late in G1 at a "restriction point", or R-point, as originally proposed by Arthur Pardee (1974). After passage of the R-point neither mitogenic signals nor intracellular protein synthesis are required by cells for further progression through the cell cycle.

Transition from G0 to G1

Most cells of an adult mammalian organism are quiescent cells. They are smaller than G1 cells and their rate of macromolecular synthesis is about one-third as that of G1 cells (Pardee, 1989). Also, nutrient deprived non-cancer cells in culture are G0 cells. Addition of growth factors, such as PDGF (platelet-derived growth factor), EGF (epidermal growth factor), IGF-1 (insulin-like growth factor-1) in a defined sequence, activates G0 cells to progress into G1. Very early changes that occur in the growth factor stimulated cell include increased transport of nutrients through the membrane, changes in chromatin structure and production of novel mRNA (Pardee, 1989). G1 for cells leaving G0, especially if they were in G0 for a long time, is much longer than for normally cycling cells, as synthesis of mRNA and proteins needed for the next cell cycle must be enhanced (Hofbauer & Denhardt, 1991).

Growth factors act by binding to their specific receptors on the cell membrane. This binding activates the receptor's tyrosine kinase activity and subsequently leads to activation of a GTP-binding protein Ras[1]. Ras resides at the inner surface of the plasma membrane and provides a link between the cell surface and the nucleus (Vojtek & Der, 1998; Ewen, 2000). Ras proteins switch between inactive GDP- and active GTP- bound states, the switch is catalysed by additional cytoplasmic signalling proteins. Activated Ras, in turn, induces a phosphorylation cascade that involves Raf kinase and MAPK (mitogen activated protein kinases) pathway (Schaeffer & Weber, 1999; Chang & Karin, 2001). Some of the phosphorylated MAPKs translocate into the nucleus (Fukuda *et al.*, 1997; Khokhlatchev *et al.*, 1998) where they, directly or indirectly, phosphorylate several different transcription factors, among others, the products of the proto-oncogens *c-jun*, *c-fos* and *c-myc* (Kerkhoff & Rapp, 1998; Ewen, 2000).

---

[1] There are multiple Ras proteins, different ones act in different cell types, and all work in the same way.

G1 phase: early- and late-response genes

c-jun, c-fos and *c-myc* are members of a set of "early-response" genes induced within the first few minutes (immediate early) or hours (delayed early) after entrance to G1 phase. Induction of early-response genes is independent of protein synthesis, because early-response proteins (often transcription factors) are present in G0 cells and are activated by post-translational modification such as phosphorylation. Fos and Jun belong to the bZIP (basic region leucine zipper proteins) family of transcription factors that bind to palindromic AP-1 regulatory elements and activate many genes involved in cell growth and differentiation (Karin *et al.*, 1997; Shaulian & Karin, 2001). Fos-Jun heterodimers interact with different transcription factors and can bind in opposite orientations to AP-1 sites, and in this way alter promoter selectivity and transcriptional activity of the transcription complexes (Ramirez-Carrozzi & Kerppola, 2001). Among the proteins regulated by AP-1 are cell cycle regulators like cyclin D1, p53, p21 (cip1/waf1) and p19 (ARF).

The *c-myc* gene product is a transcription factor that contains a basic-helix-loop-helix-leucine zipper dimerization domain (Dang, 1999). Myc upregulates activity of cyclin E/CDK2, antagonizes the function of p27^{Kip1} (Obaya *et al.*, 1999; Obaya *et al.*, 2002) and stimulates E2F-dependent transcription (Beier *et al.*, 2000a). The action of Myc in activation of cyclin E/CDK2 and E2F has been found to synergize with the action of Ras (Leone *et al.*, 1997).

Proteins encoded by early-response genes induce the transcription and expression of "late-response" genes, which appear later in G1 and are not expressed if protein synthesis is inhibited (Hofbauer &. Denhardt, 1991). Many of the late-response genes encode proteins that are linked to DNA replication – histones, thymidine kinase, thymidylate synthase, dihydrofolate reductase and ribonucleotide reductase. Transcription factor E2F, G1 phase cyclins and cyclin dependent kinases also belong to the late-response proteins. Besides the requirement for protein synthesis, the late-response genes require continuous presence of growth factors for their transcription. Withdrawal of the growth factors before expression of the late-response genes prevents the cells to pass the restriction point and to enter a new cell division.

Proteins coordinating the progression of the cell cycle

The key enzymes that regulate the transition between the various phases of the cell cycle are cyclin-dependent kinases (CDKs). Several mechanism are known that regulate their activity: interaction with cyclins, regulation of cyclin gene expression, post-translational modification of CDKs by phosphorylation-dephosphorylation, interaction of cyclin/CDK complexes with CDK inhibitor proteins (CKIs) and degradation of the intrinsic cell cycle proteins by the ubiquitin/proteasome pathway.

Cyclins and cyclin- dependent kinases during the cell cycle

Binding of a cyclin subunit to a CDK is usually the primary mechanism of CDK activation and a prerequisite for phosphorylation/dephosphorylation reactions. A cyclin/CDK complex is only partially active. It attains full activity after phosphorylation of the CDKs on specific conserved Thr residues by a CDK-activating kinase CAK and its partner cyclin H (Martinez *et al.*, 1997), followed by removal of inhibitory phosphates on Tyr or Thr residues by a phosphatase from the Cdc 25 family (Jinno *et al.*, 1994). The level of CDKs remain constant during the cell cycle, but cyclins are unstable and fluctuate in abundance during the cell cycle. In higher eukaryotes, at least eight different cyclin classes (cyclins A to H) have been identified based on sequence similarity and association with different stages of the cell cycle (Noble *et al.*, 1997). I will concentrate here on the best known and mostly investigated cyclin/CDK complexes (Fig. 2.1).
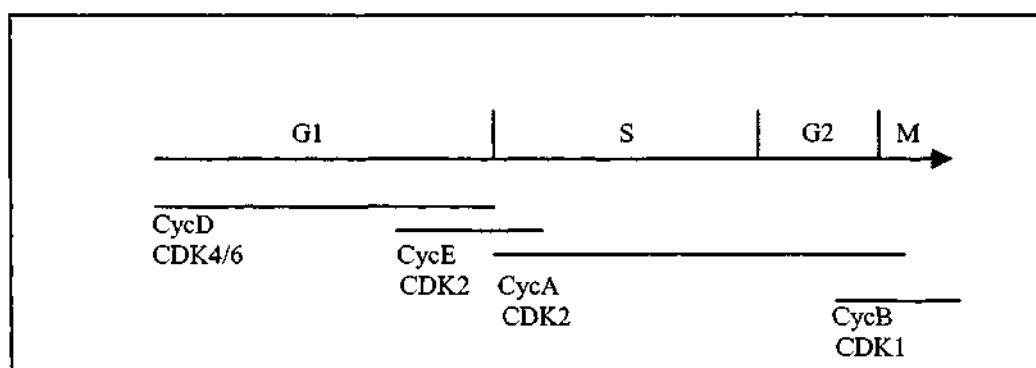
Figure 2.1. Major CDK/cyclin complexes involved in cell cycle control in mammalian cells. Modified after Morgan (1997).

The G1 progression in mammalian cells depends on three D-type cyclins (D1, D2 and D3), which assemble with either CDK4 or CDK6, and on cyclin E, which combines later in G1 with CDK2 (Sherr, 1994). The cyclin D1 promoter has been shown to be activated by bZIP superfamily of transcription factors, including c-Jun, c-Fos, ATF2 (activating transcription factor) and ATF3 (Beier, *et al.*, 1999a; Allan *et al.*, 2001). Cyclin D2 was found to be transcriptionally induced by Myc binding to a highly conserved E-box element in D2 promoter (Bouchard *et al.*, 1999).

The synthesis of D-type cyclins and their assembly with CDK4/6 depends on continuous mitogenic stimulation. Mitogen withdrawal leads to cessation of cyclin D synthesis, and because cyclins are unstable, their holoenzyme activities decay rapidly and the cells exit the cell cycle (Sherr, 1996). Absence of cyclin D-dependent kinase activity after the cells passed restriction point is without effect. Because the major substrate of cyclin D/CDK4 or cyclin D/CDK6 complexes is the Retinoblastoma protein (Rb) and the related Rb pocket proteins, p107 and p130 (Kato *et al.*, 1993), it was concluded that phosphorylated Rb proteins are required for exit from G1 (Sherr, 1996).

The synthesis of cyclin E is maximal at the G1 to S phase transition, and cyclin E associated CDK2 activity shifts pRb phosphorylation from being mitogen-dependent (cyclin D-driven) to mitogen-independent (cyclin E-driven) (Sherr, 1996). Cyclin E/CDK2 phosphorylates the Rb proteins on different sites than cyclin D/CDK4/CDK6; phosphorylation on both CDK2 and CDK4/6 sites is required for full inactivation of the Rb proteins (Kitagawa *et al.*, 1996; Zarkowska & Mittnacht, 1997). Recently, a controversial report demonstrated some cancer cells to proliferate when cyclin E/CDK2 activity was inhibited by overexpression of inhibitor protein p27 (Tetsu & McCormick, 2003), thus questioning the essential role of cyclin E/CDK2 in cell cycle control (Grim & Clurman, 2003).

Induction of cyclin A has been shown to depend on the presence of cyclin E activity (Rudolph *et* al., 1996). Cyclin A is first detectable during late G1, and functions in S-phase, G2 and mitosis. Once cells enter S phase, cyclin D and E are degraded and CDK2 forms complexes with cyclin A (Sherr, 1994; Sherr, 1996). To obtain maximal expression of cyclin A, its promoter had to be activated by transcription factors ATF2 and E2F (Beier *et al.*, 2000b). In the S-phase, cyclin A/CDK2 is believed to be important for initiation of DNA replication and to restrict the initiation to only once per cell cycle (Yam *et al.*, 2002).

Cyclin A/CDK2 activity appears to be the major cyclin/CDK activity in G2 and this activity was demonstrated to be limiting for the initiation of mitosis and for progression through mitosis until the middle of prophase (Furuno et al., 1999) when chromosomes condensate. Recently, several novel proteins interacting with cyclin A/CDK2 have been identified by a yeast triple-hybrid approach, co-immunoprecipitation and GST pull-down assays (Diederichs et al., 2004). These proteins link the cyclin A/CDK2 complex to diverse cellular processes such as DNA repair, signalling and splicing (Diederichs et al., 2004). Cyclin A also controls the half-life and degradation of cyclin B: Accumulation of cyclin B is dependent on cyclin A (Lukas et al., 1999). Levels of cyclin B increase in G2 and activity of cyclin B/CDK1 is crucial for G2/M progression and for progress of mitosis, especially for the breakdown of the nuclear envelope and formation of the mitotic spindle. An important regulator of cyclin B is its activator cdc25 phosphatase, that in turn is regulated by cytoplasmic protein 14-3-3 (Forrest & Gabrielli, 2001). The protein 14-3-3 might prevent mitosis by sequestering Cdc25 in the cytoplasm and thus separating it from its substrate cyclin/CDK1 kinase (Pines, 1999), and/or by blocking access of the substrate to the catalytic site of Cdc25 (Forrest & Gabrielli, 2001). Dephosphorylation of cyclin B/CDK1, and in consequence its activation, is inhibited by damaged or unreplicated DNA; in this way initiation of mitosis is inhibited and cells have time to repair the damage. Cyclin A is almost completely degraded before the metaphase to anaphase transition (Geley et al., 2001), while the degradation of cyclin B is delayed until cells enter anaphase and all of the chromosomes are correctly aligned on a well-formed mitotic spindle (Geley et al., 2001). Degradation of the mitotic cyclins triggers exit from mitosis and starts cytokinesis.

In summary, G1 specific CDKs activate S phase specific CDKs and promote synthesis of proteins needed for chromosome replication. In turn, S phase specific events trigger mitosis and cell division.

Cyclin-dependent kinases inhibitors (CKIs)

CKIs are small proteins that per definition impose negative control on activity of cyclin-dependent kinases. CKIs are currently divided into two families based on their structures and CDK targets (Sherr & Roberts, 1999). The first class consists of p21, p27 and p57 proteins, also known as Cip/Kip family, that can bind both to cyclin and CDK subunit of

cyclin D-, E-, and A-dependent kinases. The second class includes the INK4 proteins, that inhibit only G1 kinases - CDK4 and CDK6, but no other CDKs, and cannot bind to the cyclin subunit, as exemplified by p16 (Sherr & Roberts, 1999).

Initially it was assumed that CKIs act solely as inhibitors of cyclin-dependent kinases (Harper *et al.*, 1993; Peter & Herskowitz, 1994), later p21 was found in both catalytically active and inactive cyclin/CDK complexes (Zhang *et al.*, 1994), and was shown to promote the assembly of CDK4 with cyclin D, and activity of CDK4, but not the assembly or activity of CDK2/cyclin E (LaBaer *et al.*, 1997; Chang & McCubrey, 2001). Accordingly, low doses of γ-irradiation that induced p21, resulted in specific inhibition of CDK2, but not CDK4, phosphorylation activity (Brugarolas *et al.*, 1999). p21 was also shown to bind to PCNA (proliferating cell nuclear antigen), the processivity subunit of DNA polymerase δ, and thereby inhibit DNA replication, but not DNA repair (Chuang *et al.*, 1997).

Similarly to p21, p27 was found to be a positive regulator of the assembly, stability and nuclear localization of D-type cyclins, and a negative regulator of cyclin E-CDK2 and A-CDK2 (Cheng *et al.*, 1999). The preferential binding of p21 and p27 to D-cyclins keeps other cyclins, especially cyclin E, free from the inhibitors (Cheng *et al.*, 1999; Perez-Roger *et al.*, 1999). Mitogenic stimulation that induces CDK2 activity was shown to down-regulate p27 (Mohapatra *et al.*, 2001; Chang & McCubrey, 2001), and in mitogen-starved fibroblasts, p27, and to a minor extent p21, were found responsible for inhibition of CDK2 activity (Coats *et al.*, 1999). When expression of p27 was antisense inhibited, the cells did not respond to mitogen withdrawal by cell cycle arrest (Coats *et al.*, 1996). Low or absent p27 in tumour cells is an important clinical marker correlating with bad prognosis of disease (Clurman & Porter, 1998).

Recently, a hypothesis has been proposed that in parallel with their role in down-regulating the CDK activity, p21 and p27 inhibitor proteins might be targeted to gene promoters, interact with several transcription factors and in this way inhibit genes like Myc, E2F and AP-1, that are involved in cell cycle progression (Coqueret, 2003).

p16^{INK} accumulates progressively as cells age and approach the limit of their life span in vitro, possibly being induced by a senescence timer (Sherr & Roberts, 1999). p16 binds to CDK4/6, prevents its association with D-type cyclin and so phosphorylation of its substrate, the Retinoblastoma protein. CDK4/6 can thus be imagined to partition between the inactive INK4 bound state and the active Cip/Kip bound state (Sherr & Roberts, 1999).

Increased amounts of INK4 proteins would release Cip/Kip proteins from CK4/6 and can lead to cyclin E/CDK inhibition and to cell cycle arrest (Sherr & Roberts, 1999).

Retinoblastoma tumour suppressor protein and transcription factor E2F as substrates for cyclin /CDK complexes

The already mentioned Retinoblastoma tumour suppressor protein is one of the most important cyclin-CDK substrates for G1 progression. Early in G1, pRb and the related pocket proteins p107 and p130, are hypophosphorylated and bind to transcription factors of the E2F family. E2F proteins are heterodimers containing a subunit encoded by the E2F gene family, of which 7 are now known (E2F1-E2F7), and a subunit encoded by one of the three members of DP family (Dyson, 1998; Tamrakar et al., 2000; Bracken et al., 2004). Heterodimerization enhances the ability of E2F to bind to DNA and stabilizes the interaction between E2F and pRb (Kaelin, 1999). Because the transactivation domain of E2F1-E2F5 contains a pRb binding site, the E2F/Rb complex blocks transcription (Dyson, 1998; ; Bracken et al., 2004). The repression by pRb/E2F is also performed, at least in part, through recruitment of chromatin modifiers - histone deacetylase (HDAC), histone methyltransferase and DNA methyltransferase - to the E2F-regulated promoters (Luo et al., 1998; O'Connor et al., 2001; Bracken et al., 2004). Binding of these chromatin modifying enzymes prevents the binding of other proteins required for transactivation and promotes formation of nucleosomes that inhibit transcription (Luo et al., 1998; Brehm & Kouzarides, 1999; Bracken et al., 2004).

Phosphorylation of Rb and Rb family members (p107 and p130), at multiple sites during G1, first by cyclin D-dependent kinases, and later by cyclin E-dependent kinases, inactivates their capacity to interact with E2F transcription factors, thus allowing gene transcription and progression from the G1 to the S phase of the cell cycle. The disruption of E2F/Rb complex is also a common mechanism of action for oncoproteins encoded by DNA tumour viruses (Nevins, 1992). The purpose here is to enhance virus replication, and activation of both DHFR and thymidine kinase has been demonstrated in adenovirus and SV40-infected cells (Nevins, 1992).

The target genes of E2F include those whose products are important for DNA synthesis: dihydrofolate reductase, thymidine kinase, DNA polymerase α and PCNA (Dyson, 1998;

Brehm & Kouzarides, 1999), and also cell-cycle regulators like cyclin E, cyclin A and E2F itself. The phosphorylation of E2F by cyclin D-dependent kinase, release of free E2F and increase of cyclin E transcription results in a positive feedback loop that generates a rapid rise in E2F-dependent transcription and the onset of S phase (DeGregori *et al.*, 1995; Morgan, 1997). As cells leave S phase, the activity of E2F decreases due to its ubiquitin–directed degradation (Martelli & Livingston, 1999), and to phosphorylation of E2F/DP dimers on both subunits by cyclin A/CDK2, what inhibits DNA-binding activity of E2F (Dyson, 1998). Late in mitosis pRb is dephosphorylated by a holoenzyme complex containing protein phosphatase type 1 (PP1), thereby restoring the growth suppressive function of pRb (Rubin *et al.*, 1998). PP1 appears to be indispensable for mitotic exit (Tamrakar *et al.*, 2000).

Connection between $p16^{INK}$, $p21^{Cip}$ and tumour suppressors pRb and p53

The already mentioned $p16^{INK}$ causes hypophosphorylation of the Retinoblastoma protein, and so inactivation and prevention of G1 to S transition, through inhibition of cyclin D/CDK kinases. The genetic locus *INK4a* encodes, from a distinct promoter, another growth inhibitory protein, $p19^{ARF}$ (for alternative reading frame), that in turn controls the tumour suppressor protein p53 (Kamijo *et al.*, 1997). p53 is a key regulator in a wide range of cellular processes, including cell cycle control, apoptosis and DNA repair, and functions to integrate cellular responses to stress (Levine, 1997; May & May, 1999). p53 is a sequence specific transcriptional activator and among its transcriptional targets is the CDK inhibitor p21 (Levine, 1997). $p19^{ARF}$ is not a CDK inhibitor, but neutralizes oncoprotein Mdm2, an inhibitor of p53 (Kamijo *et al.*, 1998; Honda & Yasuda, 1999) and in this way leads to transcriptional activation of p21 and cell cycle arrest or apoptosis. The Mdm2 protein has been recently shown to interact physically with the p21 inhibitor and to facilitate proteasome-mediated p21 protein degradation independent of p53 (Zhang *et al.*, 2004).

When p53 upregulates p21 expression in response do DNA damage, p21 promotes pRb activation and cell cycle arrest in order to allow DNA to be repaired. Another link between p53 and pRb was pointed out by Bates *et al.* (1998): the transcription factor E2F1 was

shown to activate directly the expression of p14^{ARF}, the human analogue of p19^{ARF}, while the transactivation defective E2F1 mutants had no activity. The authors suggested that E2F1, a protein activated by cell-cycle progression, is part of fail-safe mechanism to protect against aberrant cell growth (Bates *et al.*, 1998). The human analogue of p19^{ARF} has been shown to interact physically with E2F1 and to inhibit its transcriptional activity (Karayan *et al.*, 2001). Thus, pARF can be viewed as a dual-acting tumour suppressor in both the p53 and pRb pathways.

Proteolysis in cell cycle control

Ubiquitin cascade

Proteolysis during the cell cycle is mediated by two distinct ubiquitin-conjugation pathways, each mediated by a different multisubunit protein complex . Ubiquitin is a small (76 residues) highly conserved protein, essentially identical from yeast to humans. The proteolytic pathways involve two successive steps: attachment of multiple ubiquitin molecules to the substrate protein, and degradation of the tagged protein by the 26S proteasome (Ciechanover, 1998, Ciechanover & Ben-Saadon, 2004). Conjugation of ubiquitin to the substrate proceeds via a three-step cascade: 1) ubiquitin is activated on its C-terminal by the ubiquitin-activating enzyme, E1. 2) several species of ubiquitin-conjugating (E2 or Ubc) enzymes, transfer the activated ubiquitin to a member of the ubiquitin ligase family (E3 or Ubl) enzymes. 3) E3 catalyzes the covalent attachment of ubiquitin to a lysine residue of the substrate protein. In most cases, the Lys residues that serve as ubiquitination sites are not specific or part of the recognition motif (Ciechanover, 1998). Recent findings indicated, that for several proteins (for example the inhibitor p21), the first ubiquitin moiety is fused not to an internal lysine, but to the α-NH$_2$ group of the N-terminal residue (Ciechanover & Ben-Saadon, 2004).

Each E2 acts with several E3s, interactions among different E2s and different E3s generate a large number of specific substrate targeting complexes (DeSalle & Pagano, 2001). The human genome contains ~1000 different E3 ligases (Ciechanover & Ben-Saadon, 2004). As shown by *in vitro* reconstitution experiments, E1, E2, E3 enzymes and the 26S proteasome are the minimal requirements for protein degradation. Several novel accessory

proteins have been described that recognize and bind to both the multi-ubiquitin chain and the proteasome and optimise degradation of many substrates (Hartmann-Petersen *et al.*, 2003). Probably the only function of polyubiquitin chain is targeting and tethering substrates to the proteasome, as it was demonstrated that localization of a substrate to the purified proteasome was sufficient for degradation in a ubiquitin-free *in vitro* system (Janse *et al.*, 2004).

SCF complex

One of the ubiquitin-conjugation protein complexes, best characterized in budding yeast, promotes passage from G1 to S phase by initiating DNA replication through degradation of S phase cyclins inhibitor (SIC in yeast, p27 in mammals). This ligase complex is designated SCF from the names of the major compound proteins: Skp1-cullin/Cdc53-Fbox (Fig.2.2).
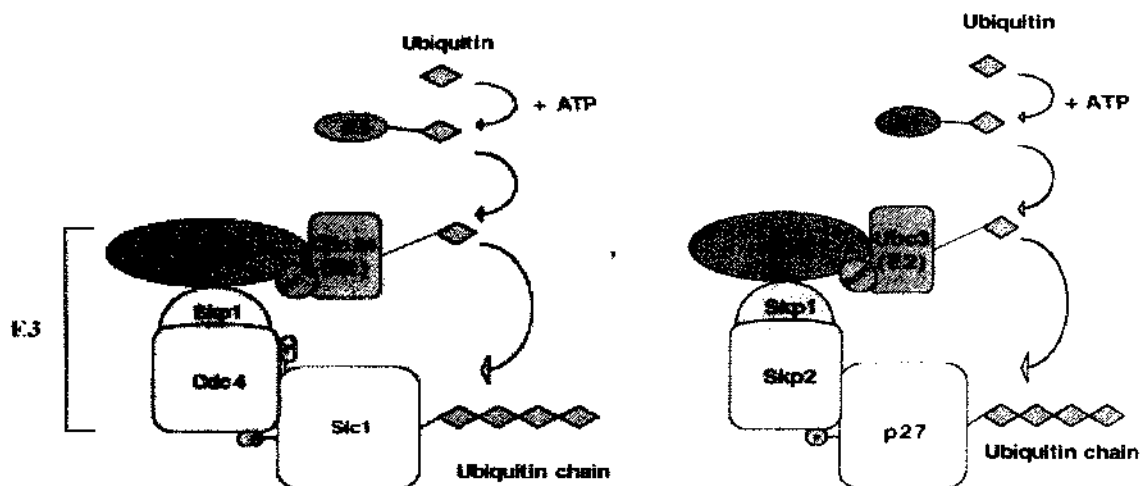
Figure 2.2. Schematic representation of yeast SCF complexed with E2 Cdc4 (left) and mammalian SCF complexed with E2 Ubc3, each binding their specific phosphorylated substrates: Cdc4 binds Sic1, and Skp2 binds p27. After DeSalle & Pagano (2001).

SCF ligases also contain a small protein subunit called Hrt1, Rbx1 or Roc1, that contain a conserved C-terminal RING finger[2], a cysteine-rich fold surrounding zinc atoms, through which Cdc53/Cullin binds the E2 enzyme (Fig. 2.2; Tyers & Jorgensen, 2000; DeSalle & Pagano, 2001). The SCF compound proteins in yeast indispensable for SIC degradation

---

[2] The RING domain mediates protein-protein interaction and is distinct from Zn-finger motif of DNA binding proteins (Saurin *et al.*, 1996).

are: Cdc53, Skp1 and Cdc4. The Cdc53 ortholog[3] in human cells is the cullin protein CUL1. Cullins are highly conserved during evolution and have been identified in organisms from *C. elegans* to humans (Peters, 1998; Ciechanover, 1998). Cdc53/CUL1 acts as a docking site on which other proteins of the SCF complex assemble (Fig. 2.2; Peters, 1998; Obaya & Sedivy, 2002). Skp1 is a critical scaffold protein, it binds both to Cdc53/cullin and to so called F-box proteins. F-box proteins serve as receptors and define specificity for ubiquitination substrates. They are linked to Skp1 by common F-box motif, a conserved domain of approximately 40 amino acids (Tyers & Jorgensen, 2000). Some of the F-box proteins, as for example Cdc4 and Met30 in yeast, bind their substrates via Trp-Asp (WD) repeats, other as Grr1 in yeast via leucine rich repeats, and SKP2 protein in human uses still another domain (Fig. 2.2; Tyers & Jorgensen, 2000; Pfleger *et al.*, 2001 and references therein). Cdc53 and Skp1 are constitutive subunits of the SCF, Cdc4 is indispensable for degradation of SIC, but dispensable for degradation of G1 cyclins, while Grr1 is essential for G1 cyclin proteolysis (Peters, 1998 and references therein). The Cdc4 protein, was not required for the SCF mediated degradation of human TK1 in yeast cells, while functions of Cdc53 and Cdc34 were required (Ke *et al.*, 2003).

The SCF ubiquitin ligases are constitutively active, regulation occurs at the level of substrate recognition, where phosphorylation of the substrate is necessary to convert it to a form susceptible to the action of the SCF ligase complex. The SIC inhibitor must be phosphorylated on several of the nine putative phosphorylation sites to be recognized by Cdc4 and ubiquitinated (Harper, 2002 and references therein). On the contrary, in HeLa cells, phosphorylation of the p27 inhibitor on a single site was necessary and sufficient for interaction with the SKP2 protein and for ubiquitin-dependent degradation by the SCF complex (Tsvetkov *et al.*, 1999). SKP2 was also suggested to be the F-box receptor for other cellular regulators like cyclin D1, p21 and E2F-1 (DeSalle & Pagano, 2001, and references therein). Also, in ovarian cancer cells exposed to vitamin D_3, decreased abundance of SKP2 resulted in increased p27 stability and arrest of cells in G1 (Li *et al.*, 2004a).

---

[3] Genes in two separate species that derive from the same ancestral gene are said to be orthologs

APC complex

The second ubiquitin-conjugation pathway initiates chromosome segregation and exit from mitosis by degrading anaphase inhibitors (now called securins) and mitotic cyclins. It involves a large protein complex (over 10 different subunits) called the anaphase-promoting complex (APC) or cyclosome, that uses a different set of E2s, than those used by the SCF pathway (DeSalle & Pagano, 2001). There are notable structural similarities between SCF and APC subunits (Fig.2.3).
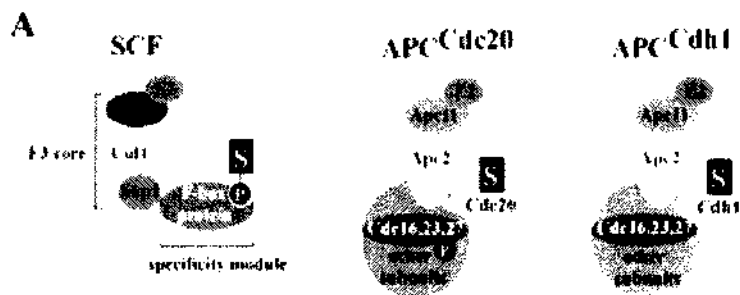
Figure 2.3. Structure comparison of SCF and APC complexes. Cullin-homology domain containing proteins are yellow, RING-H2 proteins red assemble with the E2, homologs of Skp1 adapter protein purple are unknown for APC. After Harper *et al.*, 2002.

The formation of the APC complex is nucleated on a cullin-related large protein called APC2, which shares a 180-residue cullin-homology domain with Cdc53 (Fig. 2.3; Obaya & Sedivy, 2002; Harper *et al.*, 2002). Two related WD (Trp-Asp) repeat proteins collaborate with APC and confer substrate specificity to APC-dependent degradation reactions: Cdc20 and Cdh1/Hct1 in yeast, or hCDC20 and hCDH1 in humans (Peters, 1998; Zachariae, 1999; Morgan, 1999 and references therein). Cdc20-APC has ubiquitin ligase activity towards proteins containing a nine amino acid motif RXXLXXXXN designated the "destruction box", or D-box (Pfleger *et al.*, 2001; Harper *et al.*, 2002 and references therein), while Cdh1-APC recognizes proteins that contain a D-box or a KEN box (Pfleger & Kirschner, 2000). The KEN box like the D-box is transposable to other proteins (Pfleger & Kirschner, 2000). Both Cdc20 and Cdh1 have been shown to bind the substrates via their N-termini, and not via the WD repeats (Pfleger *et al.*, 2001). Cdc20 is itself a Cdh1-APC substrate (Pfleger & Kirschner, 2000). Vertebrate Cdc20 lacks a D-box (though the budding yeast Cdc20 contains two D-boxes) and is recognized by a KEN

signal sequence. Mutations in the KEN box stabilized proteins against ubiquitination and degradation (Pfleger & Kirschner, 2000). Also, a KEN box located in the C-terminal of human TK1 has been found necessary for interaction with APC-Cdh1 and for the mitotic degradation of hTK1 (Ke & Chang, 2004).

APC and cell cycle phases

The amount of Cdc20 bound to the APC increases between S phase and mitosis, while Cdh1 is predominantly bound to the APC in G1 (Peters, 1998 and references therein). Binding of Cdc20 to APC leads to metaphase - anaphase transition through cleavage and dissociation of cohesin complex that holds the sister chromatids together. This allows them to separate and to move to opposite poles of the mitotic spindle (Fig. 2.4).
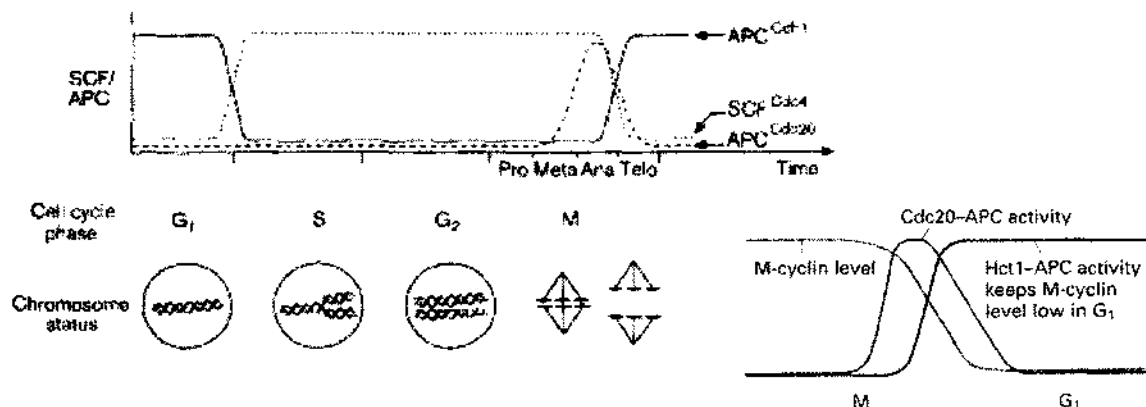
Figure 2.4. Model of cell cycle progression regulated by SCF and APC. Pro, prophase, Meta, metaphase, Ana, anaphase and Telo, telophase. After Peters, 1998. Inset: Continuous suppression of M-cyclin through inhibition of Cdc20-APC and activation of Hct1-APC (Hct1=Cdh1). From "The Cell" 4th Edition, Alberts et al., 2002.

The mitotic cyclin B/CDK1 complex stimulates Cdc20-APC activity kinase – phosphorylation of APC core subunits is required for activation of APC (Zachariae, 1999 and references therein). At the same time the active Cdc20-APC complex promotes destruction of the mitotic cyclins. Low level of the mitotic cyclins means low level of Cdc20-APC activity and rapid accumulation of new M phase cyclins for the next cycle. In order to allow for a G1 phase in the next cycle, Cdh1-APC complex starts to accumulate in late mitosis as Cdc20-APC destructs the mitotic cyclins - in contrast to the activity of Cdc20-APC, the activity of Cdh1-APC is inhibited by mitotic cyclins (Morgan, 1999 and references therein). The level of the mitotic cyclins is kept low after mitosis and during

next G1 by Cdh1-APC (Fig. 2.4). Unlike B-type cyclins, G1 cyclins are immune to the action of the APC and accumulate when Cdh1-APC activity is high.

The ability of Cdh1 to bind to APC inversely correlates with its phosphorylation state (Peters, 1998) - the cyclin proteolysis at the end of mitosis depends on Cdh1 dephosphorylation by a late mitotic gene product, the phosphatase Cdc14 (Morgan, 1999 and references therein). The accumulation of the mitotic cyclin B has been shown to require the E2F-dependent S-phase cyclin A/CDK2 mediated phosphorylation of Cdh1. This in turn prevents its assembly with APC, and blocks APC-dependent proteolysis allowing accumulation of APC targets (Lukas et al., 1999). Assembly of APC with a phosphorylation-deficient mutant of Cdh1 resulted in a complex active even in S-phase cells (Lukas et al., 1999).

Chapter 3

Regulation of TK1 expression

Several reports have demonstrated that TK1 enzyme activity is extremely low in resting cells, but is highly expressed in growing and malignant cells (Bello, 1974; Johnson *et al.*, 1982; Coppock & Pardee, 1985; Hallek *et al.*, 1992). The activity of TK1 has been found to exhibit a periodic increase pattern, starting to increase in G1, continuing to increase in S and G2, and decreasing at about the time of mitosis (Bello, 1974). Experiments using actinomycin D - inhibitor of RNA polymerase - and cycloheximide - inhibitor of protein synthesis, have shown that the increase of TK1 activity is due to *de novo* synthesis of TK1 mRNA and TK1 protein (Johnson *et al.*, 1982).

TK1 promoter and transcriptional regulation

TK1 activity and TK1 mRNA are not detectable in quiescent cells, but up to 40-fold increase in TK1 mRNA level has been shown after mitogenic stimulation or during SV40 virus infection (Stuart *et al.*, 1985; Coppock & Pardee, 1987; Roehl *et al.*, 1993). In PHA stimulated human lymphocytes the levels of TK1 activity were similarly elevated, the maximal increases reported varied between the different research groups from 20 fold to about 200 fold as compared to the unstimulated lymphocytes (references in Munch-Petersen, 1996). The level of TK1 mRNA was about 100-fold higher in PHA stimulated human lymphocytes than in quiescent lymphocytes (Kristensen *et al.*, 1994).

The significance of the human thymidine kinase gene promoter in cell cycle regulation has been demonstrated by Kim *et al.* (1988) and Travali *et al.* (1988). They have constructed hybrid genes containing the human TK1 promoter and a non cell cycle regulated reporter gene, respectively the bacterial neomycin resistance gene and the bacterial CAT (chloramphenicol acetyltransferase) gene. Expression of both bacterial genes in hamster cells was then cell cycle regulated and attained maximum as cells entered the DNA synthetic phase.

The functional promoter of the human thymidine kinase was delimited by Kreidberg & Kelly (1986) by construction of mutant plasmids with promoter deletions from both the 5' and 3' directions. The mutants were tested for the ability to transform mouse LTK⁻ cells to the TK⁺ phenotype. This analysis delimited the functional promoter to within a 83 base-pair region upstream of the mRNA cap site (Kreidberg & Kelly, 1986). This region contained the AT rich TATA box, about 20 bases upstream of the transcriptional start site where general transcription factors are bound and position RNA polymerase II for transcription initiation, several GC rich hexanucleotides involved in binding of the transcription factor SP1[1], and two CCAAT boxes in inverted orientation (Kreidberg & Kelly, 1986; Flemington et al., 1987). A schematic drawing of the regulatory elements known in the human TK1 promoter are shown in Fig. 3.1.

Figure 3.1. Regulatory elements of the human TK1 promoter. Modified after Kim et al. (1996) and Chang et al. (1999) with base numbers according to Chang et al. (1999). The sites corresponding to Yi binding sequences of Kim & Lee (1992) are the same as E2F-like binding sequences 2 and 3 of Kim et al. (1996).

---

[1] Sp1-responsive promoters usually contain multiple GC box recognition sites which are often found near binding sites for other transcription factors. This suggests that these factors act in conjunction with each other to modulate transcription (Kadonaga et al., 1987). DNA binding activity of Sp1 has been localized to three contiguous Zn finger motifs located in the C-terminal part of the protein (Kadonaga et al., 1987).

CCAAT boxes and protein NF-Y

The CCAAT boxes have been associated with transcriptional activation of genes in several systems (histone genes, Xenopus *hsp70* gene); cell cycle specific binding of nuclear proteins to the two CCAAT sequences in human TK1 promoter has been first described by Knight *et al.* (1987). Both CCAAT boxes were found important for the expression of the TK1 promoter, but the estimation of their contribution differed. In a series of experiments where TK1 promoter deletion mutants were fused with a non-cell cycle regulated gene, Roehl & Conrad (1990) and Kim & Lee (1991) identified a 70 bp region between −135 and −64 containing the distal −70 CCAAT motif, as a cell cycle regulatory unit sufficient to confer cell cycle regulation to a heterologous promoter. In a series of a similar TK1 promoter fusions, Arcot *et al.* (1989) have estimated that each CCAAT element is responsible for about 20% of the promoter strength, while the distal −70 element with one GC box maintains two-thirds of promoter activity. Mao *et al.* (1995) concluded that while the distal box could substitute for the proximal one, the converse was not true. On the other hand, transfection assays done by Chang *et al.* (1999) in HeLa cells showed, that whereas simultaneous mutations at both CCAAT boxes abolished the TK1 promoter activity, mutations at either the proximal or distal CCAAT box yielded almost 50% reduction of the promoter activity, suggesting that both boxes are functional and involved in activating the hTK1 promoter in HeLa cells (Chang *et al.*, 1999).

Arcot *et al.* (1989) demonstrated that a nuclear protein, suggested to be NF-Y, binds to both CCAAT elements, but has higher affinity for the proximal −40 element. Differential binding of a nuclear protein thought to be highly specific for the human TK1 promoter, and consequently named CBP/tk (CCAAT Binding Protein for TK gene), was also shown by Pang & Chen (1993). Here, the CBP/tk had higher affinity to the distal CCAAT box and this binding was dependent on growth phase of the cells. The DNA-protein complex was identified in young serum stimulated normal human fibroblasts, but neither in quiescent young fibroblasts, nor in old serum stimulated fibroblasts (Pang & Chen, 1993). In contrast, the expression of TK1 CCAAT-binding activity in cancer cell lines, like HeLa and promyeloleukemia HL-60 cells, was high independently of serum stimulation (Chang & Cheng, 1993). Later, this group identified the CCAAT binding protein as NF-Y using Western blot analysis and antibodies specific for the subunits of NF-Y: NF-Ya and NF-Yb (Chang & Liu, 1994). It was shown that expression of the NF-Ya subunit was high in

serum stimulated normal human fibroblasts, but absent in quiescent fibroblasts, while the expression of NF-Yb was the same before and after serum stimulation. Because both subunits[2] are required for DNA binding activity of NF-Y (Hatamochi *et al.*, 1988; Mantovani *et al.*, 1992), it was concluded that the altered expression of NF-Ya subunit after serum stimulation was responsible for the fluctuation of TK1-CCAAT binding activity in normal human fibroblasts (Chang & Liu, 1994) and so for S-phase specific transcription of TK1. Later experiments confirmed, that CBP/tk binding in human fibroblasts could be supershifted by antiserum against NF-Ya, but not by antiserum against p107, pRB, cyclin A, or a CDK2 protein (Good *et al.*, 1995), and that expression of the NF-Y was dependent on the age of human fibroblasts, being barely detectable in old cells (Good & Chen, 1996).

The sites binding NF-Y have also been shown to bind two other serum inducible proteins dbpA and dbpB, and a protein named CDP/cut[3], the CCAAT displacement protein with known transcriptional repressor activity (Kim *et al.*, 1997a). Overexpression of dbpA stimulated hTK1 promoter activity from the proximal CCAAT box, while CDP/cut could bind to both CCAAT boxes, and was found bound to the hTK1 promoter primarily in quiescent cells. Thus, a combination of suppression in quiescent cells and activation in serum stimulated cells mediated through various CCAAT binding proteins is responsible for the induction of TK1 promoter activity in the S-phase of cell cycle (Kim. *et al.*, 1997a).

Yi, E2F and Sp1 binding proteins

Another protein-binding site required for S-phase regulated transcription of human TK1, and different from the above mentioned CCAAT sequence, was identified by Kim & Lee (1992). This binding site sequence, located between –109 and –84 positions of the human TK1 promoter, contained two elements with striking resemblance to the murine protein Yi consensus binding site of the mouse promoter: CCCNCNNNCT, where N is any base (Dou

---

[2] Now it is known that NF-Y (also called CBP, CBF and CP1) consists of three different subunits A, B and C, and association of all three subunits is needed for DNA binding (Maity & Crombrugghe, 1998).
[3] The mammalian CDB/cut proteins are believed to compete for binding sites occupancy at CCAAT and Sp1 sites with activating transcription factors (Mailly *et al.*, 1996).

et al., 1991). Yi binding activity was not detected in G0 and G1 extracts, but was observed as the murine cells crossed the G1/S boundary (Dou *et al.*, 1991). Deletion of the Yi-related sequence, simultaneously with the E2F-like motif contained within the human TK1 promoter eliminated S-phase specific transcription of a reporter gene in Chinese hamster fibroblasts (Kim & Lee, 1992). Later, the Yi protein was renamed Yi2 and was found in a DNA binding protein complex related to the S phase specific E2F complex (Dou *et al.*, 1992). At the beginning of the S phase, the Yi2 complex replaced another complex named Yi1, bound to the same promoter sequence (Dou *et al.*, 1992; Dou *et al.*, 1994a). The Yi1 complex contained the murine retinoblastoma protein, cyclinD1 and CDK2 kinase, and its DNA binding sequence was different from the consensus binding sequence for E2F (Dou *et al.*, 1994a).

Binding sequences for the transcription factor E2F (Fig. 3.1) have been found both in the human (Nevins, 1992; Anderson *et al.*, 1996; Hengstschläger *et al.*, 1996) and in the mouse TK1 promoter (Dou *et al.*, 1994b). The murine E2F-like binding site was found to interact both with the human E2F protein and mouse nuclear protein complexes containing E2F, p107 (pRb-like), CDK2, cyclin A and cyclin E in a cell cycle dependent manner (Dou *et al.*, 1994b). Gel shift mobility studies with a −109 to −84 fragment of the human TK1 promoter, the fragment found critical for G1/S transition by Kim & Lee (1992), identified an S phase specific complex containing p107 and cyclin A, proteins that are known to interact with E2F (Li *et al.*, 1993). Though, not all results point to the activation of TK1 promoter by E2F. The overlapping E2F-like sites 2 and 3 at −97 to −84, in contrast to the E2F-like site 1 (Fig. 3.1), were found to be critical for the transcriptional activity of the human TK1 promoter and necessary for high specificity binding of nuclear protein complex to human TK1 promoter DNA affinity column (Kim *et al.*, 1996). But proteins in this complex were immunologically different from E2F1, E2F2, E2F3 and E2F4, what suggested presence of a novel protein with preferred binding for the human TK1 promoter (Kim *et al.*, 1996). Also, when activation of E2F target genes was measured by infection of quiescent cells with an adenovirus vector into which the E2F gene has been recombined, the TK1 gene expression increased only slightly in response to expression of E2F1, while genes encoding thymidylate synthase, proliferating cell nuclear antigen (PCNA) and ribonucleotide reductase were clearly induced (DeGregori *et al.*, 1995).

Tommasi & Pfeifer (1997) examined cell cycle specific binding of transcription factors at the human TK1 promoter by footprinting experiments in normal human fibroblasts, and identified 14 putative protein-binding sites. All of these sites, also the E2F recognition sequences, were constitutively occupied *in vivo* at all stages of the cell cycle. Similarly, both the E2F and the Sp1 binding sites were occupied throughout the cell cycle in mouse cells (Karlseder *et al.*, 1996), and because the C-terminal of Sp1 is necessary for interaction with the N-terminus of E2F protein, the TK1 promoter activity decreased with increasing distance between E2F site and Sp1 sites (Karlseder *et al.*, 1996). All four factors of the Sp1 family were able to bind E2F1 (Rotheneder *et al.*, 1999), but they differed in their ability to activate the mouse TK1 promoter: Sp1 and Sp3 were strong activators, Sp4 was a weaker one, and Sp2 showed no activation at all (Rotheneder *et al.*, 1999). In contrast to the mouse promoter, an E2F-like sequence in hamster TK1 promoter was found not to bind transcriptional factors of the E2F family, and its removal did not affect promoter activity (Sorensen & Wintersberger, 1999). The hamster promoter was regulated by a single Sp1 binding site and one CCAAT box, and in this promoter, NF-Y probably replaces the function of E2F (Sorensen & Wintersberger, 1999). Also Chang *et al.* (1999) have found that mutation at position -106/-102, the E2F-like site 1 in the human TK1 promoter, had little effect on the promoter activity in transfected HeLa cells, while mutation at the Sp1 binding site –118/-113, caused about 50% reduction of the human TK1 promoter activity (Chang *et al.*, 1999). It is then questionable to generally regard TK1 as an E2F regulated gene.

Comparison of the human, hamster and mouse TK1 promoters

Only the mouse TK1 promoter carries a genuine binding motif for E2F, and the other promoters have sequences that resemble it more (human) or less (hamster or rat) (Wintersberger, 1997). All three promoters have GC boxes for binding of Sp1, but the human and hamster promoters, in contrast to the mouse one, also carry a TATA box and CCAAT boxes for binding of NF-Y. The role of Yi complexes for the expression of the mouse promoter is not absolutely clear, but most probably the Yi1 is different from E2F, as they complex with different DNA binding proteins, and Yi1 complex contains D1/CDK2 kinase while E2F complex contains cyclin E/CDK2 and cyclin A/CDK2 kinases (Dou *et al.*, 1994a; Dou *et al.*, 1994b; Dou & Pardee, 1996).

Cyclins and CDKs

Complexes containing cyclin A, p107 and CDK2, that could bind to human TK1 promoter were also detected in the nuclear extracts from growth-stimulated hamster fibroblasts (Li *et al.*, 1993). CDKs and its regulators were found to contribute as well to the maximal activity of the human TK1 promoter. In IMR-90 normal human diploid fibroblasts the activity of the TK1 promoter increased by the constitutive overexpression of cyclin A or cyclin E, but not by cyclin D (Chang *et al.*, 1995). The sequence responsible for the transcriptional activation by cyclin E was located between −133 and -92 of the human TK1 promoter (Chang *et al.*, 1995). In HeLa cells, where the hTK1 promoter was already highly activated, no further activation could be obtained by ectopic expression of cyclin A or E (Chang *et al.*, 1995). Still, inhibition of the TK1 promoter activity by the tumour suppressor p16 could be overcome by overexpressing cyclin A in HeLa cells, and this implicated a direct role of cyclin A in activation of the TK1 promoter (Chang *et al.*, 1999). Also, activity of CDK2 was found to be required for the induction of TK1 mRNA in serum stimulated young fibroblasts, as presence of CDK2 inhibitor roscovitin abolished this induction (Chang & Huang, 2001). In contrast to the young cells, aging serum-induced fibroblasts showed no increase in the activity of CDK2, probably due to upregulation of p21 and absence of cyclin A expression (Chang & Huang, 2001). The absence of CDK2 activity resulted in 6-fold lower levels of TK1 mRNA than those in young cells, and in impairment of TK1 polypeptide synthesis (Chang & Chen, 1988; Chang & Huang, 2001). This is in agreement with results of Haidweger *et al.* (2001) that cyclin A/CDK complex phosphorylates Sp1 and that the phosphorylated Sp1 shows increased DNA binding ability which in turn enhances the expression from the Sp1 responsive promoter.

Post-transcriptional regulation

Nuclear run-on transcription assays have shown that the rate of TK1 transcription increases 2 to 7-fold between the G0 and S phases, and can therefore only partly explain the more than 40-fold increase of the level of mRNA (Coppock & Pardee, 1987; Stewart *et al.*, 1987). Gudas *et al.* (1993) synchronized hamster fibroblasts to G0, early G1 and mid G1, and investigated the mRNA level and transcription rate upon removal of the respective

proliferation block. Cells emerging from the quiescent G0 state showed a 4-fold increase in the rate of TK1 transcription, while cells restimulated from G1 arrest showed no change in the rate of TK1 transcription, but concomitant increase in mRNA abundance. The conclusion was that transcriptional and post-transcriptional control mechanisms involved in regulation of TK1 mRNA can be uncoupled, and that the transcriptional component is responsible for mRNA level during progression from G0 to G1, while the post-transcriptional component is responsible for mRNA level during progression from G1 to S phase (Gudas et al., 1993). Mechanisms of post-transcriptional regulation that account for the increase of TK1 mRNA and enzyme activity at the beginning of the S phase include increased mRNA stability (Coppock & Pardee, 1987) and increased splicing efficiency of nuclear TK1 precursor transcripts (Gudas et al., 1988). Gudas et al. (1988) have found accumulation of several high molecular weight TK1 mRNA precursors and of mature 1.4 kb transcript within the nucleus as cells entered S-phase. Several hours later, the accumulation of nuclear TK1 mRNA was followed by accumulation of cytoplasmic TK1 mRNA. The authors suggested presence of a nuclear factor synthesized or activated at the beginning of S-phase and responsible for processing of hnRNA (Gudas et al., 1988). Using reverse transcription coupled to the polymerase chain reaction, Lipson & Baserga (1989) could not detect TK1 hnRNA in quiescent human fibroblasts. The amount of TK1 hnRNA increased transiently upon serum stimulation attaining maximum early in the S-phase. At its maximum concentration the ratio of TK1 hnRNA to TK1 mRNA was 1:155 (Lipson & Baserga, 1989).

Another post-transcriptional mechanism is the growth dependent regulation of TK1 mRNA by specific TK1 introns. Rotheneder et al. (1991) have found two DNase hypersensitive sites in intron 2 of the mouse TK1 gene. Addition of these sequences, or all of intron 2, to CAT constructs led to a marked stimulation of CAT expression over that seen with the mouse promoter only (Rotheneder et al., 1991). In a further investigation, Sutterluety et al. (1998) showed the presence of a putative antisense promoter in intron 3 of the murine TK1 gene, and found that TK1 specific antisense transcription was enhanced in resting mouse fibroblasts. The ratio between sense and antisense transcription changed from higher sense transcription in cycling fibroblasts to predominant antisense transcription in serum deprived cells. Since antisense transcription is correlated with repression of TK1 mRNA in

resting cells, it is very likely to play a role in the regulation of TK1 expression (Sutterluety *et al.*, 1998).

Translational regulation

Several studies indicated that expression of TK1 enzyme activity was regulated independently of TK1 mRNA. Ito & Conrad (1990) examined mRNA, protein and enzyme expression with constructs where a human TK1 cDNA was transcribed from the human promoter, the HSV-1 promoter and the SV40 early promoter. While the promoter did affect the pattern of mRNA expression, it had only a minimal effect on the pattern of protein or enzyme expression. For example, when Rat3 (TK⁻) cells were transfected with SV40-hTK1 construct, the mRNA was fully induced after 2-4 h of serum stimulation, while protein and enzyme activity levels remained low until about 12 h following serum stimulation, despite the presence of high level of mRNA. Thus, Ito & Conrad (1990) concluded, that induction of TK1 protein was uncoupled from induction of TK1 mRNA, and suggested presence of translational or post-translational regulation. Similarly, when Mikulits & Müllner (1994) transfected mouse fibroblasts with a vector expressing human TK1 cDNA from a constitutive promoter, they found constant mRNA level, but 8-fold rise in enzyme activity at G1/S transition, followed by decline in G2. The results indicated that in the absence of transcriptional control, the TK1 activity was repressed in all periods of the cell cycle except the S phase (Mikulits & Müllner, 1994). In further investigation, Mikulits *et al.* (1996) studied the consequences of TK1 mRNA abundance on the cell cycle dependent regulation of TK1 activity in nontransformed mouse cells. Low constitutive TK1 mRNA expression still resulted in maximal TK1 activity in S-phase. Increasing concentrations of constitutive mRNA compromised cell cycle regulation of TK1 activity, and the highest constant mRNA levels (observed after hormone induction of MMTV[4] promoter) resulted in constitutively high TK1 activity. These data indicated presence of factor(s) required for repression of TK1 translation during G1 which can be titrated by high

---

[4] Mouse mammary tumor virus

mRNA concentration (Mikulits *et al.*, 1996). 5'UTR of TK1 mRNA, when present at high concentrations, could be the molecule titrating the repression factor, as presence of 5'UTR allowed *in vitro* translation of TK1 mRNA in a cap-independent manner (Chou & Chang, 2001). 5'UTR in TK1 mRNA was proposed to confer a secondary structure to regulate ribosome binding during translation without being controlled by the cap-binding step (Chou & Chang, 2001).

Increased efficiency of translation of TK1 mRNA has been demonstrated by Sherley & Kelly (1988b) to account for the increase of TK1 activity in the S-phase of cycling HeLa cells. Here, the level of TK1 mRNA increased less than 3-fold, while the amount of TK1 polypeptide increased about 15-fold between G1 and G2. Changes in the level of TK1 protein were correlated with the changes in TK1 enzyme activity. Pulse labelling experiments where the rate of incorporation of ^{35}S-methionine into TK1 polypeptide were measured, revealed that the rate of synthesis of TK1 protein was 10-fold greater in S-phase than in G1 phase. In addition, the stability of TK1 protein decreased upon cell division resulting in low enzyme level in G1 cells (Sherley & Kelly, 1988b).

Also, Gross & Merrill (1988; 1989) reported that changes in TK1 mRNA levels were much smaller than changes in TK1 protein and activity levels in mouse myoblasts withdrawing from the cell cycle. These cells maintained nearly proliferative mRNA levels, but have greatly reduced levels of TK1 activity due to a 10-fold reduction in the rate of TK1 synthesis. TK1 protein stability was unchanged in this system, and the decrease in TK1 protein synthetic rate, in the continuous presence of mRNA, indicated decrease in the translational efficiency of TK1 mRNA. Because distribution of TK1 mRNA on polysomes did not change with differentiation, the authors suggested that translational control occurred at a post-initiation level (Gross & Merrill, 1989).

Post-translational regulation

Regulation by mitosis specific degradation

As already mentioned, Sherley & Kelly (1988b) demonstrated that the TK1 protein in cycling HeLa cells was degraded during mitosis and early G1. At different times between G1 and G2 the TK1 was as stable as bulk cellular protein which had a half-life of about 40

hours. The half-life of thymidine kinase decreased to about 4 hours at the time of cell division, and decreased further to less than 1 hour in the newly divided G1 cell (Sherley& Kelly, 1988b). In agreement with the above, a significant increase in TK1 protein levels at the G1/S transition in serum stimulated Rat3 cells (stably transfected with hTK1 cDNA linked to SV40/hTK1 promoter) was concluded to be primarily the result of a stabilization of TK1 protein at G1/S (Carozza & Conrad, 1994).

The timing of TK1 degradation and the half-life of the enzyme were defined more precisely when cells were arrested between metaphase and anaphase with the reversible microtubule inhibitor nocodazole (Kauffman & Kelly, 1991). The results indicated that degradation of wild-type TK1 began between metaphase and cytokinesis and that the half-life of TK1 during nuclear/cell division could be less than 20 minutes. Deletion of 40 C-terminal amino acids of human TK1 did not alter its enzymatic activity, but abolished the specific degradation at the border of mitosis and G1, and resulted in constitutive expression of the protein during the cell cycle (Kauffman & Kelly, 1991). When TK1 mRNA was expressed from a growth-independent heterologous promoter, the same C-terminal deletion allowed expression of thymidine kinase protein and activity in quiescent G0 cells (Kauffman et al., 1991).

The evidence for changes in the stability of TK1 during cell cycle, and for the importance of the C-terminal, was also obtained with mouse enzyme. Progressive deletion of up to 20 C-terminal amino acids led to stabilization of TK1 and to a stepwise loss of its growth-dependent regulation (Sutterluety et al., 1996; Mikulits et al., 1997). Mouse TK1 with a mutated, full-length version of its C-terminal, but with 50 aberrant amino acids was inherently unstable (Mikulits et al., 1997). Removal of the C-terminal 30 residues resulted in dramatic increase in protein stability which was the same in resting and growth induced cells (Sutterluety et al., 1996). The cellular half-life estimated in these experiments with TK1Δ30[5] was more than 18 hours, while the half-life of wild type mouse TK1 in quiescent cells was about 40 minutes (Sutterluety et al., 1996). So, in both the quiescent (Sutterluety et al., 1996) and mitotic cells (Kauffman & Kelly, 1991), the low wild type TK1 enzyme levels, are at least partly due to decrease in TK1 protein stability. Later, Sutterluety & Seiser (1997) have shown that the TK1 protein half-life rose more than six-fold after addition of thymidine to resting cells, while TK1 mRNA levels and TK1 translation rate

---

[5] TK1Δ30 is a polypeptide where the last 30 amino acids were removed

were unaffected. Binding of thymidine could potentially change the TK1 protein conformation and therefore influence TK1 degradation (Sutterluety & Seiser, 1997).

More recently, Posch *et al.* (2000) reported that the ability of the murine TK1 to bind substrates was essential for both growth-dependent regulation and stabilization in proliferating fibroblasts. Mutants lacking the binding sites for either ATP or thymidine were not only enzymatically inactive, but showed no induction of protein levels in response to serum stimulation due to rapid degradation in exponentially growing cells. By cross-linking experiments it was shown that TK1 proteins with mutated substrate binding sites existed only as monomers in contrast to active TK1 enzyme (TK1 wild type and TK1Δ30) that existed as dimers and tetramers. It was concluded that substrate binding to the TK1 protein was a prerequisite for stabilization, and the availability of substrates might offer additional means for fine-tuning enzyme degradation (Posch *et al.*, 2000). These studies of *in vivo* properties of murine TK1 (Posch *et al.*, 2000) are in excellent agreement with *in vitro* studies of Munch-Petersen *et al.* (1993) on pure human lymphocyte cytosolic TK1: ATP was found to be not only a stabilizer of thymidine kinase activity, but a positive effector inducing enzyme concentration-dependent shift from a less active dimer to a tetramer with higher substrate affinity.

Recently, the mitotic degradation of human TK1 was shown to be dependent on a ubiquitin-proteasome pathway - the anaphase-promoting complex and its activator Cdh1, controlled destruction of hTK1 through a KEN box motif located in the C-terminal of the protein (Ke & Chang, 2004).

Regulation by phosphorylation

The effect of substrates on TK1 stability can be connected to results of Chang & Huang (1993) and Chang *et al.* (1994), that have found growth dependent phosphorylation of TK1 in promyeloleukemic HL60 human cells and in several other human cell lines. HL60 cells showed constitutive expression of the level of TK1 polypeptide, but TK1 activity increased during growth stimulation (Chang & Huang, 1993). This increase of activity was accompanied by increment in the extent of phosphorylation of the enzyme. Phosphoamino acid analysis indicated that serines were the only phosphorylated residues and that the serines were situated at two major and several minor phosphorylation sites (Chang & Huang, 1993). The phosphorylated sites in TK1 polypeptide in proliferating and

mitotically blocked HeLa cells were found to be distinctly different (Chang *et al.*, 1994), and the hyperphosphorylation of TK1 in mitotically blocked cells resulted in decreased affinity for thymidine: K_m was 15.5 μM as compared to 1.49 μM in proliferating cells (Chang *et al.*, 1994). The reduced thymidine affinity was suggested to be at least partly responsible for the decreased stability of TK1 (Chang *et al.*, 1994).

In a later investigation, Chang *et al.* (1998) have demonstrated that Ser[13] of the human TK1 was the primary site of mitotic phosphorylation, and that the wild type TK1, but not the mutant TK1[Ala13], could serve as a good substrate for CDK1or CDK2 kinases *in vitro*. In this study, no difference in the catalytic properties of Ser[13] phosphorylated and unphosphorylated TK1 was found, and the importance of another residue, Ser[231], for the mitotic phosphorylation could not be excluded (Chang *et al.*, 1998).

Ser[13] was also shown to play a crucial role in the mitotic degradation of human TK1 expressed in yeast (Ke *et al.*, 2003). Mutation of Ser[13] to Ala resulted in elevated accumulation of hTK1 and in reduced rate of hTK1 degradation during mitosis in both budding and fission yeasts (Ke *et al.*, 2003). Functional intactness of the yeast proteasomes, more specifically of the so-called SCF-complex, was found to have a major contribution to the degradation of hTK1 in yeast cells (Ke *et al.*, 2003). However, the SCF complex and Ser[13] were not essential for mitotic degradation of hTK1 expressed in mammalian cells. Instead, another ubiquitin ligase, anaphase-promoting complex and its activator Cdh1, were shown to control destruction of hTK1 through a KEN box motif located in the C-terminal of the protein (Ke & Chang, 2004).

Regulation at the enzymatic level

The apparent K_m values of TK1 for its substrate thymidine - 15.5 μM and 1.49 μM in mitotically blocked and in proliferating cells, respectively (Chang *et al.*, 1994) are in very good agreement with the K_m values found *in vitro* for the pure substrate-free human lymphocyte TK1: a less active dimeric form with K_m about 15 μM associated reversibly upon incubation with ATP into a more active tetrameric form with K_m about 0.7 μM (Munch-Petersen *et al.*, 1993). This transition was dependent on enzyme concentration, ATP and thymidine (Munch-Petersen *et al.*,1993; Munch-Petersen *et al.*,1995a). Because the concentration of TK1 protein fluctuates during the cell cycle, the low affinity dimer form was suggested to dominate in G1, the high affinity tetrameric form in S/G2 phases,

and the transition was considered as an additional fine tuning of TK1 activity during the cell cycle (Munch-Petersen et al.,1995a).

Perturbation in ATP-induced transition pattern from low thymidine affinity dimer to high thymidine affinity tetramer has recently been reported for a recombinant TK1 enzyme in which Ser-13 was substituted with aspartate (Li et al., 2004). The S13D substitution which mimics the mitotic phosphorylation of Ser-13 caused an equilibrium shift from a tetramer to a dimer. Because the dimeric form of TK1 is less active than the tetrameric, it was proposed that the physiological role of Ser-13 phosphorylation is to counteract ATP-dependent activation of TK1, thus attenuating its enzymatic function at G2/M phase (Li et al., 2004).

Regulation of TK1 expression in malignant cells

The overall regulation of TK1 is different in tumour derived and virus transformed cells when compared to normal cells (Hengstschläger et al., 1994a). As already mentioned, there was little regulation at the mRNA level in human HeLa cells (Sherley & Kelly, 1988b). Hengstschläger et al. (1994a, 1994b, 1994c) measured enzyme activity and mRNA levels during the cell cycle of normal fibroblasts, normal or malignant lymphocytes, other malignant cell lines like retinoblastoma and carcinoma, and of cells transformed with different viruses (polyoma, papilloma, adenovirus, SV40 or Epstein Barr virus). In normal cells both mRNA level and TK1 enzyme activity was transiently induced in early S-phase. In malignant and virally transformed cells TK1 mRNA level and TK1 activity were consistently higher in all phases of the cell cycle, and only minor fluctuations in the mRNA level during the cell cycle could be observed. In contrast to the normal cells, there was no decrease in TK1 activity during G2: both TK1 polypeptide level and enzyme activity stayed at the high S-phase levels until mitosis (Hengstschläger et al., 1994a, 1994b, 1994c).

Malignant chronic lymphatic leukaemic (CLL) cells showed a different mechanism of deregulation of TK1 mRNA synthesis and of TK1 enzyme activity (Kristensen et al., 1994). In normal PHA stimulated human lymphocytes, an about 100-fold increase in TK1 mRNA level at the entry to S phase was followed by an increase in TK1 activity.

Surprisingly, in CLL cells, the TK1 mRNA was expressed 30-300 fold higher than in quiescent lymphocytes, but the TK1 activity stayed low and at the same range as in quiescent lymphocytes (Kristensen *et al.*, 1994). A defect in processing of the enzyme in CLL cells was suggested, but the reason for this deregulation is unknown.

TK1 as a modulator of cell cycle progression

The inhibitor of CDK2/cyclinE and CDK2/cyclinA complexes, p21, is known to be a negative regulator of cell growth. As already mentioned, it could influence promoter activity of TK1 (Chang & Huang, 2001) and was also shown to suppress mitotic phosphorylation of TK1 protein (Chang *et al.*, 1998). Huang & Chang (2001) found that antiserum against human TK1 co-immunoprecipitated p21 indicating that hTK1 can form a complex with p21 polypeptide. TK1 activity was not changed by this association, but the p21-mediated inhibition of cell proliferation was markedly reversed. Using immunoprecipitation technique it was shown that elevated expression of TK1 in the cells, disrupted association of p21 with CDK2. Huang & Chang (2001) hypothesize, that elevated expression of TK1 in some tumour cells not only provides dTTP for DNA replication but also attenuates the inhibitory growth function of p21.

Chapter 4

Structure of the human thymidine kinase 1

Knowledge of protein structure is an invaluable aid for rational design of new pharmaceuticals: It should be possible to make drugs that bind tightly to the enzyme and interfere specifically with its function. Two methods, X-ray crystallography and nuclear magnetic resonance (NMR), are being widely used for investigations of protein structure. In principle, each of the methods can position all the atoms in a molecule and their relation to each other. An advantage of NMR spectroscopy is that it is carried out on molecules in solution and can show conformational variability and flexibility of protein molecules. X-ray crystallography is limited to molecules that can be crystallized and describes proteins frozen into one conformation in their crystalline form.

The crystal structure of *Herpes simplex* virus type 1 thymidine kinase (HSV1-TK) has been studied since 1995, and solved in complex with different substrates (Brown *et al.*, 1995; Wild *et al.*, 1995; 1997; Champness *et al.*, 1998; Bennett *et al.*, 1999; Prota *et al.*, 2000; Vogt *et al.*, 2000). Recently, the crystal structure of thymidine kinase from *Varicella zoster* virus (human herpes virus type 3) complexed with bromovinyl-deoxyuridine 5'-monophosphate (BVDU) and ADP has been determined (Bird *et al.*, 2003). Also, the structures of crystalline deoxynucleoside kinase from the fruit fly *Drosophila melanogaster – Dm*-dNK, and of crystalline human dGK have been elucidated (Johansson *et al.*, 2001).

The human TK1 was cloned in 1983 - 1984 (Bradshaw, 1983; Bradshaw & Deininger, 1984), allowing the necessary amounts of TK1 protein for crystallization to be produced. Unfortunately, attempts to grow TK1 crystals suitable for X-ray crystallography have so far been unsuccessful.

Large amounts of pure protein are also required for structure determination by nuclear magnetic resonance. NMR enables determination of 3D-structure of proteins in solution if their size does not exceed 30 kDa. Because the subunit size of TK1 is 25 kDa, and TK1 exists in solution as a mixture of dimers and tetramers, the available NMR methods do not offer the required resolution and sensitivity.

Although the three-dimensional structure of TK1 still awaits to be solved, I will propose a possible 3D-structure of TK1 based on its amino acid sequence, computer calculations for probabilities of formation of β-strands and α-helixes, and analogy to the structures of the already known deoxynucleoside and deoxynucleotide kinases - adenylate kinase and other NMP kinases, *Herpes simplex* virus thymidine/thymidylate kinase, the human deoxyguanosine kinase and *Drosophila melanogaster* deoxynucleoside kinase.

The primary structure of TK1

The human thymidine kinase has 234 amino acid residues and forms a subunit of 25 kDa. A search in the protein databases of NCBI (National Center for Biotechnology Information) with BLAST (Basic Local Alignment Search Tool (Altschul *et al.*, 1990; Altschul *et al.*, 1997) results in a large number of sequences from vertebrata, nonvertebrata, plants, bacteria and viruses of the pox family. The vertebrate and vaccinia virus kinases show the highest homology to the human thymidine kinase 1 (Fig. 4.1). The main difference between the vertebrate and pox virus enzymes is that the former possess additional N-terminal and C-terminal residues (15 and 42 respectively for the human TK1 (Boyle *et al.*, 1987; se also Fig. 4.1).

Based on amino acid sequences, Eriksson *et al.* (2002) divided the known deoxyribonucleoside kinases into three groups. The human TK1 with vertebrate and vaccinia kinases form a group of their own. The other mammalian kinases, dCK, dGK, TK2 and the *Drosophila* multisubstrate deoxynucleoside kinase, dNK, form a second group (the so called TK2 group), and *Herpes virus* kinases form a third group. Sequence comparison of TK1 to other cellular deoxynucleoside kinases, and to *Herpes simplex* virus thymidine kinase show low overall homology confined to the phosphate-binding motif, the so called P-loop (Eriksson *et al.*, 2002). The P-loop (Fig. 4.2) is commonly found in ATP- and GTP-binding proteins, such as mononucleotide kinases, elongation factors, myosin and Ras proteins (Saraste *et al.*, 1990). This P-loop motif is probably the most ancient motif in nucleotide-binding proteins and might be the archaic mode of interaction between nucleotides and proteins (Dreusicke & Schulz, 1986). The consensus sequence for the P-

```
Human     MSCINLPTVLPGSPSKTRGQIQVILGPMFSGKSTELMRRVRRFQIAQYKCLVIKYAKDTR 60
Mouse     MSYINLPTVLPSSPSKTRGQIQVILGPMFSGKSTELMRRVRRFQIAQYKCLVIKYAKDTR 60
Hamster   MNYINLPTVLPGSPSKTRGQIQVILGPMFSGKSTELMRRVRRFQIAQNKCLVIKYAKDTR 60
Chicken   MNCLTVPGVHPGSPGRPRGQIQVIFGPMFSGKSTELMRRVRRFQLAQYRCLLVKYAKDTR 60
Vaccinia  MN--------------GGHIQLIIGPMFSGKSTELIRRVRRYQIAQYKCVTIKYSNDNR 45
Mycoplasma MYNRF----------NKGMIEVITGPMFSGKTEELLRRFRLLNYAKAKTLLIKPAFDTR 49
Bacillus  MYLIN----------QNGWIEVICGSMFSGKSEELIRRVRRTQFAKQHAIVFKPCIBNR 49
                *            * *::* *.*****; **;**.*   : *; ; : .*  .*


Human     YSSS-FCTHDRNTMEALPACLLRDVAQEAL--GVAVIGIDEGQFFP-DIVEFCEAMANAG 116
Mouse     YSNS-FSTHDRNTMDALPACMLRDVTQEAL--GVAVIGIDEGQFFP-DIVDFCEMMANEG 116
Hamster   YSSS-FSTHDRNTMDALPACLLRDVAQEAL--GAAVIGIDEGQFFP-DIVEFCEVMANAG 116
Chicken   YCTTGVSTHDRNTMEARPACALQDVYQEAL--GSAVIGIDEGQFFP-DIVEFCEKMANTG 117
Vaccinia  YGTG-LWTHDKNNFEALEATKLCDVLESIT--DFSVIGIDEGQFFP-DIVEFCERMANEG 101
Mycoplasma FSKEEIISRAGVKTKTHSVKNTEQIRKILEKEKFDALVIDEIHFFDFDIVYLIEELANSG 109
Bacillus  YSEEDVVSHNGLKVKAVPVSASKDIFKHIT-EEMDVIAIDEVQFFDGDIVEVVQVLANRG 108
           :    . :!    . .:  .    :: :      .: *** ;** *** , : ;** *


Human     KTVIVAALDGTFQRKPFGAILNLVPLAESVVKLTAVCMECFREAAYTKRLGTEK------ 170
Mouse     KTVIVAALDGTFQRKAFGSILNLVPLAESVVKLTAVCMECFREAAYTKRLGLEK------ 170
Hamster   KTVIVAALDGTFQRKAFGSILNLVPLAESVVKLTAVCMECFREAAYTKRLGLEK------ 170
Chicken   KTVIVAALDGTFQRKAFGSILNLVPLAESVVKLNAVCMECYREASYTKRLGAER------ 171
Vaccinia  KIVIVAALDGTFQRKPFNNILNLIPLSEMVVKLTAVCMKCFKEASFSKRLGEET------ 155
Mycoplasma YHIIVSGLDQNFKREPFEVVSYLLSIAEKVTKLQAICVKCQRAATTTFRKVESK------ 163
Bacillus  YRVIVAGLDQDFRGLPFGQVPQLMAIAEHVTKLQAVCSACGSPASRTQRLIDGEPAAFDD 168
           :**:.** *:  .*  :  *:.:* *.** *:* *   *; : *


Human     EVEVIGGADKYHSVCRLCYFKKASGQPAGPDNKENCPVPGKPGEAVAARKLFAPQQILQC 230
Mouse     EVEVIGGADKYHSVCRLCYFKKSSAQTAGSDNK-NCLVLGQPGEALVVRKLFASQQVLQY 229
Hamster   EVEVIGGADKYHSVCRVCYFKKSSVQPAGPDNKENCPVLGQPGEASAVRKLFAPQQVLQH 230
Chicken   EVEVIGGADKYHSVCRACYFQKRP-QQLGSENKENVPMGVKQLDMPASRKIFAS------ 224
Vaccinia  EIEIIGGNDMYQSVCRKCYIDS-------------------------------------- 177
Mycoplasma EIKLLGDVDEYEARCRKCHIQGSKDKN-------------------------------- 190
Bacillus  PIILVGASESYEPRCRHCHAVPTKQR-------------------------------- 194
           : ::*  : *.. ** *;


Human     SPAN 234
Mouse     NSAN 233
Hamster   NSTN 234
Chicken   ----
Vaccinia  ----
Mycoplasma ----
Bacillus  ----
```

Figure 4.1. Alignment of human TK1 amino acid sequence with related enzymes (Clustal W 1.82). Symbols below the sequences are: "*" – identical residues in all sequences, ":" – conserved substitutions, "." – semiconserved substitutions. The sequences have following GenBank Identifier numbers: gi/23503074, human; gi/6678357, mouse; gi/125428, Chinese hamster; gi/125427, chicken; gi/9791018, vaccinia virus; gi/15828616, Mycoplasma; gi/21397804, Bacillus.

loop motif, originally called motif A by Walker and co-workers (Walker *et al.*, 1982) and later "Walker" A motif, is **GXXXXGKS/T**, with invariable glycines and a key lysine for providing the positive charge, and followed by a serine or threonine (Fig. 4.2). Another conserved sequence, Walker B motif, was originally identified as $R/KX_{1-4}GX_{2-4}\Phi X\Phi_2\Phi D/E$ (Walker *et al.*, 1982; Traut, 1994b) and later more simple as $\Phi\Phi\Phi\Phi D$ where Φ can be any hydrophobic amino acid and D is an invariable aspartate (Myles *et al.*, 1991). The Walker B sequence is positioned at an average distance of 60 amino acids from the P-loop (Traut, 1994b) and corresponds to VIGID in the vertebrate TK1 and vaccinia TK1 sequences (Fig. 4.1 and Fig. 4.2). The conserved aspartate has been found to coordinate a Mg^{2+} ion required for phospho-transfer reaction (se later).

A: P-loop, the phosphate-binding site (Walker A motif)

```
HSV1       (48) TLLRVYIDGPHGMGKTTTT
Human TK1  (18) RGQIQVILGPMFSGKSTEL
Human TK2  (49) KKSVICVEGNIASGKTTCL
Vaccinia   (3)  GGHIQLIIGPMFSGKSTEL
ADK        (7)  KSKIIFVVGGPGSGKGTQC
EF-Tu      (11) PHVNVGTIGHVDHGKTTLT
```

B: Mg^{2+} binding site (Walker B motif)

```
(156) ALTLIFDRHPIAAL
(91)  VAVIGIDEGQFFPD
(127) SSVRLMERSIHSAR
(76)  FSVIGIDEGQFFPD
(87)  SKGFLIDGYPREVK
(74)  RHYAHVDCPGHADY
```

C: Thymidine binding site

```
HSV1       (151) HAPPPALTLIFDRHPIAAL
Human TK1  (118) TVIVAALDGTFQRKPFGAI
Human TK2  (122) TRPQVSSVRLMERSIHSAR
Vaccinia   (103) IVIVAALDGTFQRKPFNNI
ADK             ------------------
EF-Tu           ------------------
```

D: LID motif

```
(209) PEDRHIDRLAKRQRPGER
(158) REAAYTKRLGTEKEVEVI
(185) NPETCYQRLKKRCREEEK
(143) KEASFSKRLGEETEIEII
(121) GPETMTKRLLKRGETSGR
      ---------------
```

Figure 4.2. Alignment of different thymidine kinases with adenylate kinase (ADK, from vertebrate muscle) and elongation factor EF-Tu (from *E. coli*). Conserved residues/regions are in bold; "----" means absence of the homologous motif. Sequences for human TK2 were aligned with HSV1 TK according to Eriksson et al. (2002). The sequence and numbering of HSV1-TK is after Wild *et al.* (1997), vaccinia virus TK after Black & Hruby (1990), and adenylate kinase after Schulz *et al.* (1986). Human TK1 has accession nr P04183, TK2 nr O00142 and EF-Tu nr NP_417798 respectively, in the SwissProt databank. This figure is modified after Folkers *et al.* (1991).

Alignment of different thymidine kinases revealed still another conserved site: The FQRK sequence in the human TK1 corresponding to FDRH in herpes viral thymidine kinases (Balasubramanian *et al.*, 1990; Gentry, 1992) and to F/MERS in the TK2 group of deoxynucleoside kinases (Eriksson *et al.*, 2002). On the basis of sequence analysis and enzyme kinetic results from mutants in which the conserved residues were replaced, this

site was suggested to be involved in binding of thymidine (Gentry, 1992). More recent experiments with murine TK1 confirmed importance of the P-loop and of the putative thymidine binding site: exchange of lysine 32 with isoleucine, or deletion of residues 28 to 33 (in the putative ATP binding site), and mutations of arginine 130 and phenylalanine 133 to alanine, or deletion of residues 130 to 138 (in the putative thymidine binding site) resulted in a complete loss of TK1 enzymatic activity (Posch *et al.*, 2000).

An additional region, although not very conserved in eukaryotic and pox thymidine kinases, is an arginine-rich consensus sequence RXXXRXR present in the so-called LID motif of the enzyme (the LID closes upon substrate binding like a lid). Alignment of the conserved regions found in thymidine kinases, adenylate kinase and GTP-binding elongation factor EF-Tu is shown in Fig. 4.2.

Secondary structure of TK1

It is a general belief that the sequences of amino acids determine if the peptide will form an α-helix, β-strand or a random coil, which in turn will fold up into a defined tertiary structure (Cook, 1967; Anfinsen, 1973). Different amino acids have been found with different frequencies in α-helices, β-strands, loops, or in turns that break the chain and ensure that it folds back to form the core. These differences can be essentially explained by differences in size, conformation and charge of each amino acids, but at the same time the neighbouring residues and residues involved in tertiary interactions, are found to be important in determining the secondary structure. Many sequences have been found to adopt alternative conformations: α-helix in one protein context and β-strand in another[1]. In 1983 Kabsch & Sander evaluated the accuracy of algorithms generally used for the secondary structure prediction (Lim, 1974a; Lim, 1974b; Garnier *et al.*, 1978; Chou & Fasman, 1978) to be below 56%. Increase of the prediction accuracy to 66% was obtained when alignment of homologous sequences was used simultaneously with an algorithm quantifying the degree of conservative substitution at each alignment position, including insertions and deletions (Zvelebil *et* al., 1987). Later algorithms, as for example the

---

[1] If a protein assumes inappropriate conformation for the context, pathological diseases can result as in the case with prion diseases (mad cow disease, Creutzfeldt-Jacob disease).

profile-based neural network program PHD (Profile network from HeiDelberg), used multiple sequence alignments and reached prediction accuracy greater than 70% (Rost & Sander, 1993; Rost & Sander, 1994). Since the 1990ies the prediction area is in a constant state of development. Improvements in computational tools, X-ray crystallography and NMR spectroscopy have resulted over the past few years in flood of information on protein sequence and structure homology. The present secondary structure prediction algorithms obtain their parameters from an analysis of proteins of known three-dimensional structure and then apply them to the sequence of unknown structure. Novel methods using neural networks can predict protein secondary structure at about 80% accuracy (Petersen *et al.*, 2000). I have used the secondary structure prediction provided by JPred server (Cuff *et al.*, 1998; Cuff & Barton, 1999; Cuff & Barton, 2000). JPred (or JNet) combines several prediction methods to achieve a consensus prediction that has an improved accuracy when compared to any single method. JNet is being continuously developed and the averaged Q_3 score, where Q_3 is the percentage of residues predicted correctly for the three conformational states, strand, helix and loop, increased from 72,9% (Cuff *et al.*, 1998) through 78% (Cuff & Barton, 1999) up to 84% (Cuff & Barton, 2000). As can be seen in Fig. 4.3 different methods give different results. The JPred consensus prediction differs as well from the earlier prediction of Folkers *et al.* (1991), where model of TK1 was built based on secondary structure prediction method of Garnier *et al.* (1978) and with 3D-structure of adenylate kinase as template. Still, the pattern is clear: By all methods TK1 is predicted to be an α/β protein built of alternating α-helices and β-strands, a common feature in nucleotide binding proteins.

```
              10        20        30        40        50        60        70        80        90       100       110
Human      MSCINLPTVLPGSPSKTRG.......GPMFSG.STELMRRVRRFDIAQYK.....YAKDTRYSSSFCTHDRNTMEALPACLLR....QEALGV......IDEGQFFP..VEF.EAM.NA
Chicken    MNCLTVPGVHPGSPGRPRGQIQVIFGPMFSGKSTELMRRVRRFQLAQYRCLLVKYAKDTRYTTGVSTHDRNTMEARPACALQDVYQEALGSAVIGIDEGQFFPDIVEFCEKMANT
Poxvirus   ---------------GYIKLILGPMFSGKTTELVRIVKRYKIANYKCCVIKYYNDNRCDESIVTHDGVYIDSISTLKLNDIIYEMDNVDVIGIDEGQFFNDIVEFSENMANK
C.elegans  ----------LPRCPNRV-GSITVILGPMFSGKTTELLRLHDRQIIAKRTCVLVKYAGDTRYDADLVTHSKMTGQGVKAHRLSEVQSQIFNVQVVSIDEGQFFEDLAETCEELAQR
Mycoplasma -------------SPRGWIEVICGPMFSGKTEELLRKIKRWKLAKIPVIIFKPKIDTRQQHLVKSRNGHSDEAIEINSPLEIYDTKDRFDVVAIDEAQFFSEIVEVVKSLNDL
Bacillus   -------------KQSGWLELICGSMFSGKSEELIRRVKRATYAKQEVRVFKPVIDNRYSEAVVSHNGTSMTSYAISSAADIWDHIESTDVVAVDEVQFFDQIVEVLSSLADK


Jalign     -------------------EEE--------HHHHHHHHHHHHH-----EEEE----------EE-----------------EE-------EEEE-----HHHHHHHHHH----
Jfreq      -------------------EEEEEE--------HHHHHHHHHHHHH---EEEEEE-------EEEEE------------H-HHE------EEEEE--H----HHHHHHHHHH--
Jhmm       ------EEEEE----------EEEEEE----EEEEEE-----EEE-----------EEEEEE-------E---EE----------------------EEEEEEEEEEEEE--
Jnet       ---------E----------EEEEEE-------HHHHHHHHHHHHHH---EEEEEE--------EEEEE------------HHHHHHH----EEEEE-------HHHHHHHHHH--
Jpssm      -------------------EEEEEEE-----HHHHHHHHHHHHHH---EEEEE-----------EEE----------HHHHHHH----EEEEE--------HHHHHHHHH--
Jpred      -------------------EEEEEE-------HHHHHHHHHHHHHH---EEEEE---------EEE--------------HHH------EEEEE-------HHHHHHHHHH--
                         β1              α1           β2                              α2        β3        α3


              120       130       140       150       160       170       180       190       200       210       220      230
Human      GK.....LDGTFQRKPFGAILNLV.....KLTAVCMECFREAAYTKRLGTEKE..............FKKASGQPAGPDNKENCPVPGKPGEAV...LFAPQQILQCSPAN
Chicken    GKTVIVAALDGTFQRKAFGSILNLVPLAESVVKLNAVCMECYREASYTKRLGAEREVEVIGGADKYHSVCRACYFQKRP-QQLGSENKENVPMGVKQLDMPASRKIFAS---------
Poxvirus   GKIVIVSALDATYQRKTFGNILNLIPLSEKVIKLNAICKICFNDAAFTKRLCDDTKIELIGGEDKYSSVCRKCYF---------------------------------------
C.elegans  GKVVCVAALNGTFERKPFPQISLLLPYANEIKQVTAVCVECGSQANFSFRSTLDKKVEVIGGSDTYTALCRECYVQKSEEKDAEEQMKTGC----------------------
Mycoplasma GINVIVSGLDTDFRAEPFGSIPQLLAIADKICKLDAVCNVCGQLAQRTQRIVSKNETVLIGDIEAYEPRCKLH----------------------------------------
Bacillus   GYRVIAAGLDMDFRGEPFGVVPNIMAIAESVTKLQAVCSVCGSPASRTQRLIDDDPVILVGAAESYEARCR--HHHEVPGK--------------------------------


Jalign     --EEEEEE-----------------H-HHHHHHHHHH-------H-------EEEE----E-HHH-------------------------------------------
Jfreq      --EEEEEE-------------HHHHHHHHHHHHHHHHEEE-------HHHHH----EEEEE----EEEEEE-EEE------------------------HHHHHHHH-HHHHHH-----
Jhmm       EEEEEE--------------------------------------------EEE--EEEE--------------EEEEEE------------------------
Jnet       --EEEEEE-------------HHHHHHHHHHH---EEEEE--------EE------EEEEE---EEEEEE--EEE-----------------------HHHHHHH--HHHH------
Jpssm      --EEEEE--------------HH-EEE-EE-E--EEEEEE----EEEEEE-----EEEE----EEEE-----EEE-----------------------HHEE-----------
Jpred      --EEEEEE----------------HHH-HHH---EEEEE--------EE------EEEE----EEEEE----E------------------------HHHH-----------
               β4              α4                        β5                                      α5
```

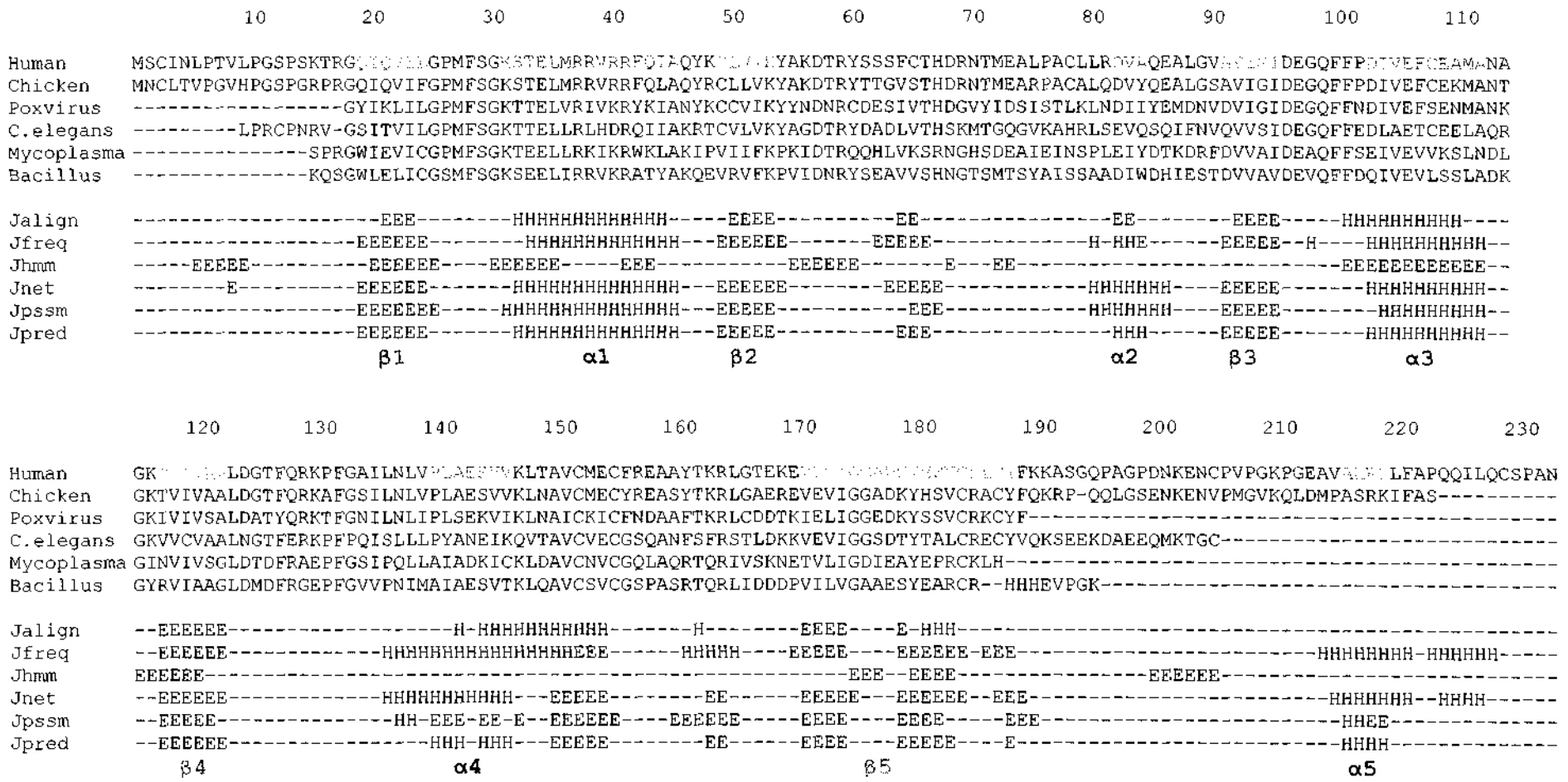Figure 4.3. Prediction of TK1 secondary structure by Jpred (se text). Jpred is the consensus prediction over all methods. The numbering refers to human TK1. The sequences have following SwissProt databank accession numbers: P04047, chicken; Q90024, poxvirus; Q9XXF3, C. elegans; P75070, Mycoplasma; Q03221, Bacillus. HHH, α-helix; EEE, β-strand; ---, residues to which no structures could be assigned.

Supersecondary structure

Most nucleotide binding proteins have a core consisting of several parallel β strands alternating with α-helices. The β-α-β unit formed by two parallel β-strands connected by an α-helix is called Rossman-fold, after Michael Rossmann who described this domain first (Rao & Rossmann, 1973; Rossmann et al., 1974), and who defined topologically equivalent residues in polypeptides that have little sequence similarity, but adopt similar folds[2] (Matthews & Rossmann, 1985). Fig. 4.4 shows a topology diagram (description of how different structures are connected) for TK1 based on JPred secondary structure prediction.
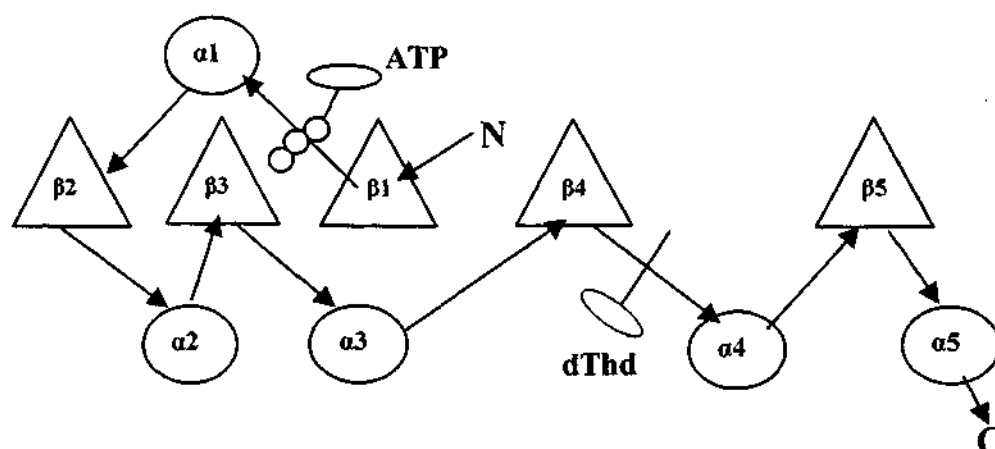
Figure 4.4. Proposed topology diagram of TK1 based on JPred prediction (se also Fig. 4.3). Binding sites for thymidine (dThd) and ATP are also shown.

For comparison, the secondary structure and topology of the Drosophila pyrimidine kinase Dm-dNK is:

β1 →α1 →β2 →α2 →α3 →α4 →β3 →α5 →α6 →β4 →α7 →α8 →β5 →α9 (Johansson et al., 2001). dNK is somewhat longer, 250 amino acids as compared to 234 of TK1, and also the total length of the secondary structure elements is higher: 150 out of 250 (about 60%) amino acids are found in α-helices or in β-strands, whereas JPred predicts only 70 out of TK1's 234 amino acids (about 30%) to be in α-helices or in β-strands. Solution of 3D-structure of TK1 may change this number.

---

[2] Folds, also called motifs, are defined as arrangements of several elements of secondary structure and the connections between them

Tertiary structure

The tertiary structure is defined as the arrangement of all the atoms of the protein in space, comprising all possible interactions.

The Protein Data Bank[3], an archive for the structures of all biological macromolecules (Bernstein et al., 1977; Westbrook et al., 2002), contained in October 2004 about 25000 proteins of known three-dimensional structures, and the number of deposited experimentally determined protein structures continues to increase exponentially.

Comparison of 3D structures of sequenced proteins revealed that: 1) homologous proteins show high degree of structural resemblance, and 2) proteins with different sequences (less than 20% residue identity) can adopt the same fold; i.e. structure is conserved more extensively than sequence (Lesk & Chothia, 1980; Chothia & Lesk, 1986; Ponting & Russell, 2002). Also, theoretical calculations estimated the number of distinct structural folds to be limited to a few thousands (Zhang, 1997). This made homology modelling possible: a three-dimensional model of a protein sequence is constructed based on known structures of related proteins. To obtain a reasonable level of accuracy, the sequence identity between the modelled sequence and at least one known sequence must be higher than 30%. An analysis of the Protein Data Bank of experimentally determined structures of proteins reveals that protein pairs with more than 30% pairwise sequence identity (for alignment length > 80 residues) have homologous three-dimensional core structures, i.e. the essential fold of the two proteins is identical, but additional loop regions may vary (Rost & Sander, 1996). In 1999 the best prediction methods based on comparative modelling could built quite accurate models: <1.5 Å rmsd (root mean square deviation of C_α positions) if the new protein showed more than 30% identity to the known homolog, and <4.4 Å rmsd if the new protein showed 17% identity (Sternberg et al., 1999). Nevertheless, a request in February 2003 at Center for Biological Sequence Analysis (www.cbs.dtu.dk) for 3D structure of TK1 resulted in following answer: "No similar sequence found in PDB". However, a similar request in October 2004, retrieved a template on which a putative model could be build. The template, PDB entry name 1u98, had 33%

---

[3] at www.rcsb.org/pdb (Research Collaboratory for Structural Bioinformatics)

identity to TK1, and surprisingly, it was none of the nucleoside kinases with solved 3D structures (se later), but ATP and DNA binding protein RecA from *E. coli*. Another modelling request in October 2004, to a META PredictProtein server of V. Eyrich and B. Rost (www.predictprotein.org), could not be carried out, because "the degree of similarity with proteins of known 3D structure may be to low". Still, the META server assigned TK1 to a protein superfamily[4] of P-loop containing nucleotide triphosphate hydrolases – obviously, the P-loop motif is the common denominator.

Overall 3D-structure of TK1

Sequence identity between HSV1-TK and enzymes belonging to the TK2 group is very low (~10%), despite this, the core structures of HSV1-TK, dGK and *Dm*-dNK overlap (Johansson *et al.*, 2001). The recently solved crystal structure of human dCK (47% sequence identity with dGK) complexed with dC and ADP shows a fold very similar to *Dm*-dNK and dGK (Sabini *et al.*, 2003). It is therefore quite probable that also TK1 has a similar structural fold shared with other nucleoside and nucleoside monophosphate kinases. This fold consists of three parts: the CORE region that includes the central 5-pleated parallel β-sheet surrounded by α-helices, the LID region, and the phosphate-binding region, the P-loop (Fig. 4.5, se also Fig. 4.10 for the open conformation of the enzyme). Comparison of the topology model of TK1 (Fig. 4.4) with the topology and the solved three-dimensional structure of *Dm*-dNK (Fig. 4.5) reveals two main differences: 1) presence of only two helices, α2 and α3, in the right part of the structure, and not five, α2 to α6, as in dNK, and 2) the LID region of *Dm*-dNK is formed by α7, α8 and the turn between them, while it seems that α4 and the following turn is the LID region of TK1. The fold of HSV-1 TK, which has an even bigger subunit of 376 amino acids built of 5 β-strands and 12 α-helices (Wild *et al.*, 1997), also shows a clear overlap with dGK and *Dm*-dNK (Johansson *et al.*, 2001).

---

[4] Proteins with significant sequence similarity, and similar structure and function belong to the same protein family. A strong evolutionary relationship is evident within a protein family (for example the globin family). If two or more families with little sequence similarity have the same structural motif and functional similarities they are grouped as superfamily. An evolutionary relationship between the families in a superfamily is considered probable.
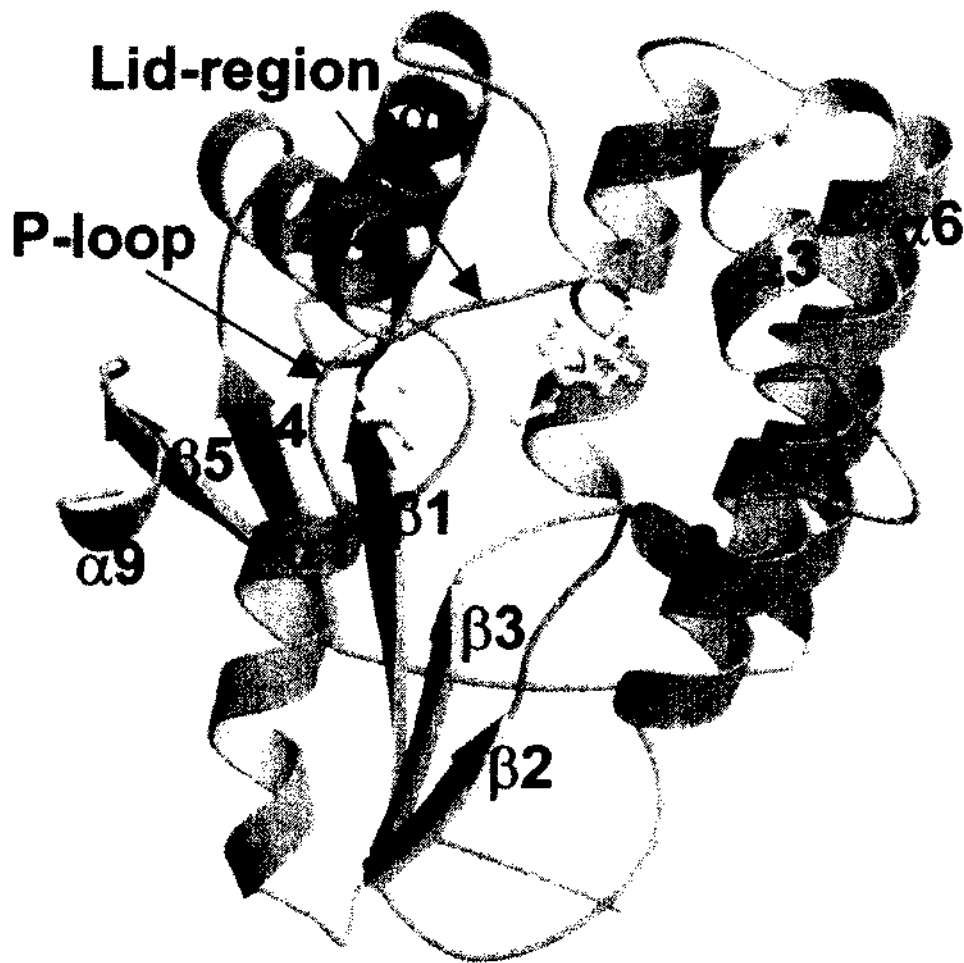
Figure 4.5. 3D structure of dNK subunit with bound deoxycytidine and sulfate ion (in yellow, in ball-and-stick representation. (After Johansson *et al.*, 2001).

The P-loop motif is present between β1 and α1 and forms the phosphate donor site together with the LID region. In analogy to *Dm*-dNK and dGK, it might be suggested that α4 of TK1 corresponds to α5 of *Dm*-dNK, and that the dimers of TK1 are formed by two helices from each subunit: α2 and α3. These two mainly hydrophobic helices contain many β-branching amino acids, and have been predicted to be involved in dimer formation and dimer-tetramer transition (Paper II, Frederiksen *et al.*, 2004).

The P-loop of the ATP binding site and binding to the phosphates

The glycine-rich P-loop makes the ATP binding site together with the LID region. The residues of the loop, GPMFSGKSTEL in the human TK1, form an anion hole that accommodates the phosphates of ATP. The glycine loop is the most conserved sequence

among nucleoside and nucleoside monophosphate kinases, presence of amino acids with a side chain at the positions occupied by glycines would probably collide with the bound phosphate, and/or make the loop structure to rigid to undergo a conformational change. The importance of the conserved residues in the P-loop of HSV1 thymidine kinase (amino acid nr 56-63, sequence: GPHGVGKT) was investigated by Liu & Summers (1988). When any one of the three glycines was changed to valine, or when lysine was changed to isoleucine, the corresponding mutant enzyme was inactive. Substitution of threonine by serine did not abolish the activity, but gave a 4-fold increase of K_m for thymidine and a 3-fold increase of K_m for ATP (Liu & Summers, 1988). Interaction of the side chain hydroxyl groups of serine and/or threonine with Mg^{2+}, as seen in the crystal structures of H-Ras p21 protein (Pai et al., 1990) and in modelling/calculational work on the Mg-binding site of HSV-1 TK (Kussmann-Gerber et al., 1998) is expected to stabilize binding of the Mg^{2+} ion (se later). The invariable lysine makes hydrogen bonds to both β- and γ-phosphate groups of ATP, as shown in the crystal structure of E. coli adenylate kinase (Müller & Schulz, 1992). Altogether, the β-phosphate makes five hydrogen bonds: three with backbone atoms of the P-loop, one with the invariable lysine (K13 of the P-loop in E. coli enzyme) and one with arginine (R123) situated at the beginning of the LID domain. The γ-phosphate, which is transferred during the reaction, is less well bound. It makes only two hydrogen bonds: to K13 and R123, while α-phosphate makes three hydrogen bonds with the backbone and side –chain atoms of the P-loop and one with R123 (Müller & Schulz, 1992). Several types of NMR experiments (1D proton, 2D – NOESY and COESY) by Byeon et al. (1995) on wild type and K21R and K21A mutants of adenylate kinase complexed with MgAp$_5$A[5] provided evidence that the main function of the invariable lysine is to orient the phosphate chain of MgATP in proper conformation for catalysis and to ensure interactions between the substrates and the active site residues. The K13Q substitution in E. coli adenylate

---

[5] Ap$_5$A, P^1,P^5-bis(5'-adenosyl)pentaphosphate, a two-substrate-mimicking inhibitor (although ATP and AMP together have four phosphate groups, their spatial separation seems to be matched by the five phosphate groups of Ap$_5$A).
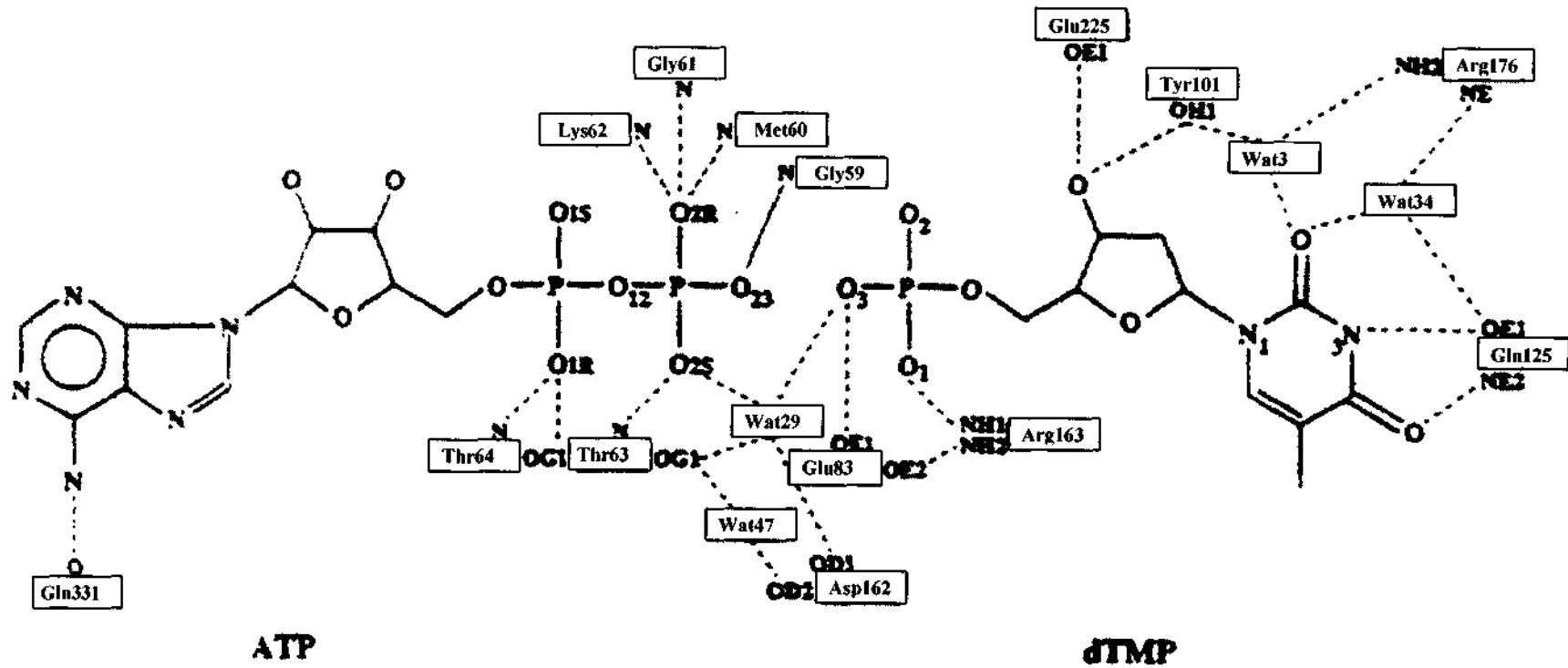
Figure 4.6. Sketch of all hydrogen bonds to the phosphate groups and the TMP substrate in the complex of HSV-1 TK with ADP and dTMP according to Wild *et al.* (1997).

kinase resulted in a mutant protein which had greatly reduced catalytic activity (K_{cat} 0.016 as compared to 305), and produced proton NMR spectra very different from that of the wild type, suggesting major rearrangements in the protein molecule (Reinstein *et al.*, 1990). Even conservative replacement of the lysine by arginine resulted in 10^5 decrease in the K_{cat}/K_m value (Byeon *et al.*, 1995). Fig. 4.6 shows the sketch of all hydrogen bonds to the phosphate groups in the complex of HSV-1 TK with ADP and dTMP according to Wild *et al.* (1997).

Mg^{2+} binding site

ATP interacts with many proteins as a complex with Mg^{2+} ion. In the early X-ray structure of yeast adenylate kinase with MgAP$_5$A complex, the Mg^{2+}ion was found in proximity to the phosphates (Egner *et al.*, 1987). An aspartate residue was proposed to coordinate Mg^{2+} as the reaction proceeds from MgATP to MgADP on the basis of studies combining NMR and X-ray structure of rabbit muscle ADK (Fry *et al.*, 1986). In the crystal structure of yeast adenylate kinase the conserved Asp89 (corresponding to the Asp93 in the mammalian muscle enzyme) was shown to hold Mg^{2+} ion *via* two water molecules (Abele & Schulz, 1995). Experiments with proton NMR indicated that binding of Mg^{2+} induced changes in the conformation of AP$_5$A when complexed to WT, but not when complexed to the D93A mutant (Yan & Tsai, 1991). Similarly, substantial differences in NMR spectra of WT AK1 were found depending on whether the enzyme was combined with AP$_5$A or with MgAP$_5$A. No such differences were observed in the D93A mutant, suggesting that Mg^{2+}induced changes in the conformation of WT but not of the D93A mutant (Yan & Tsai, 1991). Kinetically, the substitution D93A resulted in 4-fold higher K_m towards MgATP, and 650-fold decrease in K_{cat}. The conclusion was that the carboxylate group of D93 binds and fixes the Mg^{2+} ion, which in turn orients the phosphates to proper conformations (Yan & Tsai, 1991). The same role was assigned to D82 of vaccinia virus TK by mutagenesis and enzyme kinetics experiments (Black & Hruby, 1992). Modelling/calculational work of Kussman-Gerber *et al.* (1998) pointed to D162, together with T63, as the Mg^{2+} fixing residue in HSV-1 TK (Fig. 4.7A). Until now there are no crystal structures of HSV-1 TK containing Mg^{2+} ion, but crystal structures of *Dm*-dNK in ternary complex with dTTP and Mg^{2+} (Mikkelsen *et al.*, 2003; Fig. 4.7B) and of human dCK in complex with dC, ADP and Mg^{2+} (Sabini *et al.*, 2003) have been solved. E127 in dCK (Sabini *et al.*, 2003) aligns with

E104 in *Dm*-dNK and with D162 of HSV-1 TK (Eriksson *et al.*, 2002). These residues are similarly placed in the β3 strand of their respective enzymes (Wild *et al.*, 1997; Johansson *et al.*, 2001), in the ERS motif strictly conserved in dCK, dGK, TK2 and *Dm*-dNK (Eriksson *et al.*, 2002). E127 of dCK and E104 of *Dm*-dNK were shown to coordinate magnesium in their crystal structures: dCK complexed with ADP, dC and Mg^{2+} (Sabini *et al.*, 2003), and *Dm*-dNK complexed with dTTP and Mg^{2+} (Mikkelsen *et al.*, 2003; Fig. 4.7B). The analogous and therefore the putative Mg^{2+} binding residue in human TK2 is E133, and in human TK1 D97 (Fig. 4.2 and Fig. 4.3). An obvious difference between human TK1/vacinia TK and HSV-1 TK/*Dm*-dNK is that in the former the Mg^{2+} coordinating residue is not positioned in the middle of the thymidine binding site (Fig. 4.2).
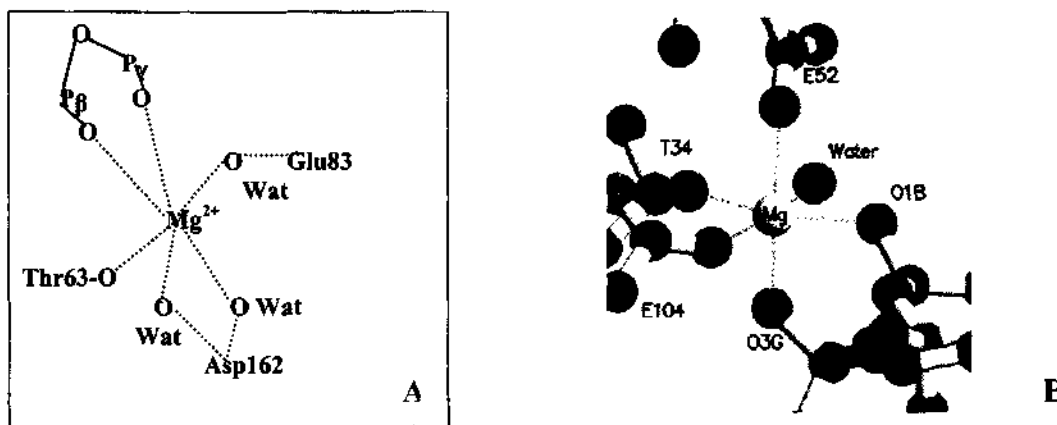
Figure 4.7. Mg^{2+} coordination. A. in HSV1 TK resulting from calculations of the most favourable interaction energies and comparison with ADK structure according to Kussmann-Gerber *et al.* (1998). B. to the dTTP phosphate oxygens (to the right), protein atoms and a water molecule in dTTP interaction with dNK (after Mikkelsen *et al.*, 2003).

LID region and interaction with substrates

The interaction of several conserved LID arginines with the phosphoryl groups of both substrates was seen in the closed crystal structure of yeast adenylate kinase, complexed with AP$_5$A (Müller & Schulz, 1992; Abele & Schulz, 1995). In the open structure, the arginines of the LID could not be reached by the phosphates of ATP. Analogous arginines are found in motif R*v*XXRXR (*v* can be valine, leucine or isoleucine) preserved in herpesviral thymidine kinases (Balasubramaniam *et al.*, 1990; Gentry, 1992) and in the

Dm-dNK-dCK-dGK-TK2 family (Johansson *et al.*, 2001). From a kinetic investigation of point mutants of human AK1, where conserved arginines were replaced with alanyl residues it was concluded that: R44 (from α2 helix in adenylate kinase, corresponding to α2 helix in *Dm*-dNK, and so not in the LID region), R138 and R149 (in the LID) make closer contact with AMP than with MgATP. R132 (in the LID) makes a similar degree of contact with both AMP and MgATP, and while R132, R138 and R149 are important for catalytic activity (mutants show substantial decrease in K_{cat}), R44 is probably involved in AMP binding only, and not in catalysis (Kim *et al.*, 1990). Two conserved aspartates, D140 and D141 in the chicken muscle adenylate kinase (the same numbering as in the human enzyme) were shown to form salt bridges to R132 and R138, and in this way to assist the arginines in stabilizing the transition state (Dahnke & Tsai, 1994). Also, several lysines conserved in mammalian species of ADK were shown to be essential for catalytic efficiency of the enzyme through interaction with both MgATP and AMP substrates (Ayabe *et al.*, 1997). The analogous arginines in uridylate kinase (UMPK) have been shown to interact with the γ-phosphate of ATP (Schlichting & Reinstein, 1997). When compared to AMPK or UMPK the eukaryotic thymidylate kinase contains two fewer arginine residues that can interact with the transferred phosphoryl groups (Ostermann *et al.*, 2000), and because activity of the isolated human TMPK is low, the authors speculate that an additional factor/protein may supply further catalytic residues and accelerate enzymatic reaction in the cell (Ostermann *et al.*, 2000 compare this situation to p21[ras] where GTPase-activating protein GAP, in addition to assembling the residues at the active site, provides a catalytic residue). Alternatively, at least in the yeast thymidylate kinase, an arginine residue present in the P-loop could make hydrogen bonds to the phosphoryl groups (Lavie *et al.*, 1998). Similarly to the thymidylate kinase, the putative LID region of the human TK1 has two arginines only, R158 and R165 (Fig. 4.2 and Fig. 4.3). The hydrogen bonds to the phosphates could then be supplied by R38, R39, R41 and R42 present in the α1-helix, just next to the P-loop and in the proximity of the reaction centre.

The adenine and ribose of ATP were found not to bind very tightly in several crystal structures of adenylate kinases (Müller & Schulz, 1992; Müller-Dieckmann & Schulz, 1994; Abele & Schulz, 1995). The ribose probably does not interact with any protein atom directly, but rather is stabilized by several water molecules in the crystal structure of deoxyguanylate kinase (Sekulic *et al.*, 2002). Similarly, in the crystal structure of *Herpes*

simplex virus type 1 thymidine kinase, the adenine ring of ATP was found to be sandwiched between a strongly conserved arginine (R216) and a glutamine (Q331), and so to make only minimal contacts with the enzyme (Evans *et al.*, 1998). Q331 makes the single hydrogen bond to the adenosine moiety of ATP, while ribose makes none (Wild *et al.*, 1997; Fig. 4.6). The weak binding of the adenosine moiety to HSV-1 TK is reflected in a rather high K_m value for ATP of 70 μM (Wild *et al.*, 1997).

In summary, the phosphate donor ATP binding site was found to be very similar in HSV-1 TK and in adenylate kinase, as well as in other so far investigated nucleoside and nucleotide kinases. The catalytic transfer of a phosphate group to the substrate must then occur at very similar sites and involve similar amino acid residues. A view of the binding pattern of ATP in HSV1-TK is shown in Fig. 4.6.

The thymidine binding site

The binding site for the phosphate accepting substrate must be responsible for the substrate specificity of the enzyme. NMP kinases interact extensively with their NMP substrates what results in specificity towards the base of the phosphate acceptors. The adenine moiety of AMP for example, is fixed by as many as five hydrogen bonds in the AMP binding pocket of adenylate kinases (Yan & Tsai, 1999). Substitution of some critical residues may result not only in decreased catalytic efficiency but also in alteration of NMP specificity. For example, replacement of conservative residues that make hydrogen bonds to the adenine ring (T39 and Q101 of the chicken muscle adenylate kinase) by alanine, caused remarkable increases in activities (K_{cat}/K_m) for UMP and CMP with concomitant decrease in activity-with AMP (Okajima *et al.*, 1993a; Okajima *et al.*, 1993b). Recent results of Knecht *et al.* (2002b) have shown that in *Drosophila melanogaster* a few amino acid substitutions convert deoxyribonucleoside kinase specificity from pyrimidines to purines. The specificity of thymidylate kinase became obvious when its crystal structure was solved: the cavity in which the base of dTMP was bound was too small to accommodate purine bases, which limited specificity to pyrimidines (Lavie *et al.*, 1997). A hydrogen bond between O4 of the thymidine base and a conserved arginine (R73 in the human thymidylate kinase) favoured thymidine and uracil over cytosine (Lavie *et al.*, 1997).

During the 1990s several studies elucidated the importance of amino acid residues at the nucleoside binding site of HSV1 TK (Munir *et al.*, 1992; Munir *et al.*, 1994; Black &

Loeb, 1993; Michael *et al.*, 1994; Michael *et al.*, 1995; Black *et al.*, 1996; Pilger *et al.*, 1999). After 1995 these data could be supported by several crystal structures (Brown *et al.*, 1995; Wild *et al.*, 1995; Wild *et al.*, 1997). The recently solved structures of *Drosophila* deoxyribonucleoside kinase, *Dm*-dNK, and of the human deoxyguanosine kinase, dGK, have explained the structure-function relationship regarding the specificity of the phosphate accepting substrate (Johansson *et al.*, 2001). The wider substrate cleft of *Dm*-dNK dominated by hydrophobic residues results in its broad substrate specificity, while the strict purine specificity of dGK is explained by its tighter binding pocket containing charged residues, and especially R118, which provides hydrogen bonding interaction with the N7 and O6 of the purine ring (Johansson *et al.*, 2001).

The substrate clefts in *Dm*-dNK, dGK and HSV1-TK have some common features: a conserved Q (Q81, Q111 and Q125 respectively) residue makes hydrogen bond to the base, a conserved E-Y couple (E172-Y70, E211-Y100 and E225-Y101 respectively) makes hydrogen bonds to the 3'OH moiety of the sugar, and a conserved E-R couple (E52-R169, E70-R208 and E83-R222 respectively) makes hydrogen bonds to the 5'OH of the sugar (Johansson *et al.*, 2001; se also Fig. 4.6 and Fig. 4.8).

The substrate binding site of *Herpes simplex* thymidine kinase is composed of amino acid side-chains derived primarily from a set of roughly parallel helices: α2 (residues 85-89), α3 (residues 96-108), α4 (residues 114-130) and α6 (residues 171-178) (Evans *et al.*, 1998). The way in which thymidine kinase of *Herpes simplex* virus accommodates its substrate has been elegantly summarized by Sulpizi *et al.* (2001). Their description and the accompanying figure (Fig. 4.8) is based on both theoretical quantum-mechanical calculations (Alber *et al.*, 1998) and experimental approaches (Brown *et al.*, 1995; Wild *et al.*, 1995; Wild *et al.*, 1997; Kussmann-Gerber *et al.*, 1998; Champness *et al.*, 1998; Bennett *et al.*, 1999; Prota *et al.*, 2000).
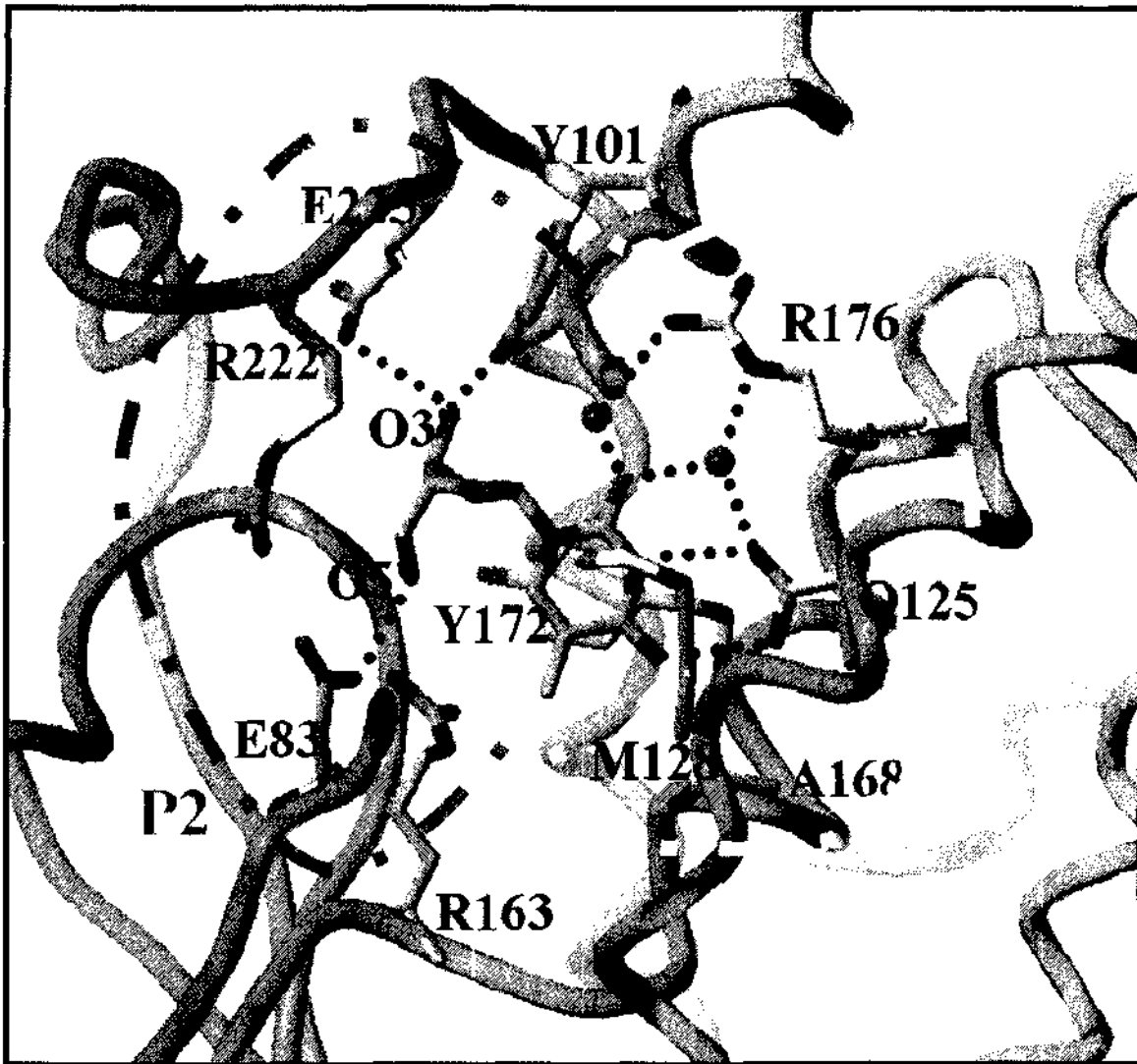
Figure 4.8. Structure for the HSV-1 thymidine kinase active site. P1 and P2 accomodate the nucleobase and the sugar moiety respectively. The water molecules are shown as *red spheres*, and the hydrogen bonds are depicted as *dotted lines*. Thymidine and amino acids are displayed as capped sticks The colour code is: *orange*, carbon; *red*, oxygen; *blue*, nitrogen; *yellow*, sulfur. Reproduced from Sulpizi *et al.* (2001).

Sulpizi *et al.* (2001) divide the thymidine binding site into two pockets, P1 and P2. In P1, the nucleobase moiety is stabilized by two hydrogen bonds (Fig. 4.6 and Fig. 4.8) with the highly conserved Q125 from the interface helix α4, and by means of two water molecules with R176 from helix α6 (Kussmann-Gerber *et al.*, 1998). The thymine is further fixed in a sandwich-like orientation by Y172 from helix α6 and M128 from helix α4 (Alber *et al.*, 1998; Wild *et al.*, 1997; Kussmann-Gerber *et al.*, 1998), and also in close proximity to A168 (Kussmann-Gerber *et al.*, 1998), found just before helix α6. Modulation of the residue size at position 128 had a direct impact on binding affinity: M128A had no

biological activity, indicating that the small alanine side chain is not sufficient to stabilize the thymine within the active site (Pilger et al., 1999). Similarly, exchange of A168 for a bigger threonine resulted in dramatic reduction in the affinity of the enzyme for substrate analogues carrying bulky substitutions on the C5, such as BVdU (E-5-bromovinyl-2'-deoxyuridine) (Darby et al., 1986). Phenotypically A168T mutant of HSV1 TK resembles that of HSV2 strain which has A168S substitution (Darby et al., 1986; Wild et al., 1997). In the crystal structure of HSV1 TK wild type enzyme the bound deoxythymidine was observed to leave a 35 Å^3 cavity close to C5 position of the base, and two water molecules close to position C2. It is proposed that the cavity and water molecules reduce substrate specificity to such an extent that TK can phosphorylate analogues with modified ribose and various thymine, guanine (Wild et al., 1997) and adenine (Vogt et al., 2000) analogues.

The second pocket in the thymidine binding site, P2 (Sulpizi et al. (2001; Fig. 4.8) accommodates the sugar moiety of the nucleoside. The 3'-OH forms specific H bonds with Y101 from helix α3 (aa 96-108) and with E225 that comes from a loop in the LID region comprising residues 219-226, whereas 5'-OH forms a direct H-bond or water-mediated interaction with R163 (positioned just before helix α6), E83 (positioned prior to helix α2) and R222 (positioned as E225 in the LID) (Fig. 4.8). The importance of E225 for catalysis was demonstrated in mutant E225L which showed 60-fold increase in K_m toward thymidine and 20-fold decrease in K_{cat} (Pilger et al., 1999). The O-4' of the sugar ring (positioned between C-1'and C-4') does not interact with polar or charged groups of the protein but points toward hydrophobic environment of side chains of W88, I97 and M128 (Sulpizi et al., 2001). From modelling and energy calculations of enzyme-substrate complexes De Winter & Herdewijn (1996) concluded that more than 80% of the total interaction energy between HSV1-TK and thymidine comes from the interaction with only six residues, namely the already mentioned four hydrogen bond partners E83, Y101, Q125 and E225, and residues M128 and Y172 which sandwich the nucleobase by means of stacking interactions. This strong binding of thymidine is in agreement with its low K_m value of 0.2 μM (Michael et al., 1995) and with its location deeply in the binding pocket (Champness et al., 1998).

The reaction mechanism

Transfer of γ-phosphate from ATP to 5'-OH of deoxynucleoside is believed to involve a nucleophilic attack of the activated 5'-OH group on the γ-phosphate of the phosphate donor. E52 in *Dm*-dNK (Johansson *et al.*, 2001) and the corresponding E83 in HSV-1 TK (Wild *et al*, 1997) supply the carboxylate group in the vicinity of the 5'-OH at the active site of the enzyme. The residues are suggested to act as a base that deprotonates 5'-OH and makes the nucleophilic attack possible (Fig. 4.9). The proximity of positive charges from arginines and the Mg^{2+} makes the phosphorus atom more electrophilic.
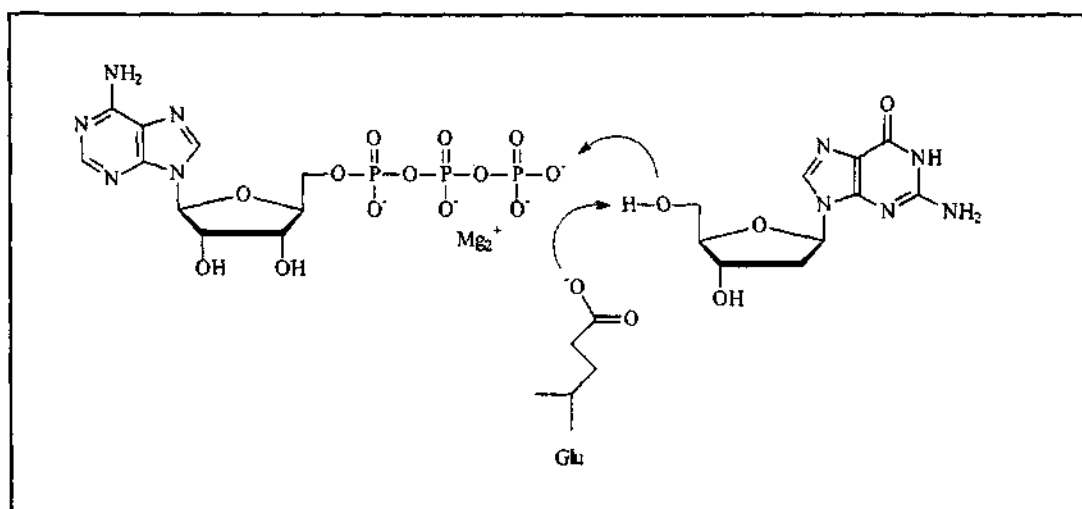
Figure. 4.9. The mechanism of action of the deoxyribonucleoside kinases. The phosphate donor ATP is to the left and the substrate dG to the right. Glu is suggested to act as a general base and Mg^{2+} as a counter ion. After Eriksson *et al.* (2002).

The catalytic glutamates in deoxynucleoside kinases with solved 3D structure, E52 in *Dm*-dNK, E70 in dGK and E83 in HSV1 TK are all placed either at the end of the β2 strand (Johansson *et al.*, 2001) or just after the β2 strand (Wild *et al.*, 1997). D58 in the human TK1 could then be the active base that deprotonates 5'-OH group (Fig. 4.3).

Conformational changes during catalysis

After elucidation of several structures of adenylate kinases (Pai *et al.*, 1977; Fry *et* al., 1986; Diederichs & Schulz, 1990; Schulz *et al.*, 1990; Schlauderer *et al.*, 1996), uridylate kinase (Müller-Dieckmann & Schulz, 1994; 1995) and guanylate kinase (Blaszczyk *et al.*, 2001; Sekulic *et al.*, 2002) with and without bound ligands it became clear that nucleotide kinases undergo large conformational changes during the catalytic cycle. Adenylate kinase
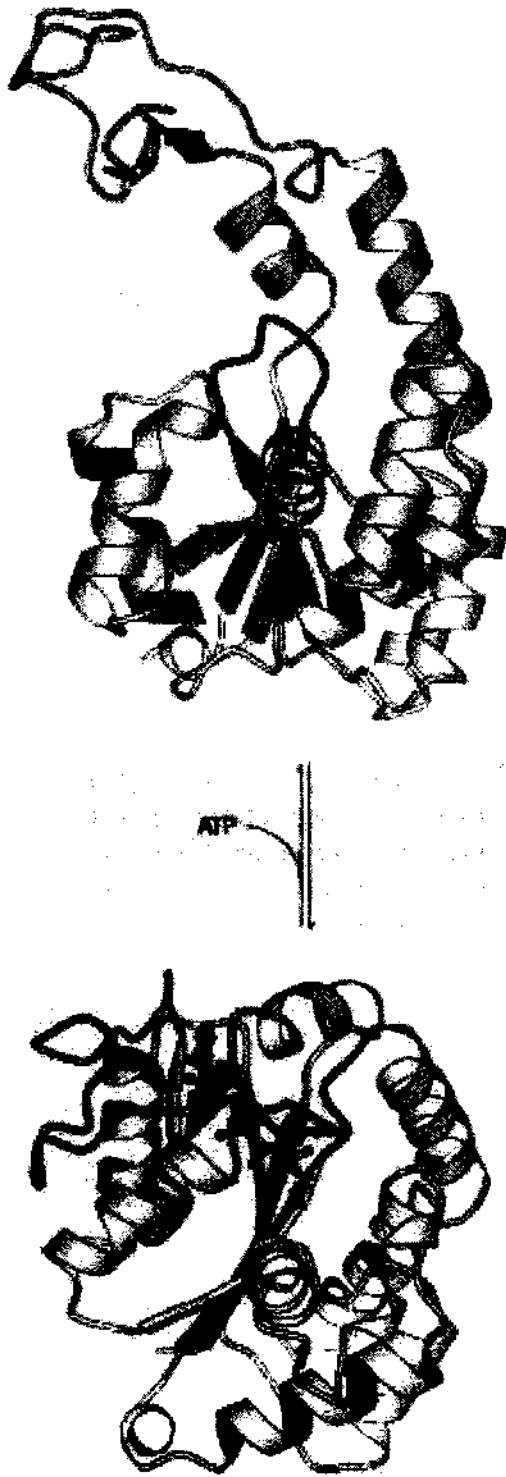
Figure 4.10. Conformatioal changes in adenylate kinase upon bnding of ATP (in ball and stick representation). P-loop is shown in green and the LID domain in yellow. Reproduced from Biochemistry, Berg *et al.* (eds.), 2002.

became a classic example of induced fit: binding of ATP induced closure of the P-loop on top of the polyphosphate chain and extensive interaction with the β-phosphoryl group (Fig. 4.10). NMR experiments of Fry *et al.* (1986) estimated that binding of the MgATP substrate involves displacement of the P-loop by 6 Å, much larger than changes in the other regions of the adenylate kinase. This movement of the P-loop permits the LID domain to move down over the bound ATP. On LID closure, its conserved aspartates form salt bridges with conserved arginines that in turn fix the phosphoryl groups, so that ATP's γ-phosphate is kept in position close to the phosphate acceptor site. Binding of the second substrate induces additional conformational changes that make possible transfer of the phosphoryl group from donor to acceptor, and not to water. Various intermediate states of the movements performed by adenylate kinase during catalysis, starting with substrate-free crystals, in complex with only one substrate, both substrates, or the bisubstrate inhibitor AP$_5$A, have been frozen and put together in a video (Vonrhein *et al.*, 1995). The thymidine kinase of *Herpes simplex* virus (HSV1) was also shown to exhibit similar structural movements on binding of thymidine and ATP (Perozzo *et al.*, 2000 – isothermal titration calorimetry study; Wurth *et al.*, 2001 – thermal denaturation and circular dichroism study).

The interface architecture

While most NMP kinases are monomeric, the thymidine/thymidylate kinase of *Herpes simplex* virus exists as a homodimer, and the structure of its dimer interface has been determined (Wild *et al.*, 1995; 1997; Evans *et al.*, 1998). The interface is dominated by helices α4, α6 and α10 which participate fully and α2 and α12 which are partially involved (Wild *et al.*, 1997). The contacts include both hydrogen bonds and hydrophobic interactions and the closest monomer-monomer contacts involve: Q185 to L91 and G92, V119 to V119 and S123, and N306 and V307 to E371 (Evans *et al.*, 1998). The bulk side chain of tryptophan W310 was found to protrude from one monomer face into the opposing monomer and to make mostly hydrophobic interactions as its five-membered ring was exposed to solvent. The polar contacts between N306 in one monomer and E371 in the other monomer reinforced the hydrophobic interactions. Additional monomer-monomer contacts were mediated by water molecules (Evans *et al.*, 1998). The interface of HSV-1 TK was found to be exceptionally non-polar, with 65% hydrophobic, 26% polar and 9%

ionogenic residues (Wild et al., 1997), as compared to the average amino acid composition of interfaces of 32 different protein dimers (Jones & Thornton, 1995).

The dimer interfaces of dNK and dGK are very similar (Johansson et al., 2001). Hydrophobic residues from helices α4 and α6 interact and form a four-helix bundle at the dimer interface. In dNK these residues include four aromatic, four aliphatic and four small residues from each subunit, whereas in dGK they contain several aromatic residues (Johansson et al., 2001; Fig. 4.11).

Figure 4.11. dNK dimer: hydrophobic residues from helices α4 and α6 in each subunit interact forming a four-helix bundle at the dimer interface. After Johansson et al. (2001).

As mentioned earlier, α2 and α3 helices (residues 83-85 and 104-113 respectively) of human TK1 presumably correspond to the interface of Dm-dNK, and so to the interface of HSV-1 TK as well. Besides, these two helices align to α4 of HSV1-TK in Clustal W, and to residues 78-98 of vaccinia TK (Fig. 4.1). The D^{89}IVEFCERMA segment of vaccinia TK was predicted to be an α helix by circular dichroism (Behrends et al., 1996), and a conserved glycine, G84, was shown to be involved in structural integrity of the enzyme, as G84V mutant lost both enzymatic activity and tetrameric conformation (Black & Hruby, 1992). The structure of the putative interface of human TK1 and the importance of residue 106 for the enzyme kinetics and conformation are investigated in details in the next chapters.

PART II

EXPERIMENTAL WORK

Chapter 5

Materials and methods

E. coli bacterial strains and their growth

KY895: F⁻, *tdk-1*, *ilv*, a thymidine kinase-deficient derivative of *E. coli* strain K12 (Igarashi *et al.*, 1967; Hiraga *et al.*, 1967) and **BL21**: F⁻, *ompT*, *hsdS* (r_B^-, m_B^-), *gal* (Amersham Pharmacia Biotech) were used for propagation of plasmids and expression of recombinant protein. BL21 is a B strain of *E. coli* that is naturally deficient in Lon protease and lacks the endoprotease OmpT due to *ompT* gene deletion (Grodberg & Dunn, 1988). The strains were grown in LB medium at 37°C (unless indicated otherwise) and with 50 µg/ml ampicillin for the plasmid containing strains.

Epicurian Coli XL1-Blue supercompetent cells: *recA1*, *endA1*, *gyrA96*, *thi-1*, *hsdR17*, *supE44*, relA1, lac[F', *proAB*, *lacI^qZΔM15*, *Tn10* (tet^R)] (Stratagene) were used for transformation of DNA after site directed mutagenesis. Epicurian Coli have much higher transformation efficiency than the parental strains thanks to EndA⁻ phenotype. The strain was grown in NZY+ (casein hydrolysate with yeast extract broth) and treated according to the instructions from the supplier.

MT102: MC1000, r⁻ m⁺, Δ*lacX74* Δ(*ara-leu*)7697 *strA* (from Ole Skovgaard, originally from Mogens Trier Hansen, Novo) was used for transformation of ligation mixtures.

DH5α max efficiency competent cells: F⁻, φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*) *recA1* *endA1* *thi-1* (Life Technologies, now Invitrogen) were used by Tina Kristensen for transformation of recombinant PCR products in construction of pGEX-2T-TK1Δ40 and TK1Δ60 deletion mutants (se later; Kristensen, 1996).

pGEX-2T plasmid vector

pGEX-2T is a pBR322 derivative that contains the ampicillin resistance gene, and it directs synthesis of foreign proteins as fusions to a glutathione S-transferase (GST) encoded by the parasitic helminth *Schistosoma japonicum* (Smith & Johnson, 1988). The vector is engineered so that the desired protein is fused to GST via a thrombin recognition site. The cloned gene is introduced into the multiple cloning site situated downstream of the coding sequence for thrombin cleavage site, and the cloned protein can be cleaved from the GST partner with thrombin (Smith & Johnson, 1988). Thrombin cleavage yields the desired protein with 2 additional amino acids (G and S) at its N-terminus, resulting from the thrombin recognition sequence (LVPR↓GS). The expression of the fusion protein is under control of lacI^Q, that synthesizes 10 times more Lac repressor than the wild type lacI gene, and under control of an IPTG-inducible tac promoter, a hybrid composed of *trp* and *lac* promoter sequences (de Boer *et al.*, 1983). pGEX-2T was purchased from Amersham Pharmacia Biotech and is shown in Fig.5.1.
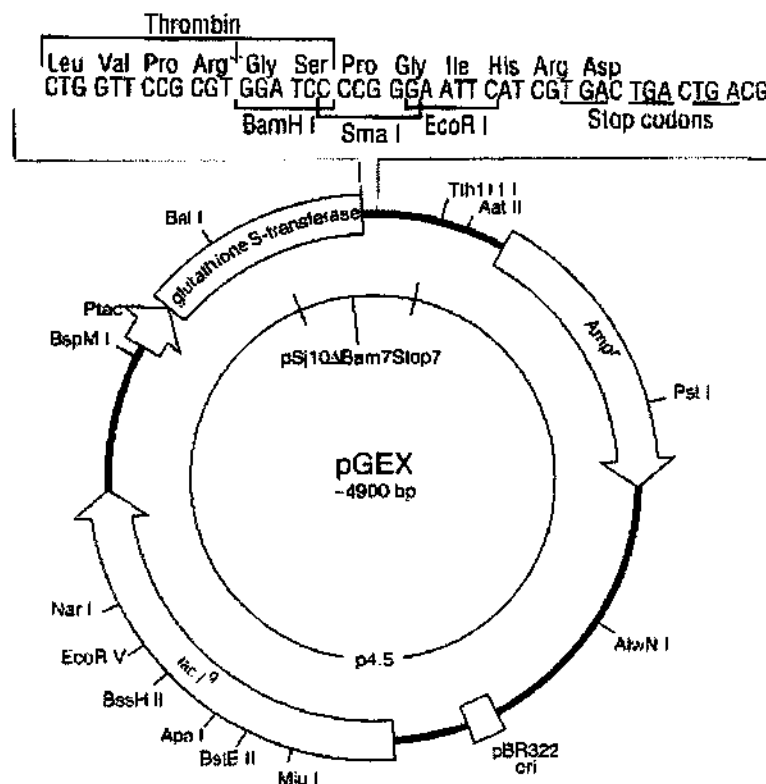
Figure 5.1. Map of the glutathione S-transferase pGEX-2T fusion vector. From GST Gene Fusion System Manual, Pharmacia Biotech (1997).

Recombinant plasmids

The following section describes construction of several recombinant plasmids. All numbering used is according to Bradshaw & Deininger (1984), where the translation initiation is at position 58, and therefore codon for amino acid 106 starts at base number 373. The altered bases are in bold type and the restriction sites, if introduced, are in italics.

Construction of pGEX-2T-TK1

pGEX-2T-TK1 contains amino acid coding region of TK1 from normal human lymphocytes. This plasmid was constructed by Dr. Hofbauer (Vienna Biocenter, Austria) and Dr. Folkers (Swiss Federal Institute of Technology, Zurich, Switzerland), and the construction procedure has been described in detail (Paper I, Berenstein *et al.*, 2000). For insertion into the *Bam*HI/*Eco*RI restriction sites of pGEX-2T and introduction of the thrombin cleavage site between the GST and the TK1 proteins, it was necessary to modify the N-terminal of TK1 to start with GSMCS instead of MCS and the C-terminal to end with ILQCMQA instead of ILQCSPAN. Neither the first 10 N-terminal amino acids, nor the C-terminal amino acids are evolutionary conserved and were therefore expected to be without importance. This expectation was confirmed by the experimental results.

The plasmid pGEX-2T-TK1 is the same as pGEX-2T-TK1^{Val106}, the latter designation is used to stress that the TK1 coding sequence contains the code for valine at amino acid position 106.

Construction of pGEX-2T-TK1^{Met106}

Two pGEX-2T-TK1^{Met106} plasmids have been constructed in our laboratory using pGEX-2T-TK1^{Val106} as the template DNA.. The codon GTG for valine corresponding to amino acid position 106 was replaced by codon ATG for methionine using the QuickChange™ site directed mutagenesis kit from Stratagene according to the protocol supplied. This mutagenesis method utilizes *PfuTurbo*™ DNA polymerase for high fidelity replication and the DNA methylation-specific *Dpn*I endonuclease to digest the parental DNA template. The cycling conditions were: 16 cycles, 95°C for 30 seconds, 55°C for 1 minute, and 68°C for 10 minutes in 50 µl volume. Construction of the first pGEX-2T-TK1^{Met106} plasmid,

here referred to as pGEX-2T-TK1^{Met106A}, was performed by a student project group (Mørch & Petersen, 1997) with the following primers: sense primer 5'-TTCCCTGACATC<u>ATG</u>*GAATTC*TGCGAGGCC(361-390)-3', and the antisense primer 5'-GGCCTCGCA*GAATTC*<u>CAT</u>GATGTCAGGGAA(390-361)-3'. The ATG codon replacing GTG at position corresponding to amino acid 106 is underlined. The base changes G to A in the sense primer and C to T in the antisense primer wee introduced in order to get a control *Eco*RI restriction site without changing the amino acid.

Construction of the second pGEX-2T-TK1^{Met106}, that did not contain a restriction site, here referred to as pGEX-2T-TK1^{Met106B}, was performed by H. Frederiksen (at that time a master student in our lab.) and me with the following primers: sense primer 5'-TTTTTCCCTGACATC<u>ATG</u>GAATTCTGCGAGGCC(358-390)-3', and antisense primer 5'-GGCCTCGCAGAATTC<u>CAT</u>GATGTCAGGGAAAAA(390-358)-3'. We judged that introduction of an extra base change and generation of a control restriction site was not necessary due to the high efficiency of generation of the correct mutants by the QuickChange^{TM} site directed mutagenesis kit. We also believe that properties of TK1^{Met106A} and TK1^{Met106B} are the same and therefore all our *in vitro* results are directly comparable.

Construction of pGEX-2T-TK1^{Val106X} mutants

The plasmid pGEX-2T-LyTK1^{Val106} was used as template DNA in PCR reactions, and mutations into the GTG codon coding for valine 106 were introduced with the Quick Change^{TM} site-directed mutagenesis kit from Stratagene (according to the instructions of the manufacturer).

The sense 5'-TTTTTCCCTGACATC<u>GTG</u>GAGTTCTGCGAGGCC(358-390)-3' and antisense 5'-GGCCTCGCAGAACTC<u>CAC</u>GATGTCAGGGAAAAA(390-358)-3'

mutagenic primers were substituted as follows in the target codon for Val-106 (underlined): GCG/CGC for Ala-106, **CAG/CTG** for Gln-106, **GGT/ACC** for Gly-106, **CAC/GTG** for His-106, **ATC/GAT** for Ile-106, **CTG/CAG** for Leu-106, **AAA/TTT** for Lys-106, **ATG/CAT** for Met-106 and **ACC/GGT** for Thr-106. The codons are given in sense/antisense primer respectively. The resulting recombinant enzymes were designated V106A, V106Q, V106G, V106H, V106I, V106L, V106K, V106M, V106T, and rLy-TK1^{Val106} was designated V106WT.

Construction of pGEX-2T-TK1[66-136(Val106)]

PCR reaction

Plasmid pGEX-2T-TK1 was used as a template in a PCR reaction with a sense primer:

5'- GGG*GGATCC*TGCACACATGACCGGAACACC(247-273)-3' designed to contain a GGG overhang, and an antisense primer:

5'-CGGCACC*GAATTC*TAGATGGCCCCAAATGGCTTCCT(480-445)-3'.

*Bam*HI restriction site was introduced to the sense primer, and *Eco*RI restriction site and an UAG stop codon (underlined) to the antisense primer. Reactions were run under following conditions: 30 cycles, 94°C for 1 min, 55°C for 1 min, and 72°C for 1 min in 25 µl volume with *Thermus aquaticus* DNA polymerase (Stratagene), and contained 4 µg/ml template DNA, 3mM $MgCl_2$, 0.2 mM of each dNTP and 0.36 µM of each primer in 10 mM Tris-HCl buffer (pH 8.3). Two PCR reactions were run independently of each other.

Purification of the PCR products

The PCR product was run on 1.5% TAE agarose. The DNA fragments of the right size were purified by freezing the agarose block and extracting the water phase with phenol. DNA was then ethanol precipitated, resuspended in E-buffer and visualized on 1% TBE agarose gel. The concentration of DNA was at this stage about 1ng/µl.

Ligation of the PCR product with the pGEX-2T vector and transformation into competent E. coli MT102

pGEX-2T and the PCR product were separately digested with *Bam*HI and *Eco*RI, mixed at the insert : vector ratio of 2 : 1, and ligated with T4 ligase at 16°C over night. The ligation mixture was transformed into competent MT102 strain.

Screening the transformants for the insert

The transformants were grown up in LB with 100 µg/ml ampicillin and plasmids isolated with Wizard Plus Minipreps (Promega). The plasmid DNA was digested with *Bam*HI and *Pst*I in order to estimate the size of the insert (the pGEX-2T vector has one restriction site for *Bam*HI at position 930 and one restriction site for *Pst*I at position 1897, so the digested vector without insert will give a fragment of 967 base pairs and a fragment of about 4000 base pairs. Because the coding region of the cloned interface was about 220 base pairs, fragments of about 1190 base pairs and 4000 base pairs were expected). Plasmid minipreps from 26 transformants obtained in the first PCR reaction have shown 25 transformant without any insert, and 1 transformant that gave fragments of 1090 base pairs and 4000

base pairs. Sequencing of the clone has shown a G to A mutation at base position 378, resulting in a sequence GTGG*AATTC*(373-381) with an *Eco*RI restriction site about 100 bases earlier than designed. The resulting clone had no AUG stop signal, and was missing over 30 amino acids.

*Bam*HI/*Pst*I digestion of plasmid DNA from 20 transformants obtained in the second PCR reaction, resulted in 4 clones with the correct insert size, but subsequent sequencing of the clones has shown no UAG stop signal, nor the expected *Eco*RI recognition site, but the original template sequence. Because, with the exception of the stop signal, the sequence of the clones was correct, I have decided to introduce the UAG stop signal to one of the clones instead of the codon CTG(466-468) by site directed mutagenesis.

Site directed mutagenesis

Site directed mutagenesis was performed with "QuickChange™ Site-Directed Mutagenesis Kit".(Stratagene) according to the manufacturers instructions. This mutagenesis method makes use of *PfuTurbo*™ DNA polymerase which has 6-fold higher fidelity in DNA synthesis than *Taq* DNA polymerase. Following mutagenic primers were used: sense primer: 5'-CCATTTGGGGCCATC<u>TAG</u>AACCTGGTGCCGCTG(451-483)-3' and antisense primer: 5'-CAGCGGCACCAGGTT<u>CTA</u>GATGGCCCCAAATGG(483-451)-3'. The stop codon is underlined. Plasmid DNA was prepared and sequenced from 10 Epicurian Coli XL-1-Blue transformants, 9 of the clones had the correct base sequence. One of the correct clones was used to transform KY895 (se below).

Construction of pGEX-2T-TK1[66-136(Met106)]

The codon GTG at positions 373-375 (underlined), corresponding to amino acid 106, was mutated to ATG by the "QuickChange™ Site-Directed Mutagenesis Kit" (Stratagene) with following primers :sense primer:

5'-CAGTTTTTCCCTGACATC<u>ATG</u>GAGTTCTGCGAGGCCATG(355-393)-3'

and antisense primer:

5'-CATGGCCTCGCAGAACTC<u>CAT</u>GATGTCAGGGAAAAACTG(393-355)-3'.

The potentially correct plasmids were transformed into Epicurian Coli XL1-Blue supercompetent cells, plasmids were re-isolated, transformed into BL21 and sequenced.

Construction of pGEX-2T-TK1Δ40 and pGEX-2TΔ60

The two plasmids have been constructed by Tina Kristensen, a former Ph.D. student at our lab. Because the cloning procedure has not been published, I will describe it shortly here.

The method used was recombination PCR for site-specific mutagenesis developed by Jones & Howard (1991) and Jones & Winistorfer (1992). For each mutant two separate PCR reactions were run with pGEX-2T-TK1 template (Fig. 5.2). Each PCR reaction employed a mutagenic primer (primer 1 and primer 3 respectively), which introduced a stop codon and a control *Sac*I restriction site, and a non-mutagenic primer used to delimit the amplified fragment (primer 2 and primer 4). The primer set in the second reaction was complementary in base sequence and antiparallel to the primer set in the first reaction, so that ends of one PCR product were homologous to the ends of the other PCR product (Fig. 5.2). The recombinant of interest resulted from recombination *in vivo*.
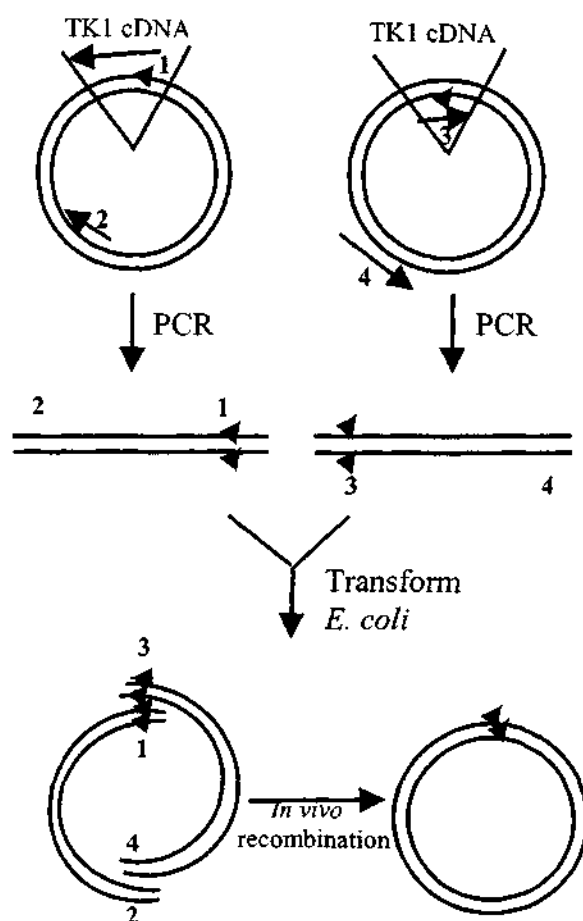
Figure 5.2. A sketch of recombination PCR. Numbering refers to the number of primer. Mismatches are indicated by filled triangles. The two PCR reactions result in two fragments with homologous ends. After Jones & Winistorfer (1992) and Kristensen (1996).

Following mutagenic primers were used: sense primer

5'-TACTTCAAGAAGGCC*TGAGCT*CAGCCTCCGGGCCG(622-657)-3' and antisense

primer: 5'-CGGCCCGGCAGGCT*GAGCTCA*GGCCTTCTTGAAGTA(657-622)-3' to

construct pGEX-2T-TK1Δ40, and sense primer:

5'-GTCGAGGTGATTGGG*TGAGCT*CACAAGTACCACTCC(571-606)-3' and antisense

primer:5'-GGAGTGGTACTTGT*GAGCTCA*CCCAATCACCTCGAC(606-571)-3' to

construct pGEX-2T-TK1Δ60. The stop codon is underlined, *Sac*I restriction site is in

italics. The non-mutagenic primers were:

sense primer: 5'-CTTATCAGACCGTTTCCCGCGTGG(3343-3366) and

antisense primer:5'-CCACGCGGGAAACGGTCTGATAAG(3366-3343). The numbering

of the non-mutagenic primers refers to the numbering of pGEX-2T bases according to

Pharmacia Biotech.

The PCR reactions were run in a total volume of 50 µl at following conditions: 0.24 µg/ml

linearized template DNA, 0.36 µM of each primer, 200 µM of each dNTP, 1.5 mM MgCl$_2$,

1.5 units *Thermus aquaticus* DNA polymerase in Tris-HCl buffer (pH 8.3) containing 50

mM KCl, 0.015% gelatine and 0.1% Tween 20; 30 cycles; 1 min at 95 C, 1 min at 58 C

and 3.5 min at 72 C. The PCR products were purified from low-melting agarose gel and

used to transform DH5α competent cells according to the instructions of the vendor (Gibco

BRL/Life Technologies). The potentially correct plasmids were transformed into *E. coli*

KY895 and sequenced on both strands with Sequenase™ version 2.0 DNA Sequencing Kit

(United States Biochemical).

Transformation into *E. coli*

Transformation into Epicurian Coli XL1-Blue Supercompetent Cells

The PCR products obtained by the Quick Change™ site-directed mutagenesis kit were,

according to the manufacturers instructions, digested with the *Dpn*I restriction enzyme in

order to get rid off the methylated, non-mutated parental DNA. The nicked, digested

plasmid DNA was then transformed into Epicurian Coli XL1-Blue Supercompetent cells

and plated on LB-ampicillin plates according to the manufacturers transformation protocol.

Transformation into *E. coli* BL21

Plasmid DNA was prepared from Epicurian Coli XL1 clones obtained by the QuickChange™ site directed mutagenesis procedure, and transformed into *E. coli* BL21 by $CaCl_2$ standard method. After propagation in *E. coli* BL21, plasmid DNA was isolated by the Wizard kit (Promega) and sequenced.

Transformation into *E. coli* MT102

MT102 was made competent by $CaCl_2$ standard method, and used for transformation of ligation mixtures. Transformation reactions were spread on LB plates containing 100 μg/ml ampicillin.

Transformation into *E. coli* KY895

KY895 was made competent by $CaCl_2$ standard method and transformed with pGEX-2T-TK1$^{66-136(Val106)}$ plasmid DNA from one of the clones coding for the correct putative interface sequence. The strain containing this plasmid was named KY895(pGEX-2T-TK1$^{66-136(Val106)}$).

DNA sequencing

pGEX-2T-TK1^{Val106}, pGEX-2T-TK1^{Met106A} and pGEX-2T-TK1$^{66-136(Val106)}$ plasmids were sequenced on both strands by the 2',3'-dideoxynucleoside 5'-triphosphates (ddNTP) chain-termination method according to the "Step-By-Step Protocols for DNA Sequencing with Sequenase Version 2.0 T7 DNA Polymerase" (United States Biochemical) with "Sequenase™ Version 2.0 DNA Sequencing Kit" (United States Biochemical), and with [α-^{35}S]-dATP (Amersham Pharmacia Biotech).

Plasmids pGEX-2T-TK1^{Val106}, pGEX-2T-TK1^{Met106B}, pGEX-2T-TK1^{Val106X} and pGEX-2T-TK1$^{66-136(Met106)}$ were sequenced on both strands by the ddNTP chain termination method with "Thermo Sequenase radiolabeled terminator cycle sequencing kit" (Amersham Biosciences) according to the instructions of the manufacturer. This kit uses four [α-^{33}P] dideoxynucleotide terminators (ddNTPs), so the radioactive label is incorporated at the 3'end of only properly terminated sequencing reaction products.

Expression of recombinant proteins

Expression of recombinant TK1 enzymes

BL21(pGEX-2T-TK1) bacterial cultures were grown overnight at 37°C, diluted to A_{600} = 0.6 into LB medium with 50 µg/ml ampicillin and adjusted to room temperature at the laboratory bench. Expression of the GST-TK1 fusion proteins was induced with 0.1 mM IPTG for 6 hours at 25°C and strong aeration. The bacterial cultures were centrifuged at 4°C and the pellet stored at -80°C.

Expression of recombinant TK1$^{66-136}$ proteins

KY895(pGEX-2T-TK1$^{66-136(Val106)}$) and BL21(pGEX-2T-TK1$^{66-136(Met106)}$) were grown overnight at 37°C in LB medium with 100µg/ml ampicillin. (The switch to BL21 as the host strain for the expression of the recombinant protein was due to the possibility of obtaining higher protein yield; the decision was taken after the experiments with TK1$^{66-136(Val106)}$ were finished.) The overnight bacterial cultures were diluted into fresh medium to A_{600} = 0.6 and left at the laboratory bench for temperature adjustment. They were then grown overnight (KY895 strains) or for approximately 6 hours (BL21 strains) at 25°C with strong aeration and in the presence of 0.1 mM IPTG (Sigma) to induce the GST fusion protein. The cells were harvested by centrifugation and stored at ÷ 80°C.

Purification of recombinant TK1 enzymes and TK1$^{66-136}$ proteins

Lysis of bacterial cells and preparation of crude bacterial extract

The bacterial pellet was resuspended in 1/20 of the original culture volume in lysis buffer (buffer A: 20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 1 mM DTT, 10% glycerol, 1% Triton X-100, 0.1 mM PMSF, 5 mM benzamidine and 50 mM ε-aminocaproic acid) and disrupted in French Press (SLM Aminco). The bacterial lysate was centrifuged at 4°C for 20 min and filtered through Whatman GF/C glass microfibre filter and MFS cellulose acetate 0.45µm filter in order to remove bacterial debris.

Glutathione Sepharose chromatography

Glutathione Sepharose 4B beads (Amersham Biosciences) were home-packed into 25 mm x 200 mm column and pre-equilibrated at 4° C with PBS, phosphate–buffered saline (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3) containing 1 mM DTT, 10% glycerol, 1% Triton X-100, 0.1 mM PMSF, 5 mM benzamidin and 50 mM ε-aminocaproic acid (so called PBS wash buffer, or buffer B). The bacterial filtrate was slowly applied to the column and the unbound proteins were removed by washing with 10-20 volumes of wash buffer. Thereafter, the column was equilibrated with phosphate-buffered saline with 0.1% Triton X-100, and moved to room temperature. The fusion protein was cleaved at room temperature for 1.5 – 2 hours by recirculation of 1 bed volume PBS containing 0.1% Triton X-100 and 50 units/ml thrombin (Amersham Biosciences). The eluate was continuously collected on ice and was named cleavage fraction 1 (cl.#1). After the cl.#1 was collected, the column was washed 4 times with one column volume of phosphate-buffered saline containing 0.1% Triton X-100. The fractions were collected on ice and named consecutively cl.#2 to cl.#5. After addition of 10% glycerol, 5 mM DTT, 5 mM $MgCl_2$ and 1% Triton X-100, the thrombin cleavage fractions were stored at -80°C. For kinetic and stability experiments where $TK1^{Val106}$, $TK1^{Met106}$ and $TK1\Delta40$ are compared, the proteins were desalted by G-25 chromatography and purified further by CM Sepharose chromatography.

Sephadex G-25 chromatography

G-25 Sephadex column (Amersham Biosciences) was equilibrated with buffer C: 10 mM K-phosphate buffer, pH 6.0, containing 5 mM $MgCl_2$, 2 mM DTT, 0.5 mM Chaps and 10% glycerol. Thrombin cleavage fractions were applied to the column and eluted with the same buffer. Protein concentration and conductivity of the eluate was continuously monitored.

CM-sepharose chromatography

CM-sepharose CL6B (Amersham Pharmacia) column was equilibrated with buffer D: 10 mM K-phosphate, pH 6.0, containing 5 mM $MgCl_2$, 2 mM DTT, 0.5 mM CHAPS and 10% glycerol. The desalted thrombin fractions were applied, and the column was washed with 50 column volumes of the equilibration buffer D in order to remove the unbound proteins.

The TK1 proteins were eluted with elution buffer E: 50 mM K-phosphate, pH 8.0, 5 mM MgCl$_2$, 2 mM DTT, 0.5 mM Chaps and 10% glycerol.

Determination of protein concentration

The method of Bradford (1976) with BSA dilutions of 20-40-60-80-100-120-140-160 and 180 µg/ml as standard was used, unless stated otherwise.

ATP incubation and storage of the purified TK1 enzymes

The thrombin cleavage fractions were diluted to 5 µg/ml in dilution buffer F (50 mM Tris-HCl, pH 7.5, 5 mM MgCl$_2$, 0.1 M KCl, 2 mM CHAPS, 10 % glycerol and 5 mM DTT) and the CM fractions were diluted to 5 µg/ml in dilution buffer G (50 mM Tris-HCl, pH 7.5, 1 mM CHAPS, 3 mg/ml BSA and 10% glycerol) with or without 2.5 mM ATP/ MgCl$_2$, and incubated on ice for 2 hours before storage at -80° C. The enzymes incubated and stored with and without ATP are referred to as the +ATP and –ATP forms, respectively.

Estimation of subunit molecular size and protein purity by SDS-polyacrylamide gel electrophoresis

Estimation of subunit size of recombinant TK1 enzymes

Protein samples were denatured for 2 minutes at 94° C in protein loading buffer: 30% glycerol, 0.4 M Tris-HCl, pH 6.8, 10 mM DTT, 1% SDS, and 0.01% bromphenol blue as tracking dye, and loaded onto the gel. The samples were run by standard discontinuous gel electrophoresis in Tris-HCl with 4.5% stacking gel, pH 6.8, 15% separation gel, pH 8.8, and with Tris-glycine running buffer The separated polypeptides were stained by silver staining procedure. Protein marker kit from BIO-RAD contained 2 mg/ml of each of the

marker protein (molecular mass from 14.4 to 94 kDa). It was diluted 100-fold, denatured, and 2 to 8 µl were applied to the gel.

Estimation of subunit size of recombinant TK1$^{66-136}$ proteins by tricine/ethylene glycol/SDS-polyacrylamide gel electrophoresis

Because the standard SDS-PAGE methods resulted in diffuse protein bands and unsufficient resolution of the relatively small TK1$^{66-136}$ peptide (< 8 kDa), a modified method was developed. The gel included 30% ethylene glycol in the gel slurry and was run with tricine instead of glycine in the electrode buffer (Separation Technique File No. 112 from Pharmacia, made available to me through Ulrik Cordes (personal communication); Schägger & von Jagow, 1987).

The upper, stacking gel contained 4.5% acrylamide /0.12% bisacrylamide (Protogel, National Diagnostics), the lower, separating gel contained 13.5% acrylamide /0.35% bisacrylamide. The buffer system in both gels consisted of 0.112 M acetate, 0.112 M Tris, pH 6.5 and both gels included 33% ethylene glycol. The electrode buffer contained 0.2 M tricine, 0.2 M Tris, pH 8.1-8.3, and 0.55% SDS. Protein samples were diluted approximately 10-fold into protein sample buffer (0.125 M Tris, pH 6.8, 25% glycerol, 1% SDS, 10 mM DTT and 0.01% bromphenol blue), denatured for 2-3 minutes at 95°C and loaded. The gel was pre-run for about 30 minutes at 80 V and run for about 45 minutes at 120 V. It was critical that the gel was not left in fixing and swelling solutions for longer than 30 minutes and 2 times 15 minutes respectively before silver staining procedure was performed. "Peptide marker kit MW 2.512 – 16949" (Amersham Pharmacia Biotech) was used as molecular mass marker. It was resuspended in protein sample buffer to 1 mg/ml (corresponding to about 0.17 µg/µl of each standard protein), stored at ÷ 20°C in small aliquots, diluted 10-fold before denaturation and 4 to 6 µl were applied to the gel.

Estimation of native molecular size by gel filtration chromatography

Estimation of native size of recombinant TK1 enzymes

The apparent native molecular size was determined on a Superose-12 or Superdex-200 columns (10 x 300 mm) connected to a Gradiphrac automatic sampler (Amersham

Pharmacia). The column was equilibrated and eluted with 50 mM imidazol-HCl buffer, pH 7.5, containing 5 mM $MgCl_2$, 0.1 M KCl, 5 mM DTT, and 2 mM Chaps (buffer H). The last two ingredients were added just before the start of the experiment. The columns were standardized with following proteins: β-amylase (200 kDa), bovine serum albumin (66 kDa), ovalbumin (45 kDa), carbonic anhydrase (29 kDa) and cytochrome C (12,4 kDa). The protein samples (200 µl, 5 µg/ml), were injected into the column together with Blue Dextran 2000 (4 µl from 10 mg/ml) as the internal standard. 200 µl aliquots were sampled on ice, into glasses containing 100 µl buffer H with 30% glycerol, 2 mM CHAPS and 2.6 mM ATP for preservation of enzyme activity. The final glycerol concentration was 10%, and the samples could be frozen for later estimation of enzyme activity. The fractions were assayed according to the standard assay for TK1 activity at 10 or at100 µM thymidine.

Estimation of native size of TK1$^{66-136}$ proteins

The apparent native molecular size was estimated on Superose 12 column (10 x 300 mm) connected to a Gradifrac automatic sampler (Amersham Pharmacia Biochem). The column was pre-equilibrated with elution buffer H without Chaps (DTT was added to the buffer just before the start of the experiment). The column was standardized as described above. 200 µl aliquots from the cleavage fractions were injected into the column together with Blue Dextran 2000 (4 µl from 10 mg/ml). 0.5 ml fractions were collected for estimation of protein concentration by the method of Bradford (1976).

Thymidine kinase assay

The standard assay conditions for thymidine kinase are: 50 mM Tris-HCl, pH 8.0, 2.5 mM $MgCl_2$, 2.5 mM ATP, 10 mM DTT, 0.5 mM CHAPS, 3 mg/ml BSA and the indicated concentration of [methyl-^{3}H]-thymidine in 50 µl total volume. The stock solution of radioactive thymidine (925 GBq/mmol, Amersham) was diluted with non-radioactive thymidine in order to get the required specific activity and concentration. The stored enzymes (at 5 µg/ml) were further diluted, just before the start of the reaction, into 50 mM Tris-HCl, pH 7.5, 1 mM CHAPS, and 3 mg/ml BSA, with or without 2.5 mM ATP/ $MgCl_2$. The TK1 assay activity was performed by the DE-81 filter paper method: aliquots

of the assay mix are put on DE filters at different reaction times, the filters are dried, and the non-phosphorylated thymidine is washed of the DE filter by three washes with ammonium formiate and one wash with water. The ^{3}H-TMP bound to the filter is eluted from the paper by a mixture of 0.2 M KCl and 0.1 M HCl, and counted in a liquid scintillation counter as previously described (Munch-Petersen et al., 1991).

Enzyme kinetics

The experimental data were fitted to Hill equation, which is obtained from Michaelis-Menten equation by raising the substrate concentration and the K_m value in to the power of n, the Hill constant:

$$v = \frac{V \cdot s^n}{K_{0.5}^n + s^n}$$

The kinetic parameters and reaction mechanism were determined using the non-linear regression software from Graphpad Prism®. V is the maximal velocity and $K_{0.5}$ like K_m in the Michaelis-Menten equation, defines the substrate concentration s where $v=0.5$ V_{max} (Cornish-Bowden, 1995).

Detection of GST-TK1$^{66-136}$ fusion proteins

Because the cloned TK1$^{66-136}$ had no enzymatic activity, the IPTG induction of the fusion protein was detected by measuring the product of enzymatic reaction between GST and its substrates CDNB (1-chloro-2,4-dinitrobenzene) and glutathione. 5 to 50 µl of crude bacterial extracts were added to 100 mM phosphate buffer, pH 6.5, mixed with 1 mM reduced glutathione and 1 mM CDNB, and the absorbance A$_{340}$ recorded at 1 min intervals for at least 5 min. Usually, the GST fusion protein was detected in small volumes of bacterial culture, e.g. 2 ml, taken before addition of IPTG and just before harvest of the cells. Such small volumes could not be subjected to French Press nor the filtration procedure, but were sonicated (50 W, 3 X 10 sec) instead and centrifuged to remove bacterial debris.

Enzyme stability

The stored CM fractions of $TK1^{Val106}$ and $TK1^{Met106}$ enzymes (at 5 µg/ml) were diluted 200-fold into 50 mM K-phosphate buffer, pH 7.5 with 0.5 mM CHAPS and 5 mM DTT, and incubated at 15° C. 10 µl aliquots were withdrawn at different incubation times, added to 40 µl standard assay mix with 100 µM thymidine at 37° C, and at 3, 6 and 9 minutes afterwards, 13 µl aliquots were spotted on the DE-filter. Filters were washed and counted as described before.

Lymphocytes, cell lines and cell culture conditions

Lymphocytes were isolated by the Isopaque-Ficoll technique from peripheral blood from healthy volunteers at Roskilde University, Roskilde, Denmark. If destined for isolation of mRNA, the lymphocytes were grown for 72 hours with PHA as described by Munch-Petersen et al. (1973). Following cell lines were bought from American Type Culture Collection: Raji, CEM-C, Molt-3, Reh, K-562, KG-1, RS4;11, AML-193 (the leukemic cell lines), DLD-1, Het-116, LoVo, SW40 (the colon cancer cell lines), and the HeLa cells. WI38, the human fibroblast cell line, and WI26 VA4, the SV40 transformed human fibroblast cell line, were bought from European Collection of Cell Cultures. The cell lines were cultured in RPMI 1640 medium (Life Technologies, Inc.) containing L-glutamine, 15% fetal bovine serum (Life Technologies, Inc.), and 1% penicillin-streptomycin in atmospheric air with 5% CO_2 at 37° C. Subculturing was performed at 10^6 cells/ml. All handling of the cell lines was performed by Jacob F. Christensen (Christensen, 2000)

Isolation of mRNA and amplification of TK1 cDNA

Total RNA was isolated with RNAqueous™ isolation kit (Ambion, Inc.) from eight leukemic cell lines, non-transformed and SV40 transformed fibroblasts, and lymphocytes from five healthy donors. The cDNA was amplified with Titan™ one-tube reverse

transcription-PCR kit (Roche Molecular Biochemicals) using following primers: sense primer: 5'-GAGAGTACTCGGGTTCGTGAA(22-42)-3', and antisense primer: 5'-ATGCAGGGCAGCGTCCAGTAG(825-805)-3' (the numbers are according to Bradshaw & Deininger, 1984). This resulted in 804 base pair fragment containing the entire coding region of TK1. The isolation of mRNA and amplification of cDNA was performed by Jacob F. Christensen (Christensen, 2000).

Isolation of a genomic TK1 DNA fragment containing the codon for amino acid 106 and amplification of this fragment by PCR

Total genomic DNA was isolated from four colon cancer cell lines, HeLa cells, and lymphocytes from four healthy donors by the SDS lysis-proteinase K treatment-phenol extraction method (Sambrook et al., 1989). Fragments of 166 bp containing the codon for amino acid 106 were amplified by PCR with the sense primer 5'-AGCGTCTTCGCTGGGGCTCC(11824-11843)-3' and the antisense primer 5'-TTCCTCTGGAAGGTCCCATCC(11989-11969)-3' at the following conditions: 25-30 cycles, 94° C for 1 min, 55° C for 2 min, and 72° C for 3 min. The numbering is according to Flemington et al. (1987). Isolation and PCR amplification of genomic TK1 DNA was performed by Jacob F. Christensen (Christensen, 2000).

Sequencing of the genomic TK1 DNA fragments and of TK1 cDNA

Both strands were sequenced with the Thermo Sequenase radiolabeled terminator cycle sequencing kit (Amersham Biosciences) with the dITP nucleotide master mix and [α-^{32}P] dideoxynucleotides (Amersham Biosciences).

NMR experiments

When the induced $TK1^{66-136}$ proteins were destined for NMR, the bacterial pellets were lysed in buffer A1 of following composition: 20 mM phosphate buffer, pH 7.5, 5 mM EDTA, 1 mM DTT, 0.1 mM PMSF, 5 mM benzamidine and 50 mM ε-aminocaproic acid (it was especially important that no glycerol was added and to avoid traces of Tris-HCl). The proteins were purified by glutathione affinity chromatography as described earlier, but the buffers contained neither glycerol, nor Triton. Thrombin cleavage was performed without Triton, in PBS only, and the cleavage time was extended to 2 hours at room temperature. After thrombin cleavage 5 mM DTT and 5 mM $MgCl_2$ were added to the cleavage fraction of $TK1^{66-136(Val106)}$, while during later purification of $TK1^{66-136(Met106)}$, the concentration of DTT was reduced to 2 mM. The proteins were concentrated by centrifugation at 4°C through Centrisart membranes (Sartorius) with molecular weight cut-off size of 5000. The centrifugation time required to concentrate the proteins in the cleavage fractions approximately 10-fold (from protein concentration 200 - 500 µg/ml to 4 mg/ml) was 3 days. (About 10% of the protein applied to the Centrisart passed through the membrane into filtrate). Afterwards, 0.5 M D_2-MSH (Aldrich) was added to final concentration of 2 mM, and D_2O (Aldrich) was added to final concentration of 10%. The samples were stored in 5 mm Wilmad NMR tubes (Fluorochem, England) at 4°C and were run during one week after they have been prepared.

All experiments have used proton NMR (^{1}H-NMR) spectroscopy and 600 MHz Varian Unity-Inova NMR spectrometer with a triple resonance gradient probe. All spectra were run at 10°C with TMS (tetramethylsilane) as reference.

Chapter 6

The role of amino acid 106 in the catalytic mechanism and oligomerization pattern of human thymidine kinase 1

Introduction

The first successful cloning of a functional human thymidine kinase 1 gene was performed by Bradshaw (1983) who isolated genomic TK1 DNA by transformation of TK-deficient mouse L cells with total HeLa cell DNA. A fragment of this genomic clone was used by Bradshaw & Deininger (1984) as a hybridisation probe for screening the Okayama-Berg cDNA library constructed with mRNA from SV40 transformed human fibroblasts (Okayama & Berg, 1983). A hybridizing clone, called pTK11, with cDNA insert of ca. 1.5 kb and capable of transforming LTK cells to TK$^+$, was sequenced and found to have an open reading frame of 702 bp corresponding to 234 amino acids (Bradshaw & Deininger, 1984). The whole genomic clone of Bradshaw (1983) was sequenced by Flemington *et al.* (1987) who found the human TK gene to be composed of 7 exons. No base differences were detected between the amino acid coding sequence from HeLa cells (Flemington *et al.*, 1987) and the sequence of the pTK11 clone of Bradshaw & Deininger (1984).

The first TK1 clone that arrived at our laboratory was the pTK11 clone of Bradshaw & Deininger (1984). The amino acid coding region of pTK11 was PCR amplified, inserted into a pET3a vector, and the pET3a-TK1 plasmid used to transform *E. coli* BL21(DE3) strain (Jensen, 1994). This expression system was T7 RNA polymerase dependent and IPTG inducible (Jensen, 1994; Munch-Petersen *et al.*, 1995a). The recombinant protein, later named rTFi-TK1[1], was purified, characterized, and found to have the same subunit size and specific activity as the native TK1 isolated from human lymphocytes, but to differ from it with respect to regulatory effect of ATP pre-exposure on the kinetic mechanism

---

[1] Shortage for recombinant TK1 expressed from cDNA derived from SV40 Transformed human Fibroblasts

and oligomerization properties (Jensen, 1994; Jensen & Munch-Petersen, 1995; Munch-Petersen et al., 1995a). Thus, pre-assay exposure to ATP induced a reversible enzyme concentration-dependent transition of the native lymphocyte TK1 from a low activity dimer (molecular size about 50 kDa, thymidine $K_{0.5}$ about 15 μM) to a high activity tetramer (molecular size about 100 kDa, thymidine $K_{0.5}$ about 0.7 μM). Both forms had the same k_{cat} value of about 4 s^{-1} when calculated per subunit of 24 kDa (Munch-Petersen, 1996). However, irrespective of pre-exposure to ATP, the rTFi-TK1 behaved as the low $K_{0.5}$ tetramer, with K_m about 0.4 μM (Munch-Petersen et al., 1995a). At that time, these differences were attributed to absence of post-translational modification mechanism in the bacterial expression system (Munch-Petersen et al., 1995a).

Because the yield of this expression system was poor, approximately 1 mg of TK1 protein per litre of bacterial culture, (Jensen, 1994), it was decided to use the pGEX-2T expression system (Pharmacia Biotech) instead. As described in Chapter 5 (Materials and Methods) and Paper I (Berenstein et al., 2000), the amino acid coding sequence of TK1 inserted into the multicloning site of pGEX-2T was not from HeLa cells, but from normal human lymphocytes, and accordingly the expressed TK1 was named rLy-TK1[2] (Paper I, Berenstein et al., 2000).

Preliminary kinetic and gel filtration experiments have shown that rLy-TK1 behaved as the native TK1 isolated from human lymphocytes with respect to the regulatory effect of ATP on the kinetic and oligomerization properties of the enzyme (Kristensen, 1996). Therefore, absence of post-translational modification in the recombinant enzyme could not explain the divergent properties of rTFi-TK1 and rLy-TK1. By comparison of the sequence of lymphocyte TK1 cDNA with that of pTK11 clone of Bradshaw & Deininger (1984), two differences were discovered: 1) the lymphocyte TK1 cDNA had a GTG codon for valine at amino acid position 106, while pTK11 cDNA had an ATG codon for methionine, and 2) the lymphocyte TK1 cDNA had an AGG codon for arginine at position 211, while pTK11 had an AAG codon for lysine (Kristensen, 1996; Paper I, Berenstein et al., 2000). Because the amino acid 211 is positioned in the non-conserved C-terminal of thymidine kinases, and furthermore, the replacement of lysine with arginine is conservative, this change was not considered to be of importance for the properties of the enzyme. The amino acid at site 106, on the contrary, is located in a highly conserved area (Folkers et al., 1991; Paper I, Berenstein et al., 2000; se Fig. 4.1), and is valine in all published TK1 sequences of different mammalian enzymes, in chicken and in vaccinia TK1, but not in the published

human TK1 sequences that have methionine at position 106 (Bradshaw & Deininger, 1984; Flemington *et al.*, 1987).

This study was undertaken in order to clarify which amino acid is naturally occurring at the position 106, and describe the impact on the size, conformation and polarity of the amino acid 106 on the structure and function of human TK1. The results presented in this chapter are already published (Paper I, Berenstein *et al.*, 2000; Paper II, Frederiksen *et al.*, 2004).

Results

Expression and purification of recombinant TK1 enzymes

When recombinant proteins are synthesized by *E. coli*, they often accumulate in insoluble inclusion bodies (Marston, 1986). Use of gene fusions that link the eukaryotic gene to *e.g.* bacterial β-galactosidase gene (Marston, 1986) or to GST gene from *S. japonicum* (Smith & Johnson, 1988) increases solubilization of heterologous proteins in *E. coli*. Formation of the inclusion bodies could be prevented even further by cultivating *E. coli* at 25°C instead of 37°C (Fetzer & Folkers, 1992). Accordingly, the IPTG induction of the recombinant TK1 proteins was carried out at 25°C.

The recombinant enzymes were expressed as GST fusion proteins and purified by glutathione sepharose affinity chromatography combined with *in situ* thrombin cleavage. The purity of the thrombin cleavage fractions was estimated to be >90% by SDS-PAGE (Fig. 6.1).
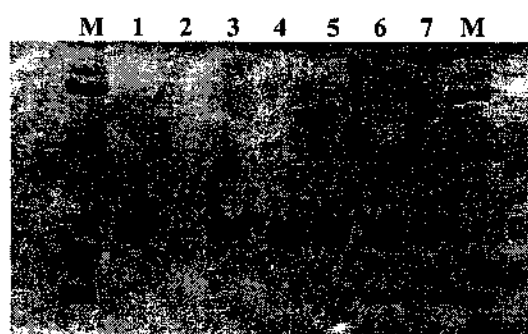
Figure 6.1. SDS gel electrophoresis of thrombin cleavage fractions. M lanes – marker proteins with 80 ng protein in each band, from bottom to top (kDa): α-lactalbumine (14.4), soybean trypsin inhibitor (20.1), carbonic anhydrase (30), ovalbumine (43), bovine serum albumine (67) and phosphorylase B (94). Lanes 1 to 7 contain 0.15 – 0.2 µg protein of respectively: cl.#2 V106WT, no sample, cl.#1 V106WT, cl.#2 V106T, cl.#1 V106T, cl.#2 V106Q and cl.#1 V106Q.

---

[2] Shortage for recombinant TK1 expressed from cDNA derived from normal human lymphocytes

For kinetic and stability experiments where TK1(V106WT) and TK1(V106M) are compared, the proteins were purified further by CM Sepharose chromatography. This procedure results in pure (99%) recombinant protein (Fig. 6.2).
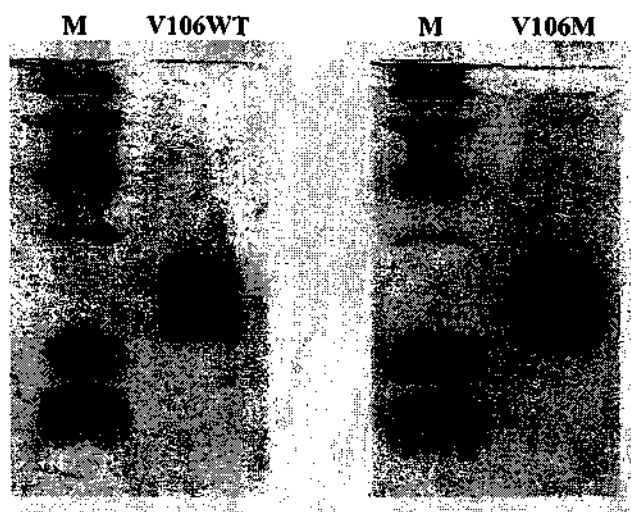
Figure 6.2. SDS gel electrophoresis of TK1(V106WT) and TK1(V106M) CM-sepharose fractions. Approximately 0.55 μg of each protein was applied to the gel. M lanes - marker proteins, from bottom to top (kDa): α-lactalbumine (14.4), soybean trypsin inhibitor (20.1), carbonic anhydrase (30), ovalbumine (43), bovine serum albumine (67) and phosphorylase B (94).

The yield of recombinant TK1 and the purification achieved from 300 ml bacterial culture are shown in Table 6.1 with V106M as an example.

Table 6.1. Purification of recombinant TK1(V106M)

	Volume	Activity		Protein		Specific activity	Yield
	ml	Units/ ml	Units total	μg/ml	mg total	Units/mg	%
Crude extract							
	50	917	45850	8615	431	106	100
GSH Sepharose (top fraction)							
	13	579	7527	179	2.3	3272	16
CM Sepharose (top fraction)							
	1.7	1589	2701	429	0.729	3710	5.9

All data are related to 1 litre of bacterial culture
Units: nmol min^{-1}

The yield of the recombinant protein was much higher if the purification procedure started with bigger volume of bacterial culture (>1 litre). The amount of protein related to 1 litre of bacterial culture was 4 – 8 mg in the thrombin cleavage fractions, and 2 – 4 mg (at concentration 4 – 8 mg/ml) in the CM top fraction.

The significance of amino acid 106 was investigated by comparing the kinetic and oligomerization properties of the rLy-TK1^{Val106} (V106WT) with rLy-TK1^{Met106} (V106M) and rLy-TK1^{Val106X} (V106X) mutants.

Enzyme kinetics

The thymidine substrate kinetics of TK1(V106WT) is compared to the thymidine substrate kinetics of TK1(V106M) at saturating concentrations of ATP in Fig. 6.3.
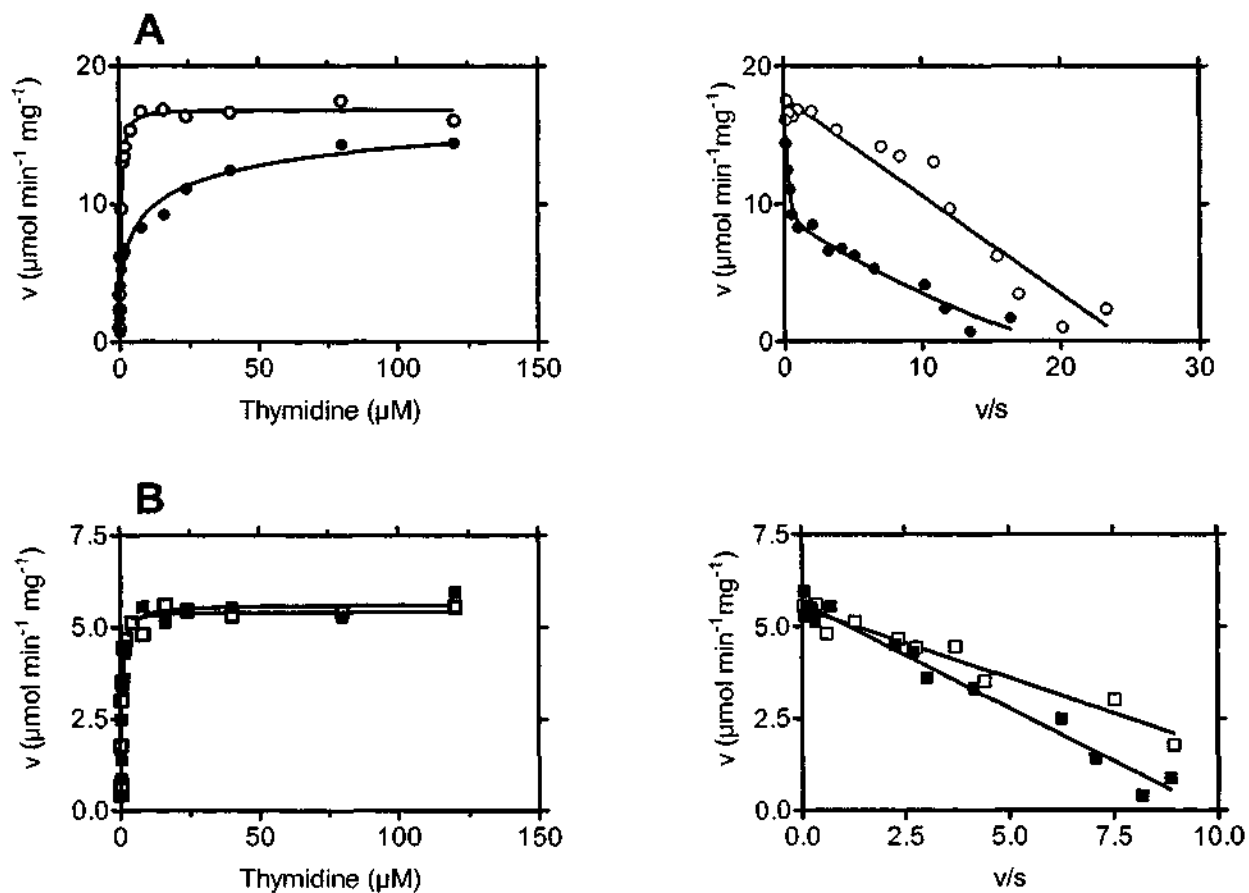
Figure 6.3. Thymidine substrate kinetics. A, TK1(V106WT), B, TK1(106M). The initial velocity of dTMP formation as a function of thymidine concentration is shown to the left, Hofstee plot of the same data is shown to the right. Open symbols, +ATP forms, closed symbols –ATP forms. The experiments were performed on CM-sepharose fractions at 5 ng/ml enzyme protein in the reaction mixture.

The recombinant TK1(V106WT) behaves essentially as the previously described endogenous TK1 isolated from human lymphocytes (Munch-Petersen *et al.*, 1993): the – ATP form of the enzyme displays a nonhyperbolic, "creeping" curve and a biphasic kinetic pattern in the Hofstee plot, indicating a negative cooperative reaction mechanism (n value 0.4), while the +ATP form displays a rectangular hyperbolic kinetics (n value somewhat above 1), and a straight line in the Hofstee plot (Fig. 6.3A). Exchange of Val106 to methionine (Fig. 6.3B) resulted in a hyperbolic rectangular reaction mechanism with a Hill coefficient (n) about 1, and a straight line in Hofstee plot, independent of preceding exposure to ATP. The kinetic properties of TK1(V106M) turned out to be the same as those described previously by Munch-Petersen *et al.* (1995a) for rTFi-TK1 derived from the pTK11 cDNA of Bradshaw & Deininger (1984).

The kinetic values of TK1(V106WT) and TK1(V106M) are compared in Table 6.2.

Table 6.2. Kinetic values of TK1(V106WT) and TK1(V106M).

	TK1(V106WT)		TK1(V106M)	
	-ATP	+ATP	-ATP	+ATP
V_{max} (µmol/min/mg)	20	17	5.6	5.4
K_{cat}[a] (sec^{-1})	8.5	7.2	2.4	2.3
$K_{0.5}$ (µM)	15	0.57	0.55	0.4
$K_{cat}/K_{0.5}$ (sec^{-1}µM^{-1})	0.57	12.6	4.4	5.7
n	0.4	1.3	1	1.1

[a]K_{cat} value was obtained using the calculated MW of TK1 equal to 25500

The $K_{0.5}$ value for thymidine of TK1(V106WT) –ATP form is about 30-fold higher than the $K_{0.5}$ value of its +ATP form. In contrast, the $K_{0.5}$ values of the –ATP and +ATP form of TK1(V106M) are equally low, about 0.5, the same as the value of TK1(V106WT) +ATP form. Though, due to lower V_{max} of TK1(V106M), its turnover number K_{cat} is 3 to 4 fold lower than the K_{cat} of TK1(V106WT), and its kinetic efficiency measured as $K_{cat}/K_{0.5}$ is also lower than the $K_{cat}/K_{0.5}$ of TK1(V106WT)+ATP form. The physiological importance of transition between a low and high activity form of TK1 will be discussed later.

The effect of replacement of Val106 with several other amino acids on the kinetic properties of the enzyme is shown in Table 6.3 and Fig. 6.4.

The thymidine substrate kinetics of V106A, V106I and V106T is essentially the same as that previously observed for V106WT: the –ATP form of these enzymes displayed

while their corresponding +ATP forms had low $K_{0.5}$ values and n values slightly above 1 (Fig. 6.4A). The substrate kinetics of V106G, V106H, V106K, V106L and V106Q was essentially the same as that of V106M: both the -ATP and +ATP forms seem to have a hyperbolic reaction mechanism with low $K_{0.5}$ values and n values above 1, that may indicate positive cooperativity (Fig. 6.4B and Table 6.3). The ratios between the $K_{0.5}$ values for the -ATP and +ATP forms supports the division of mutant enzymes into 2 groups. This ratio is 30 or higher for the enzymes belonging to Group I: V106WT, V106A, V106I and V106T, whereas it is about 1 for the enzymes belonging to Group II: V106G, V106H, V106K, V106L, V106M and V106Q (Table 6.3).

Table 6.3. Kinetic parameters of TK1(V106WT) and TK1(V106X) mutant enzymes

Enzyme	V_{max}		$K_{0.5}$		n		$K_{0.5}(-ATP)/$
	μmol min^{-1} mg^{-1}		μM				$K_{0.5}(+ATP)$
	-ATP	+ATP	-ATP	+ATP	-ATP	+ATP	
Group I V106WT-like enzymes							
V106WT	11.0±1.3	9.4±0.2	27.1±8.5	0.6±0.05	0.8±0.07	1.2±0.05	45
V106A	8.0±0.8	6.6±0.2	12.7±3.2	0.3±0.03	0.8±0.06	1.3±0.1	42
V106I	10.3±1.8	8.3±0.3	29.4±16.4	0.9±0.1	0.8±0.1	1.2±0.1	33
V106T	8.4±1.9	7.1±0.2	43±28	0.4±0.04	0.8±0.1	1.2±0.1	108
Group II V106M-like enzymes							
V106G	3.6±0.2	3.0±0.1	0.4±0.1	0.3±0.1	1.9±0.4	1.7±0.4	1.3
V106H	3.7±0.2	3.1±0.1	0.3±0.2	0.4±0.1	1.6±0.7	2.0±0.4	0.8
V106K	2.2±0.06	4.1±0.1	0.7±0.1	1.2±0.1	1.4±0.1	1.4±0.1	0.6
V106L	4.1±0.1	3.9±0.1	0.4±0.1	0.7±0.1	1.9±0.2	1.8±0.1	0.6
V106M	2.9±0.06	2.9±0.06	0.6±0.08	0.6±0.1	1.5±0.1	1.5±0.1	1.0
V106Q	4.9±0.2	4.2±0.1	0.8±0.1	0.8±0.1	1.5±0.1	1.6±0.1	1.0

The enzymes within each group are arranged in alphabetical order of their one-letter code. The best fit ± SE to all data is given.

A

B
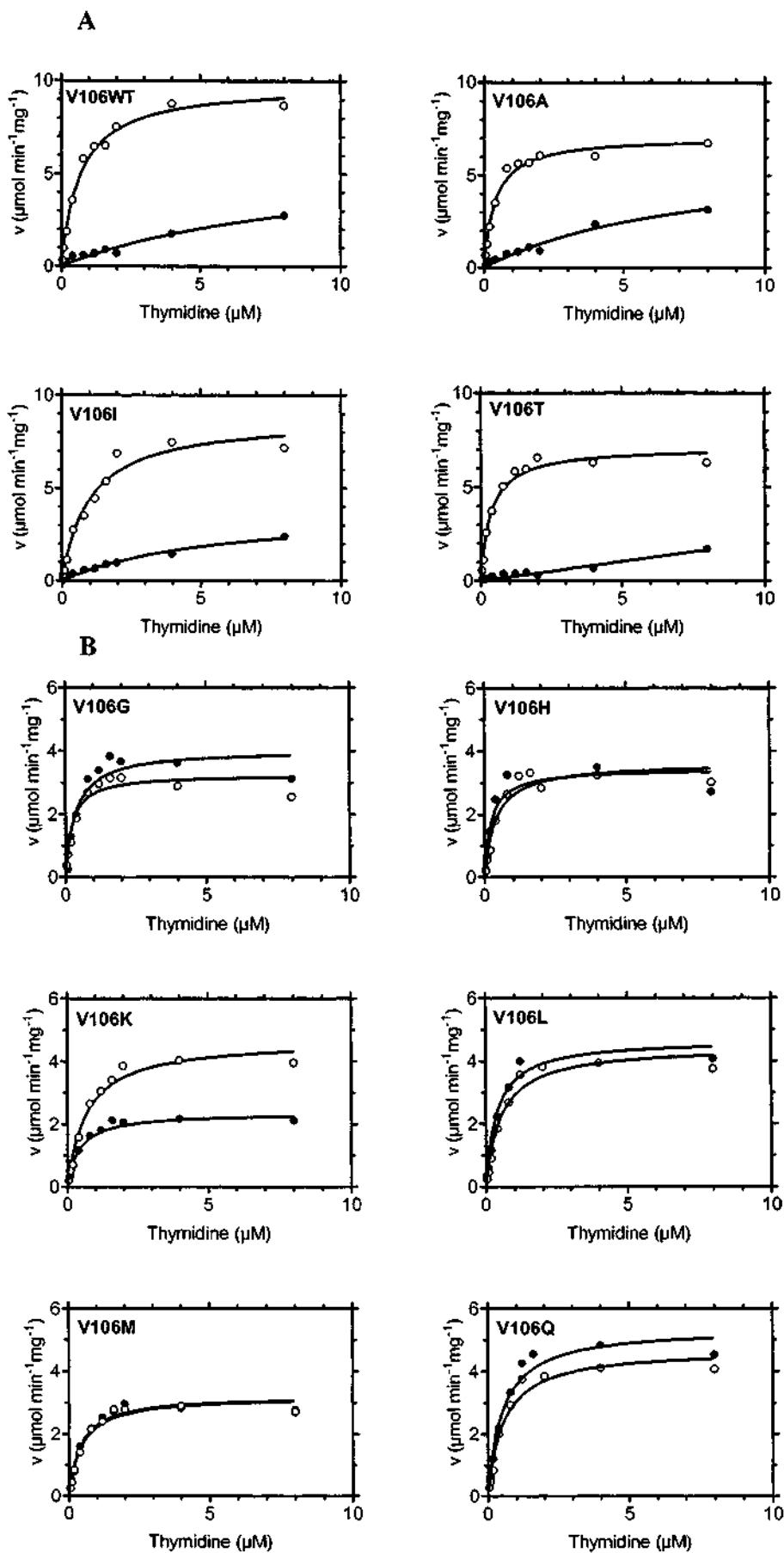
Figure 6.4. The initial velocity of dTMP formation as a function of thymidine concentration for TK1 (V106WT) and TK1 (V106X) enzymes. Open symbols, +ATP forms; closed symbols –ATP forms. A, V106WT-like enzymes and B, V106M-like enzymes. The experiments were performed on thrombin cleavage fractions at 5 ng/ml enzyme protein in the reaction mixture.

Native molecular size

Initially, the effect of replacement of Val106 with methionine on the native molecular size under conditions of gel filtration experiments was investigated (Fig. 6.5)
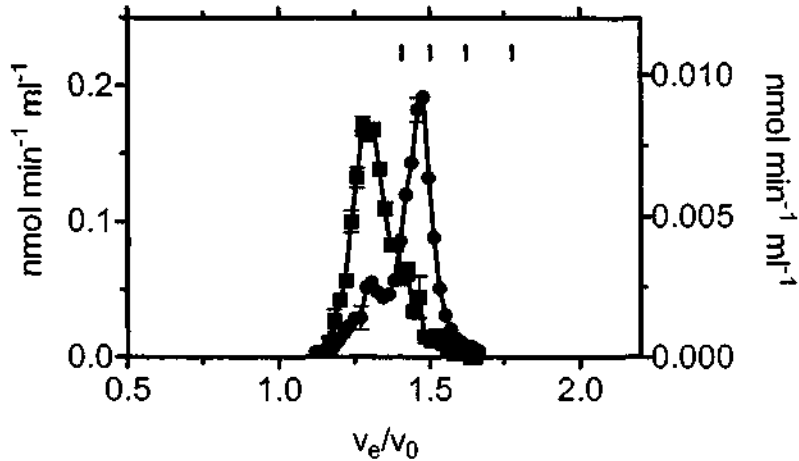
Figure 6.5. Gel filtration of TK1(V106WT) and TK1(V106M). 0.6 µg protein from thrombin cleavage fractions was applied on Superose column. Left hand scale, V106WT (circles), right hand scale, V106M (squares). The enzymatic activity was estimated at 10 µM thymidine. The molecular mass markers (|) are (in kDa): BSA (66), ovalbumine(45), carbonic anhydrase (29) and cytochrome c (12.4).

The apparent sizes of the –ATP form of V106WT and V106M was respectively 50 and 100 kDa, and according to the subunit size of 24-25 kDa, the V106WT eluted predominantly as a dimer, while V106M eluted predominantly as a tetramer. In the presence of ATP V106WT enzyme eluted as a tetramer (Fig. 6.6).
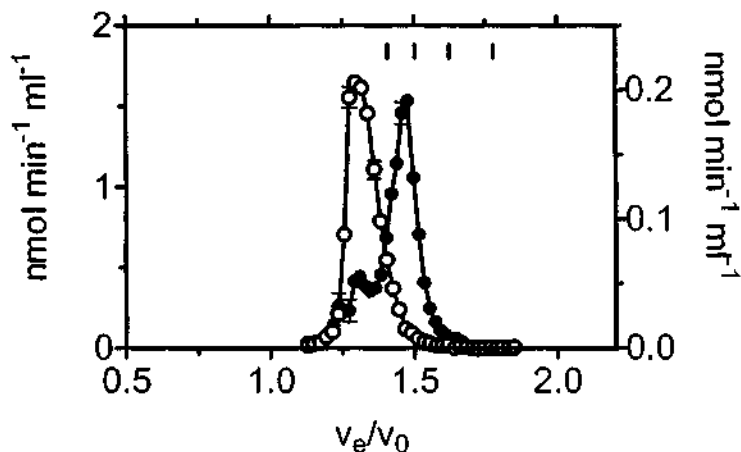
Figure 6.6. Gel filtration of TK1(V106WT) +ATP and -ATP forms. 1 µg and 0.6 µg from respectively +ATP and ATP form protein from thrombin cleavage fractions were applied on Superose column. Left hand scale, +ATP form (open symbols); right hand scale, -ATP form (closed symbols). The enzymatic activity was estimated at 10 µM thymidine. The molecular mass markers (|) are (in kDa): BSA (66), ovalbumine(45), carbonic anhydrase (29) and cytochrome c (12.4).

The behaviour of the recombinant enzymes was in agreement with the previous observation for the native Ly-TK1 (Munch-Petersen *et al.*, 1993) and the rTFi-TK1 (Munch-Petersen *et al.*, 1995a) derived from the cDNA clone of Bradshaw & Deininger (1984).

Because the activity of the dimeric form of TK1-ATP does not attain V_{max} at 10 μM thymidine, we determined the proportion of enzyme present as dimer/tetramer by estimation of the enzymatic activity in gel filtration fractions of different V106X mutants at the almost saturating thymidine concentration of 100 μM. The results are shown in Fig.6.7.

The enzyme concentration applied to the column was 3.25 μg/ml, because we wanted to mimic the physiological concentration of TK1 protein calculated to be about 4 μg/ml in S phase cells (Munch-Petersen *et al.*, 1993; Munch-Petersen *et al.*, 1995). The recombinant mutant enzymes could be divided into two groups according to the oligomerization pattern of their –ATP form. V106A, V106I and V106T eluted from Sephadex column essentially as V106WT, and were called "dimeric enzymes", because a substantial part of each enzyme eluted with an approximate size of 50 kDa, i.e. as a dimer (Fig. 6.7A). In contrast, V106G, V106H, V106K, V106L and V106Q eluted similarly to V106M with an approximate size of 100 kDa, i.e. as tetramers, and therefore called "tetrameric enzymes". In each case, small, varying proportion of enzyme protein eluted as a dimer (Fig. 6.7B).
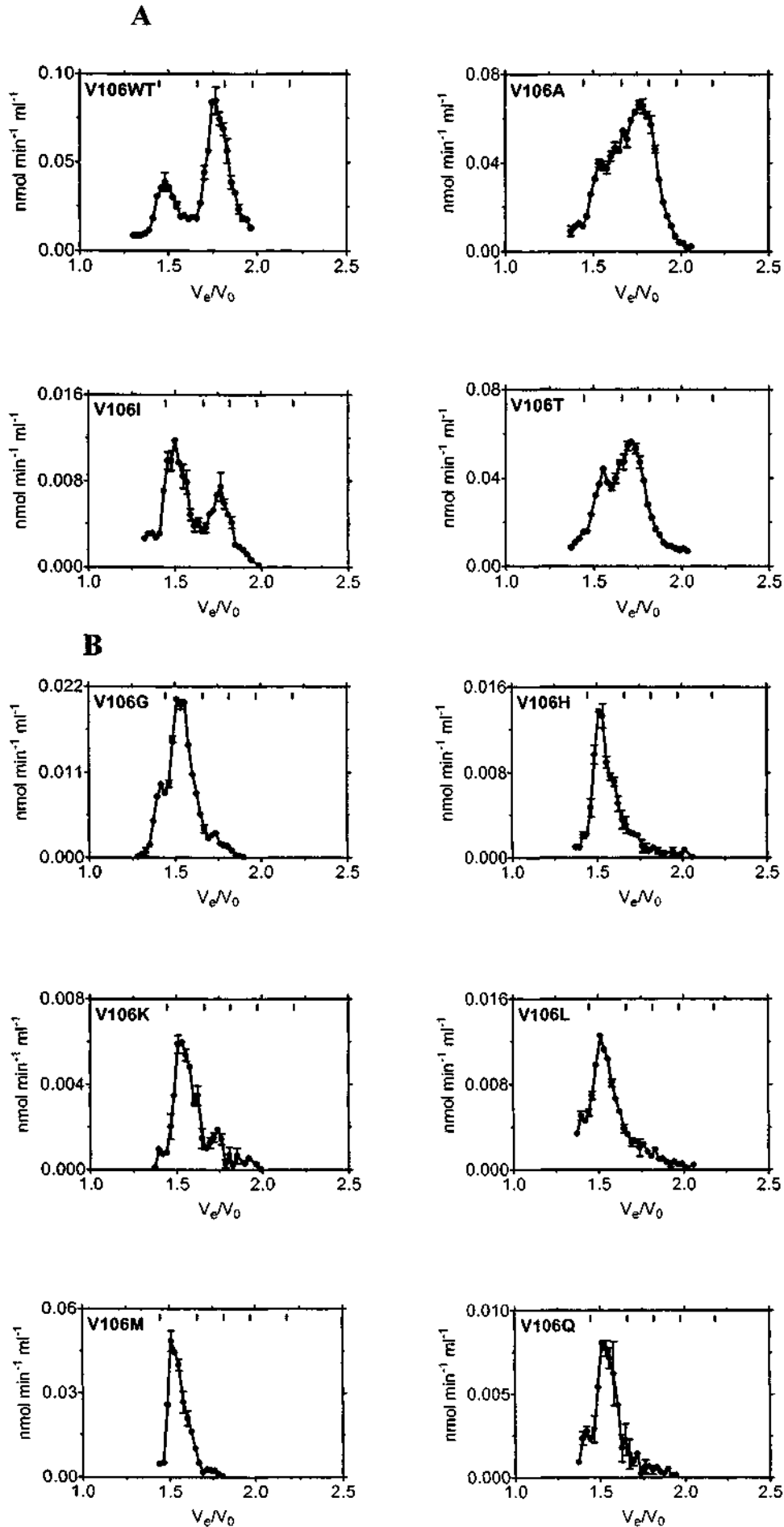
Figure. 6.7. Gel filtration of TK1(V106X) mutants on Superdex column. **A**, dimeric enzymes, **B**, tetrameric enzymes. Approximately 0.65 μg protein from thrombin cleavage fractions was injected. The molecular mass markers (|) are (in kDa) from left to right: β-amylase (200), BSA (66), ovalbumine (45), carbonic anhydrase (29) and cytochrome c (12.4).

Enzyme stability

In Fig. 6.8 the stability of V106WT with V106M at 15°C is compared. Since V106M lost more than 50% of its activity within 4 minutes at 25°C, stability experiments were not performed at this temperature. Also at 15°C, V106M was much less stable than V106WT, with $t_{1/2}$ values of about 40 and 390 minutes respectively. Analogously, the dimeric enzymes were found to be much more stable than the tetrameric enzymes when the respective –ATP form was diluted to 25 ng/ml into enzyme dilution buffer and incubated at 4°C for 30 minutes: the dimeric enzymes retained 80-100% of their enzymatic activity, while the tetrameric enzymes retained only 0.5-3%.
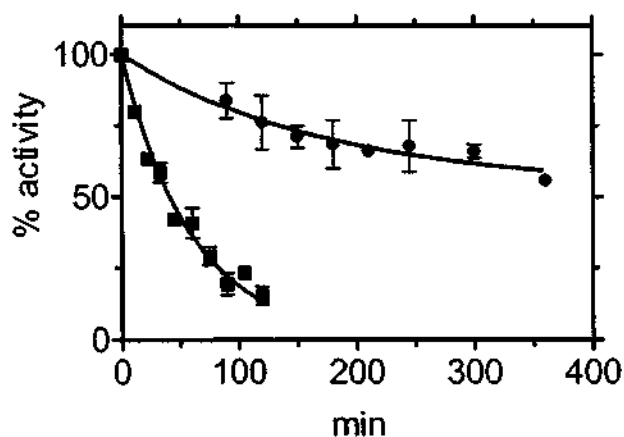
Figure 6. 8. Enzyme stability of TK1(V106WT), circles, and of TK1(V106M), squares, at 15° C in phosphate buffer. The bars show standard deviation derived from two or three independent determinations. CM-sepharose fractions were used for these experiments.

Polymorphisms in the human TK1 gene

In order to uncover naturally occurring bases and the resulting amino acids at positions 106 and 211 in the human TK1 gene, sequencing reactions were performed on:

(a) genomic DNA fragments of 166 base pairs containing the codon for amino acid 106 from four colon cancer cell lines, HeLa cells, and lymphocytes from four healthy donors,

(b) the complete coding region of TK1 cDNAs from eight leukemic cell lines, lymphocytes from five healthy donors, and one nontransformed and one SV40 transformed fibroblast cell line. Altogether, the codon for amino acid 106 was examined in genomic and cDNA from seven healthy donors and 15 cell lines. The results are summarized in Table 6.4.

Table 6.4   Polymorphisms in DNA and amino acid sequence of TK1 cDNAs

Source of cDNA	Nucleotide position						
	90	115	279	282	373	651	689
pTK11	CCC	CAG	GAG	GCG	ATG	GCC	AAG
	Pro	Gln	Glu	Ala	Met	Ala	Lys
Lymphocytes from healthy donors							
Donor 1	CCC	CAG	GAG	GCG	**GTG**	GCC	**AGG**
	Pro	Gln	**Gln**	Ala	**Val**	Ala	**Arg**
Donor 2	CC(C/T)	CAG	GA(G/A)	GC(G/A)	**GTG**	GCC	AAG
	Pro/Pro	Gln	Glu/Glu	Ala/Ala	**Val**	Ala	Lys
Donor 3	CC(C/T)	CAG	GA(G/A)	GCG	**GTG**	GCC	AAG
	Pro/Pro	Gln	Glu/Glu	Ala	**Val**	Ala	Lys
Donor 4	CC(C/T)	CAG	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro/Pro	Gln	Glu	Ala/Ala	**Val**	Ala	Lys
Donor 5	CC(C/T)	CAG	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro/Pro	Gln	Glu	Ala/Ala	**Val**	Ala	Lys
Leukemic cell lines							
RS4;11	**CCT**	CAG	GAG	**GCA**	**GTG**	GCC	AAG
	Pro	Gln	Glu	Ala	**Val**	Ala	Lys
Molt-3	CC(C/T)	CAG	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro/Pro	Gln	Glu	Ala/Ala	**Val**	Ala	Lys
Reh	CC(C/T)	**(C/T)AG**	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro/Pro	**Gln/Stop**	Glu	Ala/Ala	**Val**	Ala	Lys
KG-1	CCC	CAG	GAG	GCG	**GTG**	GCC	AAG
	Pro	Gln	Glu	Ala	**Val**	Ala	Lys
AML-193	CC(C/T)	CAG	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro/Pro	Gln	Glu	Ala/Ala	**Val**	Ala	Lys
K562	**CCT**	CAG	GAG	**GCA**	**GTG**	GCC	AAG
	Pro	Gln	Glu	Ala	**Val**	Ala	Lys
Raji	CC(C/T)	CAG	GAG	GCG	**GTG**	**GC(C/T)**	AAG
	Pro/Pro	Gln	Glu	Ala	**Val**	**Ala/Ala**	Lys
CEM-C	**CCT**	CAG	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro	Gln	Glu	Ala/Ala	**Val**	Ala	Lys
Fibroblasts							
WI38	**CCT**	CAG	GAG	GC(G/A)	**GTG**	GCC	AAG
	Pro	Gln	Glu	Ala/Ala	**Val**	Ala	Lys
WI26	CCC	CAG	GAG	GCG	**GTG**	GCC	AAG
	Pro	Gln	Glu	Ala	**Val**	Ala	Lys

The numbers are according to Bradshaw & Deininger (1984) and correspond to the underlined bases, where the polymorphism is located. Bases/amino acids that differ from the sequence of Bradshaw & Deininger (1984) are in bold. Differences in only one allele are in parentheses.

In each case the GTG codon for valine 106 was found on both alleles, and in one case – from the very first donor, called donor 1, from which the pGEX-2T-TK1 plasmid was derived – a substitution of A with G at nucleotide position 689 in both alleles, resulted in change of lysine to arginine at amino acid position 211. In Reh cells, C substituted with T in one of the alleles at nucleotide position 115, changed the CAG codon for glutamine to the stop codon TAG at amino acid position 20. The other polymorphisms revealed did not involve any change of the amino acids. Nucleotide positions 90 and 282 appeared to be more prone to polymorphism and have been previously described in the human lymphoblastoid cell line, TK6 (Grosovsky *et al.*, 1993), and recently in breast and colorectal cancer solid tumours (Gilles *et al.*, 2003). Our results regarding the nucleotide position 373 was confirmed by both groups: G instead of A in both alleles was invariably found, which resulted in valine codon for amino acid 106 (Grosovsky *et al.*, 1993; Gilles *et al.*, 2003).

Discussion

The results illustrate two important points: 1) it is most likely valine and not methionine that is amino acid 106 in TK1, and 2) amino acid 106 has profound effect on the kinetic and oligomerization properties of TK1. It remains to be explained why methionine at position 106 was found in previous investigations of TK1 derived from SV40 transformed fibroblasts (Bradshaw & Deininger, 1984) and from HeLa cells (Flemington *et al.*, 1987). As the differences might be due to different origins of the mRNA from which the cDNA was synthesized, cDNA and genomic DNA from various human cells and cell lines, also from SV40 transformed fibroblasts and HeLa cells, was sequenced. In each of the 22 independent isolations of genomic DNA and cDNA valine was found at position 106. The findings have later been supported by TK1 sequences from Burkitt lymphoma lymphocytes and from uterus tissue deposited in Genbank in 2001 by Strausberg for the NIH Mammalian Gene Collection (MGC) Project (gi:23503074) and by the already mentioned work of Gilles *et al.* (2003). The implications of the found polymorphisms for the function of TK1 *in vivo* are presented in General Discussion (Chapter 9) at the end of this thesis.

Initially, it was shown that mutation of valine 106 to methionine changed TK1 to a less stable enzyme, and a permanent tetramer with low $K_{0.5}$ towards thymidine ($K_{0.5}$ 0.4-0.6

μM, molecular size 100 kDa) irrespective of pre-exposure to ATP. For further investigation of the role of amino acid 106 for the function and structure of TK1 valine 106 was mutated to amino acids differing in size, conformation and polarity. After expression and purification, the effect of the mutated amino acids on kinetic and oligomerization properties of TK1 was examined.

The recombinant enzymes could be divided into two groups. Group I, the dimeric enzymes, contained V106A, V106I and V106T, and shared their properties with V106WT: *i.e.* in gel filtration experiments they eluted as both dimers and tetramers, with native molecular sizes of respectively 50 and 100 kDa, their –ATP form had high $K_{0.5}$ values 13 to 43, nonhyperbolic velocity vs. thymidine substrate curves, and the Hill coefficients below 0.8, indicating a negative co-operative reaction mechanisms. Hence, the hydrophobicity of the residue at site 106 was not critical for the function and conformation of TK1. Group II, the tetrameric enzymes, which contained V106G, V106H, V106K, V106L, and V106Q, shared their properties with V106M: *i.e.* they showed decreased stability when compared to Group I enzymes, in gel filtration experiments they eluted essentially as tetramers (native molecular size of 100 kDa), both their –ATP and +ATP forms had low $K_{0.5}$ values 0.3-1.2 μM, and the Hill coefficients were between 1.4 and 2 indicating positive co-operative reaction mechanism. If a mutant enzyme elutes with high proportion of its activity as a tetramer, it could be expected to have a lower $K_{0.5}$ value than an enzyme eluting with higher proportion of its activity as a dimer. Comparison of Fig. 6.8A and Table 6.2 shows that this is not the case. An explanation can be offered by the predicted secondary structure of TK1 as compared to enzymes of dNK/TK2 group (Chapter 4) whose 3D structure has been solved. Because TK1 is smaller than the other enzymes, one side of α-helix (probably α3) may constitute subunit interface, and the other side of the same helix may be involved in substrate interaction. In this way, a mutation at amino acid position 106 results not only in perturbation of dimer-tetramer transition, but also interferes with the substrate interaction.

The amino acids at site 106 in Group II enzymes differ in size, as well as in conformation and polarity. Thus, these properties do not appear to explain the decreased stability, the decreased $K_{0.5}$ value, nor the conformational changes. On the other hand, the amino acids at site 106 in Group I enzymes, are except for alanine of similar size and conformation, but of different polarities. Grouping of V106I with V106WT was expected, but absence of

nonpolar V106L and presence of polar V106T, was not. A property shared by the side chains of valine, isoleucine and threonine is branching at their β-carbon atom. β-branching amino acids are considered to destabilize α-helices because of steric clashes. Although leucine has the same hydrophobicity, size and conformation as valine and isoleucine, it differs from these amino acids by having a branch not on β-, but on its γ-carbon atom. The long hydrophobic side chain of leucine resembles the side chain of methionine in its length and the absence of the branched β-carbon. Also the α-helical propensity of leucine is very different from that of valine, isoleucine and threonine, but similar to that of methionine, and nearly as high as that of alanine. This may explain why the properties of V106L are the same as of V106M and the other tetrameric enzymes. V106A belongs clearly to the dimeric enzymes, but with its somewhat lower K_m for thymidine than seen for the other Group I enzymes, V106A might be an intermediate between the dimeric and tetrameric enzymes.

Initially, it was attempted to replace Val-106 with phenylalanine (codon GTG was changed to TTC), but it was not possible to express the V106F mutant protein. This may be due to the steric distortion of the α-helix by the bulky side chain of phenylalanine, giving a misfolded enzyme that is broken down in *E. coli*. On the other hand, a small side chain, like glycine, and to some degree alanine, will not sufficiently serve to stabilize the local, presumably helical structure.

Dimer-tetramer transition, which is dependent on enzyme concentration and pre-exposure to ATP (Munch-Petersen *et al.*, 1993; Munch-Petersen *et al.*, 1995a; Paper I, Berenstein *et al.*, 2000), has been suggested to be of significance for the fluctuation of TK1 activity during the cell cycle, with a low activity dimer dominating at low TK1 concentration during G1, and a high activity tetramer dominating in the S phase with high TK1 concentration (Munch-Petersen *et al.*, 1995a). Such transition would require an enzyme with substantial conformational flexibility. β-branching amino acids may have a regulatory role in such conformation dependent transitions, as they are known to increase the strain within an α-helix, and so to destabilize helix-helix interaction (Dao-pin *et al.*, 1990; Deber *et al.*, 1992; Deber *et al.*, 1993). Actually, valine-106 in human TK1 is preceded by another β-branched residue, isoleucine-105, and among the thirty residues in the segment 91-121, there are three α-helix breaking glycines and ten β-branching amino acids. Hence,

the whole segment is likely to maintain conformational flexibility of TK1 in response to cell cycle and substrate binding.

Sequence homology of TK1 to deoxyribonucleoside kinases with solved 3D structures – HSV1-TK and the cellular kinases dNK, dGK (Johansson et al., 2001) and dCK (Sabini et al., 2003) – is too low (about 10%) for a reliable homology model. However, despite the low sequence identity (10%) between HSV1-TK and the cellular kinases, they had very similar core structures (Eriksson et al., 2002). Therefore, TK1 may have the same overall structure as the other nucleoside kinases. A prediction of the secondary structure of TK1 (Cuff et al., 1998; Cuff & Barton, 1999; Cuff & Barton 2000, se Chapter 4) places Val106 in the middle of an α-helix which in aligns in CLUSTAL W (Higgins & Thompson, 1996) with α-helix 4, one of the interface helices of HSV1-TK (Wild et al., 1997). This suggested that the area surrounding Val-106 may be integrated into the oligomerization interface. The role of amino acid 106 and the sequence surrounding it for dimer-tetramer transition has been confirmed by cloning the putative interface fragment of TK1, TK1$^{66-136}$, and investigating the influence of V106M mutation on the oligomerization properties of the fragment. As described in the next chapter (Chapter 7), TK1$^{66-136(Val106)}$ eluted as a mixture of dimer and tetramer in gel filtration experiments, whereas TK1$^{66-136(Met106)}$ eluted as tetramer only. Such oligomerization pattern supported my assumption that the fragment TK1$^{66-136}$ is an integral part of TK1 oligomerization interface, and that amino acid 106 is indeed of significance for the subunit arrangement of the enzyme molecule.

Evolutionary conservation of valine at amino acid position 106 (or at position corresponding to this number) in TK1 sequences originating from viral and animal species, further supports its importance for the structure and function of TK1.

The size and conformation of amino acid 106 were concluded to be more important than polarity for structure and function of TK1. Valine-106 not only contributes to hydrophobicity of the putative α-helical interface, but also has the capacity to modulate subunit interaction that underlies the substrate and cell cycle regulated activity of TK1, known to be crucial for balancing the DNA precursor pools.

It will be extremely interesting to compare these results with the that of an X-ray structure of TK1 once it will become available.

Chapter 7

The putative interface fragment of human TK1

Background for the present investigation

The two methods commonly used for determination of protein structure are X-ray crystallography and nuclear magnetic resonance (NMR). The human thymidine kinase has been up to now refractory to crystallization attempts. NMR techniques do not require crystals and the protein structure can be examined in solution, but the size of the protein limits the use of NMR. Today, depending on the method and on the NMR spectrometer, the upper limit is for proteins of 20-30 kDa. Because the subunit size of TK1 is 25 kDa, and the enzyme exists as a dimer and/or tetramer, it is to big for NMR analysis. Besides, the NMR methods demand 1-2 mM protein solutions, i.e. 10 mg/ml for a 10 kDa protein, with the additional requirements of greater than 95% purity and complete solubility at these high concentrations.

In order to get information about the structure of TK1 and not to be limited by the size of the enzyme, I have cloned and expressed in E. coli a fragment of the TK1 gene coding for amino acids 66-136, and named it TK1$^{66-136(Val106)}$. Substitution of Val-106 with methionine by site directed mutagenesis gave TK1$^{66-136(Met106)}$. The TK1$^{66-136}$ fragment was supposed to be the interface domain of TK1, as it aligned in CLUSTAL W (Higgins & Thompson, 1996) with one of the dimer interface α-helices of the Herpes TK (α-helix 4). The interface domain of TK1 might be involved in subunit interaction between monomers and/or between pairs of dimers during the formation of dimers and/or tetramers, respectively. Besides Val-106 which was shown to be critical for dimer/tetramer formation (Paper I, Berenstein et al., 2000), this TK1 domain contains other putative function-essential sites: 1) the magnesium binding site with the conserved Asp-97 and 2) Phe-128,Gln-129,Arg-130,Lys-131, a group of conserved amino acids in the putative thymidine binding site. The TK1$^{66-136(Val106)}$ and TK1$^{66-136(Met106)}$ peptides were characterized by gel filtration and NMR.

Introduction

Many biological functions involve binding of proteins to other proteins and association of subunits into homo- or hetero-oligomeric proteins. Understanding of sources of affinity and specificity at interfaces can provide insights into principles of protein interaction. Modification and blocking of protein-protein interactions, as well as predicting protein-protein interaction sites is of direct relevance to the design of drugs against diseases such as BSE and Alzheimer, cancer, viral and other infectious diseases. Nearly two decades ago, Wlodawer *et al.* (1989) suggested a molecule inhibiting dimerization of HIV protease as a drug against the HIV virus, and 10 years ago, dimerization inhibitors of human glutathione reductase were suggested as possible antimalarial agents (Nordhoff *et al.*, 1993). Today, many potent inhibitors of dimerization of HIV protease, HIV reverse transcriptase, HIV ribonucleotide reductase, and Herpes virus DNA polymerase have been described (Zutshi *et al.*, 1998, and references therein; Archakov *et al.*, 2003, and references therein).

Interfaces and their properties

Protein-protein associations are to a large degree stabilized by hydrophobic interactions, but because hydrophobic forces are not specific, they would lead to all kinds of incorrect interactions in a cell (Chothia & Janin, 1975). Another fundamental aspect of protein-protein interactions is complementarity that applies both to the shape and to the charge of groups on the surface of proteins (Jones & Thornton, 1995). Only complementary surfaces can form proper hydrogen bonds and van der Waals contacts in order to close-pack together. Incorrect associations would result in poor packing and loss of hydrogen bonds to water (Chothia & Janin, 1975).

The ability of certain proteins to form complexes depends on the properties of their interfaces. The definition of which residues form the interface relies on the distance between any two atoms belonging to two residues from different chains: if the distance is less than 5.0 Å, or less than the sum of the corresponding van der Waals radii plus 0.5 Å, the residues are considered as belonging to the interface (Tsai *et al.*, 1996). The interfaces

can be imagined to stabilize both the association between the subunits, and the tertiary structure of the individual subunits (Miller *et al.*, 1987). Some aspects of the recognition sites in oligomeric proteins (as opposed to enzyme-inhibitor, or antibody-antigen complexes) will be shortly described here based on the literature from the last 3 decades. It is striking that although continuously larger data sets from protein data bases and more complicated algorithms are being used, the general conclusions remain valid.

Size of interfaces

When two proteins interact, they bury part of their solvent accessible surface into each other. Argos (1988) calculated water-accessible surface area (in Å^2) for isolated subunits from several proteins and then the accessible surface area for the respective oligomers. His calculations suggested that dimers lose about 12% of their accessible area (the average loss was 5-6% per monomer bound), trimers lose over 17% and tetramers nearly 21%. The percent area lost by oligomerization decreased as the number of amino acids in the monomer increased, because a minimum area is required for stable binding and smaller domains need to associate a greater percentage of their water accessible surface area (Argos, 1988). Later analysis of Jones & Thornton (1996) show an average interface area of about 1000 Å^2 for each monomer of homodimers.

Composition of interfaces

The analysis of amino acid composition of the interfaces has shown that the hydropathy of the interfaces lies between that of the buried (hydrophobic) and exposed (hydrophilic) protein surfaces (Argos, 1988; Janin *et al.*, 1988; Jones & Thornton, 1996). Jones & Thornton (1995) calculated that on average interfaces contain 68% non-polar atoms (carbon) and 32% polar (nitrogen, oxygen and sulphur), and in term of residues contain 47% hydrophobic, 31% polar and 22% charged. Homodimers, that rarely occur or function as monomers, were observed to have interfaces of higher hydrophobicity than heterodimers and antibody-antigen complexes (Jones & Thornton, 1996; Jones & Thornton, 1997). All interfaces had strong bias for large hydrophobic residues (Phe, Trp, Tyr) as well as Met and to a lesser extent, Leu (Argos, 1988; Ofran & Rost, 2003). Non-polar residues Ile, Leu, Phe, Val, Cys and Met contributed together with 33% to the subunit interface area, while charged residues Asp, Glu and Lys contributed much more to the accessible surface (33%)

than to interfaces (14%) or to buried surface (15%) (Janin *et al.*, 1988). Leucine and arginine were pointed out to be those residues that make the largest contribution (about 10% each) to the subunit interface area (Janin *et al.*, 1988). Especially Arg, in contrast to Lys, occurs often at the interfaces, about 5 times as frequent at subunit interfaces as inside the subunits (Janin *et al.*, 1988; Ofran & Rost, 2003). Ser, Ala and Gly are on average underrepresented (Ofran & Rost, 2003).

Energetic interactions between residues of interfaces

As mentioned before, hydrophobic side chains are the critical driving force in association of protein molecules. When percentage frequency of amino acid residues in the interface was plotted against their hydrophobicity, the interfaces were revealed as the hydrophobic patches on the surfaces of proteins (Jones & Thornton, 1995). Also, when proteins available as X-ray structures in their complexed form were analysed for clusters of interface residues it was shown, that the binding sites correspond to a large extent to the strongest hydrophobic patches (Young *et al.*, 1994).

Analysis of energetic contributions of individual amino acids to protein–protein binding (measurements of the change in free energy of binding upon mutation of different amino acids to Ala) revealed that the free energy was not distributed evenly across interface; instead, there were so called "hot spots" of binding energy made up of small subset of residues in the dimer interface (Bogan & Thorn, 1998). These hot spots were enriched in Trp, Tyr, Arg and Ile (Bogan & Thorn, 1998). (Ile appeared in hot spots 10 times as frequently as leucine and no reason for this could be found (Bogan & Thorn, 1998)). The aromatic residues are often found protruding from one subunit to the other, thus physically anchoring the two subunits together (Jones & Thornton, 1995). Arginine is capable of multiple types of favourable interactions: The guanidium group of Arg is the H-bond donor in 33% of all polar interactions between subunits (Janin *et al.*, 1988); Arg is also capable of forming a salt bridge with its guanidium motif and hydrophobic interactions with its three methylene carbon atoms (Bogan & Thorn, 1998). Removal from the solvent was a necessary condition for a residue to be energetically important for binding and could be accomplished by a surrounding seal of energetically unimportant residues (Bogan & Thorn, 1998). Also hydrophilic bridges across the binding interface have been found to

show positive correlation with the binding free energy and to significantly stabilize complexes in several cases (Xu *et al.*, 1997).

Preferences on residue pairing across subunit interface

The overall trends on residue pairing preferences were summarized by Glaser *et al.* (2001): the largest residue-residue preferences were recorded for pairs of large hydrophobic residues, as Trp and Leu, and the smallest preferences for small residues like Gly and Ala, presumably because the interaction strength is proportional to the interaction area (Glaser *et al.*, 2001). Hydrophobic-hydrophobic interactions were most common, but also contacts between the charged residues were observed. They followed charge complementarity, as Arg-Glu pairs, or could make hydrophobic contacts with their aliphatic chains, while their charged groups were solvent exposed and/or involved in hydrogen bonding (Glaser *et al.*, 2001). Contacts between residues in Cys-Cys pairs could be divided into two groups – with or without disulfide bonds (Glaser *et al.*, 2001). Cysteine bridges were observed more often than expected in interfaces of both homo-and hetero-complexes (Ofran & Rost, 2003). All other amino acids showed preference for interaction between identical residues only at interfaces of homo-dimers (Ofran & Rost, 2003). On average, identical residues with the same sequence position in the two interacting subunits made less than 2.5% of all interactions (Argos, 1988). Met was the amino acid that was remarkably involved in self-interactions (ratio of composition of this amino acid in self-interaction to its composition in all proteins was 4.88), and was followed by His and Phe (the ratio 2.55 and 2.11). The ratio for Val was 1.76 (Argos, 1988).

Secondary structure of interfaces

Interfaces differ a lot with respect to their secondary structures. The dimer interfaces of deoxyribonucleoside kinases consist, for example, exclusively of α-helixes (Eriksson *et al.*, 2002), while the dimer interface of thymidylate synthase is formed by an association between five-stranded β-sheets present in each monomer (Hardy *et al.*, 1987). Miller (1989) and Tsai *et al.* (1996) using a nonredundant set of proteins generalized that an average interface consists of an uncharged central region that could be formed by α-helixes, β-sheets, or α/ β (and α+β) structures, and an outer, charged region formed by loops that make hydrogen bonds both to the main chain and side chain from another

subunit. Based on the percentage frequency of α and β secondary structure in the interface residues, Jones & Thornton (1995) classified 53.1% of the protein interfaces in their data set as α, 21.8% as β, 12.5% as α/β, and 12.5% as coil. A motif that occurs frequently in protein-protein interfaces is the four-helix bundle (Lin et al., 1995), which has already been described in Chapter 4 for the interfaces of dNK and dGK (Johansson et al., 2001). Helices make a bundle if they have sufficient mutual contact and geometry (Lin et al., 1995).

Results

Induction of GST- TK1$^{66-136}$ fusion proteins and detection by CDNB assay

The cloned TK1$^{66-136}$ proteins had no enzymatic activity, and therefore the magnitude of induction by IPTG was estimated by measurements of expression of the GST part of the fusion protein. Fig.7.1 shows the reaction between GST-TK1$^{66-136}$ fusion proteins and GST substrate, CDNB.
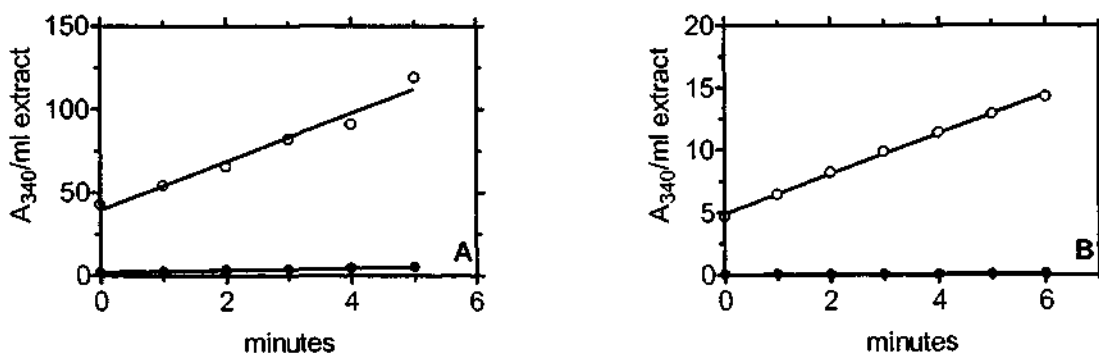
Figure 7.1. CDNB assay for GST-TK1$^{36-136}$ fusion proteins. A. TK1$^{66-136(\text{Val}106)}$ (KY895 as host bacteria, overnight induction with IPTG), B. TK1$^{66-136(\text{Met}106)}$ (BL21 as host bacteria, 6 hours induction with IPTG). The assay was performed on 5 or 50 µl bacterial extract as described in Materials and methods. A_{340} values were normalized to 1 ml bacterial extract.. Filled symbols, before addition of IPTG; open symbols, after incubation with IPTG.

The fold of induction in different experiments, as estimated with the CDNB assay, varied from about 200 to 600 for both TK1$^{66-136(\text{Val}106)}$ with KY895 as host bacteria and TK1$^{66-136(\text{Met}106)}$ with BL21 as host bacteria. The density of the bacterial culture increased no more than 10-fold during the same time period.

Purification of TK1$^{66-136}$ polypeptides by glutathione Sepharose chromatography

The purification procedure is summarized in Table 7.1.

Table 7.1. Purification of TK1$^{66-136(Val106)}$ and of TK1$^{66-136(Met106)}$

	TK1$^{66-136(Val106)}$			TK1$^{66-136(Met106)}$		
	Volume	Protein		Volume	Protein	
	ml	mg/ml	mg total	ml	mg/ml	mg total
Crude	50	13.06	653	50	5.59	279
Set	50	12.31	616	50	4.13	207
Pooled wash/equilibrium fractions						
	90		25	266		21
Cleavage fractions #1 and #2						
	3.2	0.81	2.6	8.9	0.56	4.9
Cleavage fractions #3, #4 and #5						
	4.8		0.087	13		0.2
Reduced GSH wash						
	4.8		16	13		2.9

All data are related to 1 litre of bacterial culture
The amount of protein in "crude" and "set" fractions of TK1$^{66-136(Val106)}$ is most probably overestimated due to the presence of 1% Triton. TK1$^{66-136(Met106)}$ was purified without Triton.

From 1 litre KY895 and BL21 cultures 2 to 3 mg of TK1$^{66-136(Val106)}$ and 4 to 6 mg of TK1$^{66-136(Met106)}$ respectively, usually was obtained.

Estimation of subunit molecular size and protein purity of TK1$^{66-136}$ proteins

The mass of the cloned TK1$^{66-136}$ was calculated from the amino acid composition: C^{66}THDRNTMEALPACLLRDVAQEALGVAVIGIDEGQFFPDIV/MEFCEAMANAGK TVIVAALDGTFQRKPFGAI^{136} to be 7.7 kDa (including the two additional amino acids, G and S, resulting from the thrombin recognition sequence). Standard SDS-PAGE methods are not suitable for peptides smaller than 10 kDa, and accordingly, I have initially tried to increase acrylamide concentrations up to 20% in the separation gel, and to include about 7 M urea in the gel and/or elecrode buffer (Swank & Munkres, 1971; Schleif & Wensink, 1987) without success. Instead, including 30% ethylene glycol in the gel, and replacing glycine with tricine in the electrode buffer gave good results (Fig. 7.2). The molecular size of the heavy protein band corresponds to 7.9 kDa, in agreement with the value calculated from the amino acid composition.
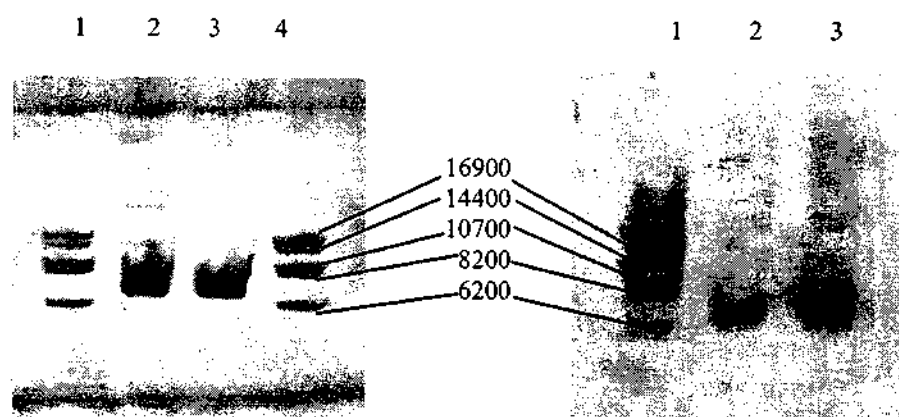
138

Figure 7.2. SDS-ethylene glycol gel electrophoresis of cleavage fractions after glutathione sepharose chromatography. Left gel: TK1$^{66-136(Val106)}$, right gel: TK1$^{66-136(Met106)}$. Lanes 1 and 4, Pharmacia marker proteins consisting of a mixture of myoglobines (in kDa from top to bottom: 16.9, 14.4, 10.7, 8.2 and 6.2). Lanes 2 and 3, 1 µl of the first two cleavage fractions (cl. #1 and cl.#2) corresponding to 0.8 and 0.7 µg protein for TK1$^{66-136(Val106)}$, and 0.4 and 0.6 µg for TK1$^{66-136(Met106)}$.

Further purification of the cloned TK1$^{66-136}$

As could be seen from Fig. 7.2 several contaminating bands were present in the cleavage fractions. Due to the hydrophobic nature of the cloned interface further purification of the TK1$^{66-136(Val106)}$ cleavage fraction was attempted on prepacked hydrophobic columns - Butyl Sepharose and Phenyl Sepharose (Pharmacia). Interaction of TK1$^{66-136(Val106)}$ with Butyl Sepharose was too weak at the conditions applied and 90% of the protein appeared in the wash fractions, while interaction with Phenyl Sepharose was so strong that the protein did not elute with low salt buffer. However, it was eluted with addition of 30% isopropanol or 20% acetonitrile, but unfortunately precipitated upon storage.

Because the percentage of contaminating proteins in the cleavage fractions was judged to be less than 10% (Fig. 7.2), the degree of purity of the protein preparations was judged to be sufficient for further analysis, and no further purification attempts were performed.

Estimation of apparent molecular size by gel filtration

It can be seen from Fig. 7.3, that TK1$^{66-136(Val106)}$ was eluted as two separate peaks of about 29 kDa and 12 kDa., while TK1$^{66-136(Met106)}$ was eluted as a single sharp peak of about 29 kDa. According to subunit size of 7.7 kDa, the TK1$^{66-136(Val106)}$ is judged to elute as a dimer and a tetramer, while TK1$^{66-136(Met106)}$ elutes exclusively as a tetramer. These results are in agreement with the results obtained for the active recombinant enzymes Ly-TK1^{Val106} and

Ly-TK1^{Met106}, as described in the previous chapter and in Paper I (Berenstein *et al.*, 2000) and Paper II (Frederiksen *et al.*, 2004).
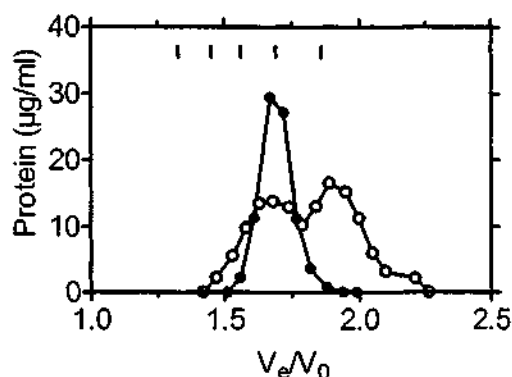
Figure. 7.3. Gel filtration of TK1$^{66-136}$ proteins. About 100 µg of TK1$^{66-136(Val106)}$ (open symbols) and TK1$^{66-136(Met106)}$ (closed symbols) from the thrombin cleavage fractions were injected into Superose 12 column. The molecular mass markers (|) are (from left to right in kDa): β-amylase, 200; bovine serum albumin, 66; ovalbumine, 45; carbonic anhydrase, 29; lysozyme, 12.4.

NMR analysis[1]

The 1D (one-dimensional) spectra are the first step in characterization of proteins by NMR. The spectra are shown in Fig. 7.4 for the TK1$^{66-136(Val106)}$ and in Fig. 7.5 for the TK1$^{66-136(Met106)}$.
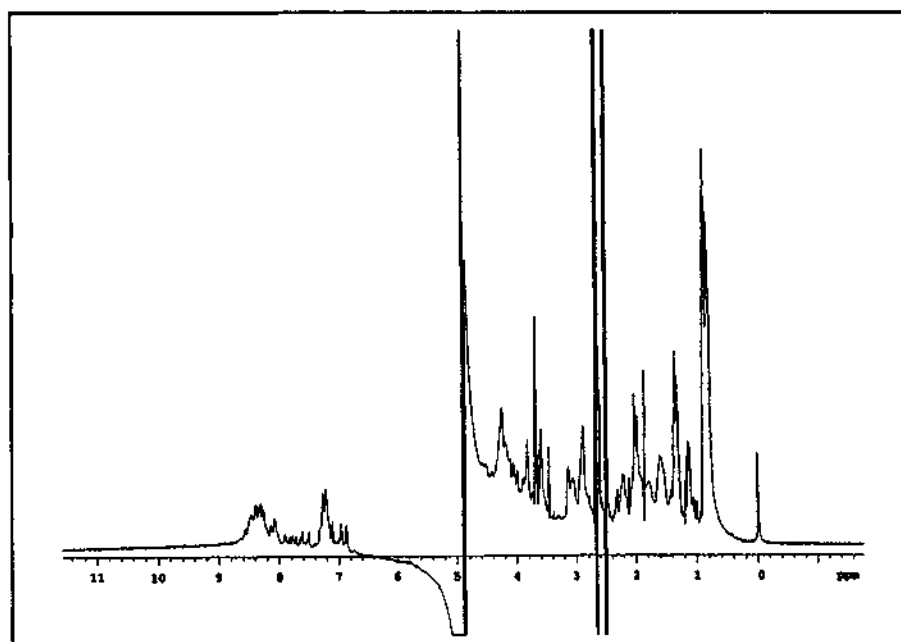
Figure 7.4. 1D spectrum of TK1$^{66-136(Val106)}$.

---

[1] Se Appendix 1 for introduction to NMR principles
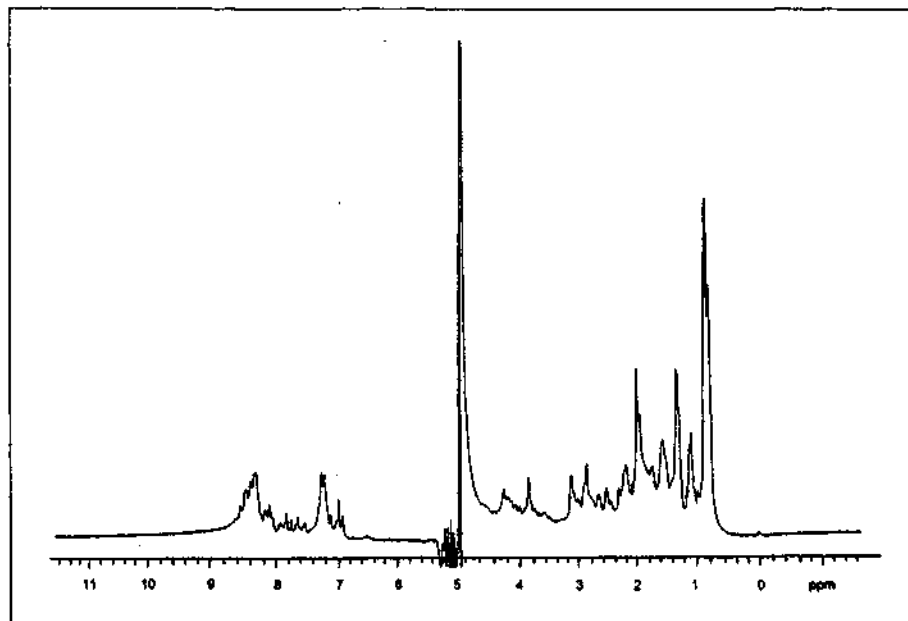
Figure 7.5. 1D spectrum of TK1[66-136(Met106)]

The signal-to-noise ratio is satisfactory and the chemical shift dispersion appears to be quite good. The purification of TK1[66-136(Met106)] was successful, as apart from amino acid resonances only a signal for water can be seen (at about 4.8 ppm). TK1[66-136(Val106)] spectrum shows two additional signals - at 2.5 and 2.6 ppm, probably residual impurities from methylene groups of ε-aminocaproic acid).

Resonances of several chemical groups can be identified in the 1D spectra. The strongest resonances in 1D spectra are observed in the region 1 ppm, where most of the resonances from valine, threonine, leucine and isoleucine methyl groups are overlapped. At the extreme left from about 7 to 11 ppm are the resonances of the amide protons, and in-between methylene groups are observed at 2-3 ppm, Cα protons at 4-5 ppm, and the protons of the aromatic rings at 6-7.5 ppm. The resonances for aromatic amino acids show partial overlap with side chain and backbone amide protons (Wüthrich, 1986). Spectra of both TK1[66-136] proteins indicate that the proteins might be either folded into α-helix, or are random coils. When residues are found in β-strand or extended conformation their ^{1}H chemical shift for Cα protons and for amide protons in the 1D spectrum is on average moved about 0.4 ppm to the higher ppm values (Wishart, et al., 1991).

Either COSY or NOESY experiment is usually run in order to assess if the protein of interest will give 2D spectra of required quality. Here NOESY was chosen, the results are shown in Fig. 7.6 and Fig. 7.7. As can be seen, the cross-peaks that are present overlap and are very broad. In small molecules, with molecular weights up to 10000, the line widths of

the resonances are usually narrow, they get broader in larger molecules (Wüthrich, 1995). Because resolution depends inversely on resonance line width, overlap of cross-peaks in 2D NMR spectra is a major problem in studies of larger proteins.
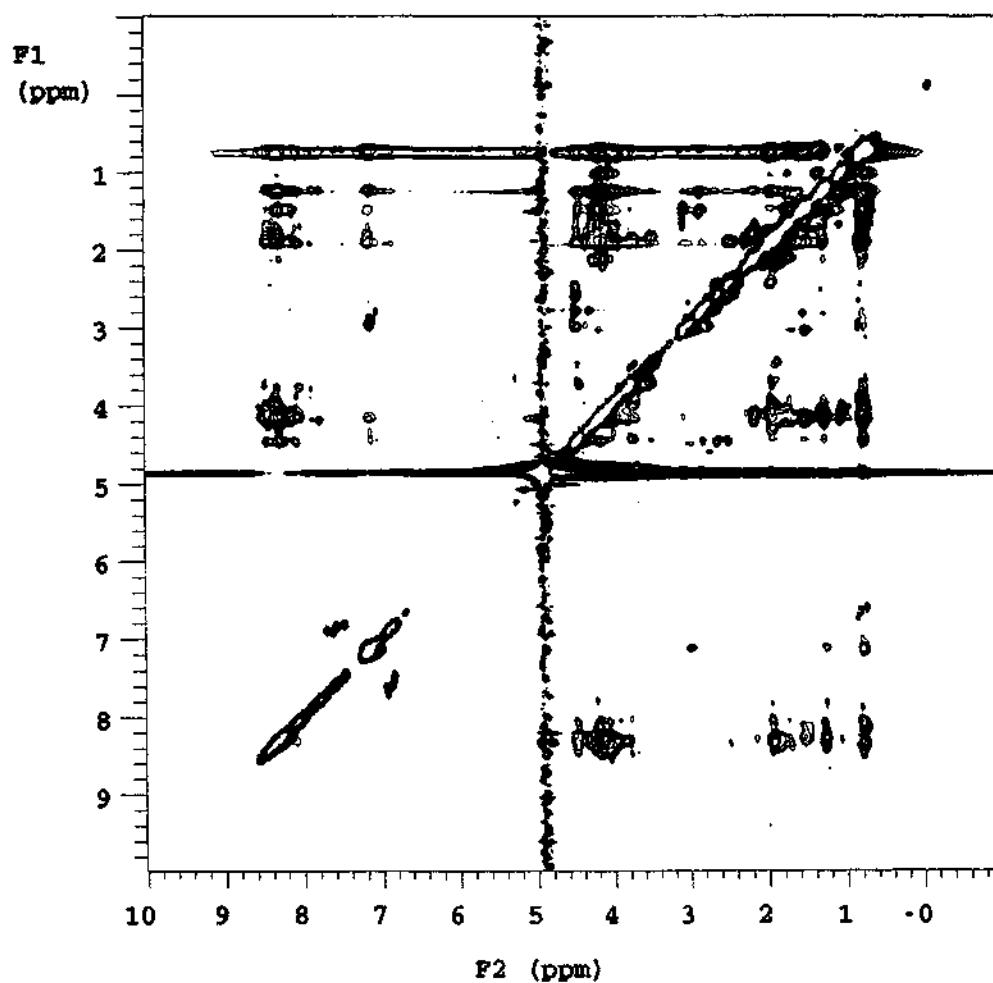
Figure 7.6. 2D NOESY spectrum of TK1[66-136(Val106)]
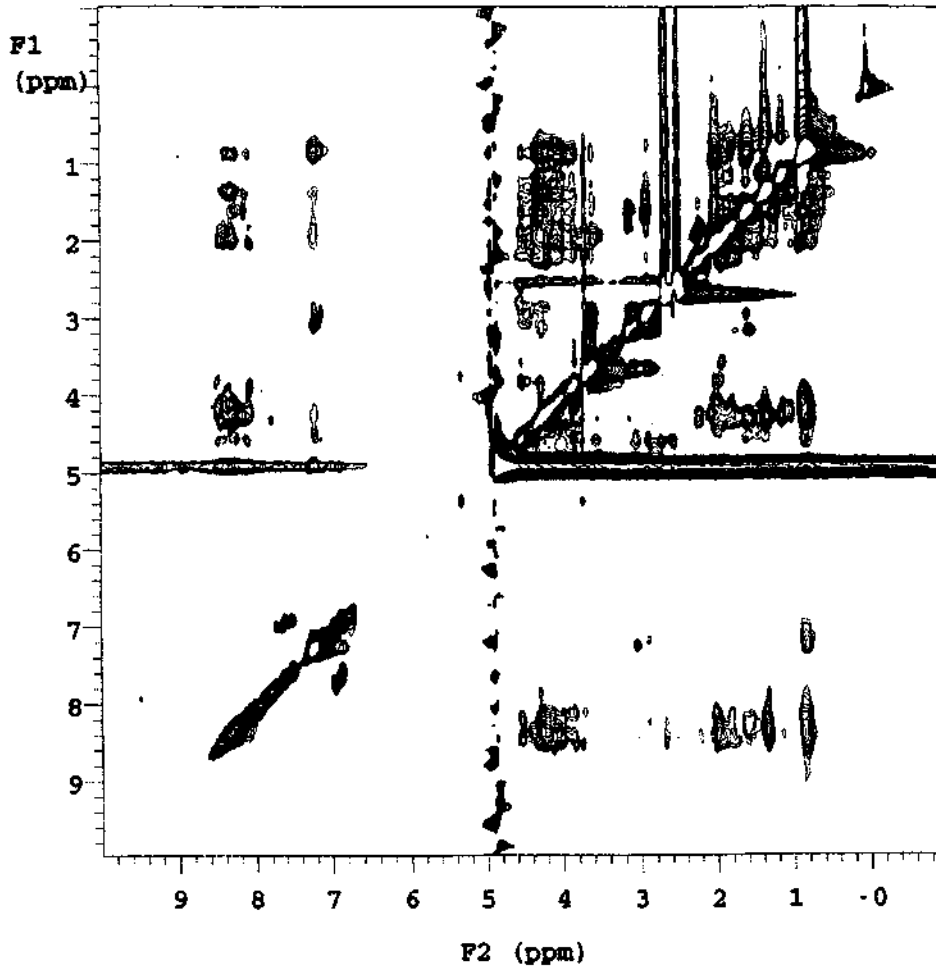
Figure 7.7. 2D NOESY spectrum of TK1$^{66\text{-}136(Met106)}$

An obvious reason for broadened cross-peaks in Fig. 7.6 and 7.7 is the self-assembly of the cloned interface into dimers and tetramers (TK1$^{66\text{-}136(Val106)}$) and into tetramers (TK1$^{66\text{-}136(Met106)}$). Each subunit contributes signals to the NMR spectra, and it is difficult to distinguish between the intra-subunit and inter-subunits cross-peaks (Ferentz & Wagner, 2000). The residual contaminating proteins in preparations of TK1$^{66\text{-}136}$ or the residual thrombin could not have any effect on my results. The concentration of TK1$^{66\text{-}136}$ was 0.5 mM, while thrombin concentration was about 1 μM, and the concentration of the high MW contaminating proteins was judged to be below 5% of the total protein. Even in-cell (*in vivo*) NMR-experiments where all proteins in a bacterial cell were labelled with an NMR-active isotope, only resonances from the protein of choice, overexpressed by induction[2], were visible (Serber & Dötsch, 2001).

---

[2] The highest overexpression levels possible are up to 30-50%, and can be obtained in yeast or in insect cells.

Discussion

The amino acid composition of the cloned TK1$^{66-136(Val106)}$ fragment and its predicted secondary structure is shown in Fig. 7.8. It contains 71 amino acids plus Gly and Ser that come from the thrombin recognition site (Fig. 7.8).

```
          1--------11--------21--------31--------41--------51--------61--------71-
A
          GSCTHDRNTMEALPACLLKIVAQEALGVN TV:DEGQFFPD!YEK KAMANAGK  IV ;LDGTFQRKPFGAI
Jalign    ------------------------EEE------H--HHHHHHH------EEEEEE-------------
jfreq     --E-------HHHH------HE------EEEEE--------HHHHHHHHH----EEEEEE-----------
jhmm      ----EEE-------EEEEE------------E-------EEEEEE----EEEEEEEEEE----------
jnet      --E------------HHHHHH----EEEEE-------HHHHHHHHH----EEEEEE------------
jpssm     -----------------HHHHHHH---EEEEE--------HHHHHHHHHH---EEEEE------------

jpred     ----------------HHHH------EEEEE--------HHHHHHHHH----EEEEEE-----------

                         α2         β3            β4       β1

B
          GSCTHDRNTMEALPACLLKIVAQFALGV: :-:DEGQFFPD!My F KAMANAGK:. .:..LDGTFQRKPFGAI
jalign    ------------------------EEEE-----H--HHHHHHH------EEEEEE-----------
jfreq     --E--------H-H-----HHEEH----EEEEE--------HHHHHHHH----EEEEEE-----------
jhmm      -----EEE-------EEEEEE-----------E------EEEEEEE----EEEEEEEEEE----------
jnet      --E-------------HHHHHH----EEEEE--------HHHHHHHHH----EEEEEE-----------
jpssm     ------------------HHHHHHHH---EEEEE--------HHHHHHHHHH---EEEEEE----------

jpred     ------------------HHHHHHH----EEEEE--------HHHHHHHHHH----EEEEEE-----------

Chou-Fasman                                    HHHHHHHHHHHHHHHHHHH
                         α2         β3            β4       β4
```
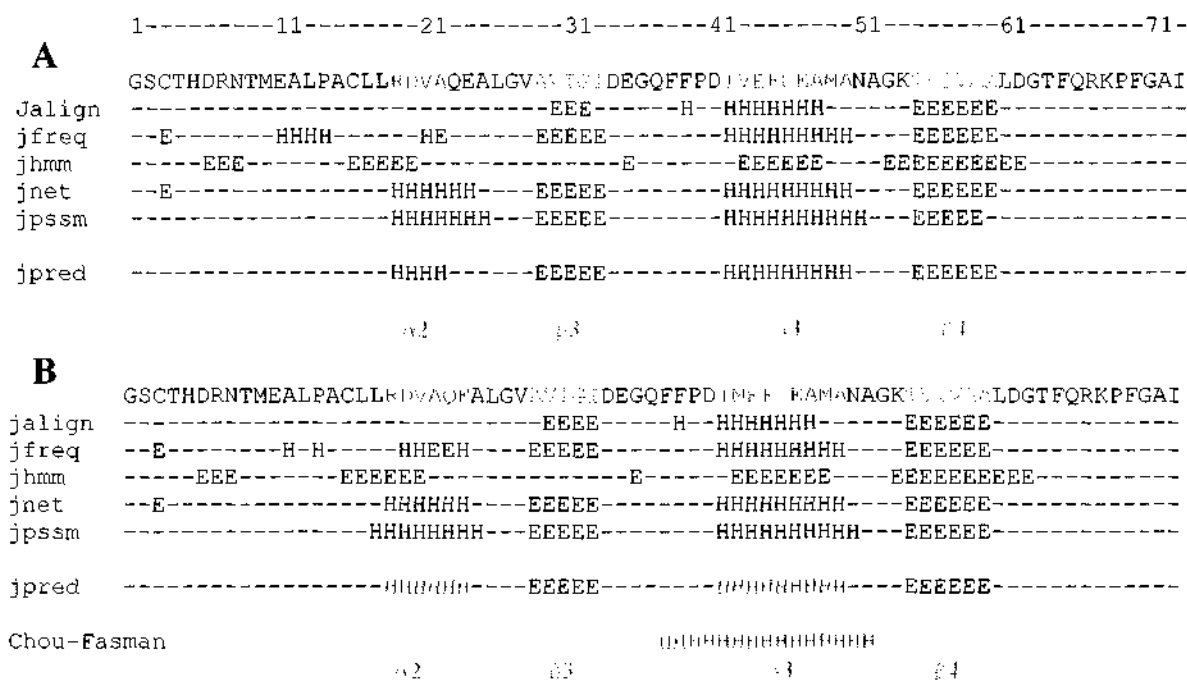
Figure 7.8. Amino acid sequence and the predicted secondary structure of the cloned putative interface fragment TK1$^{66-136(Val106)}$ (A) and TK1$^{66-136(Met106)}$ (B) by the Jpred method of Cuff & Barton (2000). H – α-helix, E – β-strand, „-„ means residues that cannot be classified as either α- or β-structure. Chou & Fassman (1978) prediction for a region corresponding to α3 is also shown. The numbers of α-helixes and β-strands are the same as in Fig. 4.3 for the entire TK1 subunit.

43 out of the 71 residues, or 60%, are non-polar (Gly and Ser that come from the thrombin recognition site are not taken into account while discussing properties of the putative TK1 interface). Leaving out the two Met residues with the somewhat nucleophilic sulphur gives 41 out of 71, or 58% of hydrophobic residues. This high number of hydrophobic residues, among them 5 Phe, 5 Leu, 5 Ile and 6 Val, as well as presence of charged residues 5 Glu, 5 Asp, 3 Arg and 2 Lys, correspond well to the description of amino acid content in interfaces. The secondary structure algorithm predicts the putative interface to form two α-helixes and two β-strands (Fig. 7.8). According to the model of 3D structure of TK1 described in Chapter 4, the two β-strands belong to the core of the enzyme and the interface is probably made by the α-helixes, α2 and α3 and by loops that surround them. The helices account for 12 out of the 71 residues, or 17% of the total cloned fragment. The helices though might be

rather weak. $\alpha 2$ of TK1$^{66-136}$ consists of only 4 residues and has no stabilizing hydrogen bond between the CO group and NH group of the main chain (from residue n to residue $n+4$) $\alpha 2$ of TK1$^{66-136(Met106)}$ consists of 6 residues and contains two hydrogen bond between the CO and NH groups of the main chain. Still, it is a weak helix compared to helices of over 8 amino acids, where the central residues have two hydrogen bonds between the corresponding CO and NH groups. Such central residues must be more geometrically constrained than the end residues. Jpred prediction of α-helix 3 (IV^{106}EFCEAMA) as a genuine α-helix is supported by CD spectroscopy study of synthetic peptides (Behrends et al., 1996). Though, both TK1$^{66-136}$ peptides contain two negatively charged residues Glu-107 and Glu-110 spaced 3 positions from each other. Because of the twist of the α-helix they must be close to each other in space and the repulsive forces probably destabilize the helix. Presence of β-branching amino acids in both $\alpha 2$ and $\alpha 3$ helices (Val-84, Ileu-105, Val-106) is another helix destabilizing factor. This effect is pronounced in aqueous solution but not in the hydrophobic environment (Li & Deber, 1994). It can be imagined that the remainder of TK1 protein creates the proper environment necessary for the correct folding of the interface in vivo.

These $\alpha 2$ and $\alpha 3$ helices are homologous to the $\alpha 4$- and $\alpha 6$-helices from each subunit of the dimer interface of dNK and dGK, that together make the four-helix bundle. The interacting residues of the dNK interface are four aromatic, four aliphatic and four small. $\alpha 2$- and $\alpha 3$-helices of the putative TK1 interface contain one aromatic and six aliphatic residues that together with the charged Asp, Glu and Arg might make the interacting residues of putative TK1 interface. The putative TK1 interface helices are shorter, 4 and 9 amino acids respectively, as compared to 19 and 18 amino acids in the dNK interface helices and cannot be classified as the four-helix bundle until the structure of TK1 has been elucidated.

When the $\alpha 3$-helix (aa 105-114 in the fragment, but 104-113 in the whole TK1) is subjected to helical wheel analysis (Fig. 7.9) it becomes obvious that the helix is polar on one side and hydrophobic on the other side.
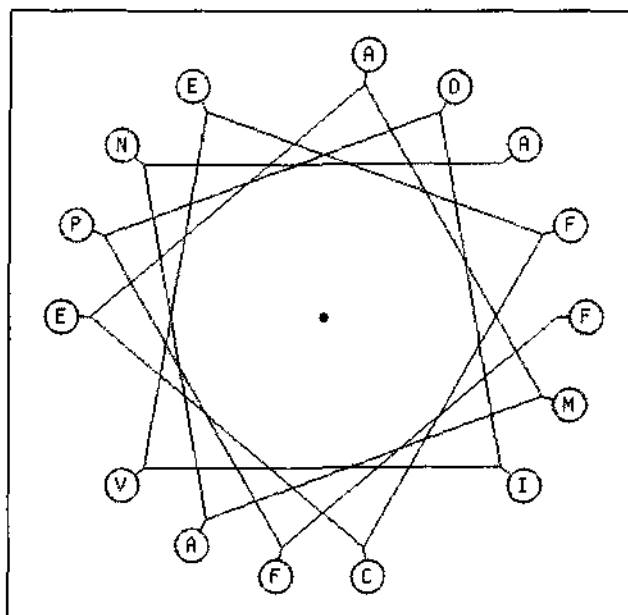
Figure 7.9. Helical wheel analysis of TK1 sequence [101]FFPDIVEFCEAMANA[115] (performed by a package from: http://marqusee9.berkeley.edu/kael/helical.htm)

As could be seen from the gel filtration experiments of $TK1^{66-136(Val106)}$ and $TK1^{66-136(Met106)}$, the exchange of Val to Met results in a permanent tetramer form of the fragment. Therefore the residue at position 106 and the neighbouring residues may be strongly assumed to belong to the oligomerization interface. We do not know if the tetramer is a dimer of dimers, where another interface region is involved, or if the tetramer is formed through interaction of the same two helices from each of the four subunits. These two possibilities should be distinguished by comparison of well-resolved (3D or 4D) NMR resonances for $TK1^{66-136(Val106)}$ and $TK1^{66-136(Met106)}$. If possibility 1) is correct, then the interactions that lead to the formation of tetramers cause less environmental change for the residues at the studied interface than if the possibility 2) is correct. The environmental change for the interface residues and/or the conformational change induced by complex formation will be reflected in changed chemical shifts. Such shift perturbations measurement have been described to locate interfaces on the individual binding partners (Gronenhorn & Clore, 1993; Zuiderweg, 2002).

Chapter 8

Investigation of the C-terminally truncated TK1 - TK1Δ40

Introduction

The C-terminal residues in human and mouse TK1 are known to be essential for the specific degradation of the enzyme at mitosis and for the cell cycle regulated expression of TK1 (Kauffman & Kelly, 1991; Kauffman *et al.*, 1991; Sutterluety *et al.*, 1996; Ke & Chang, 2004). Sequence KEN present in the C-terminal part of human TK1 (Fig. 8.1) has been shown to be responsible for mitotic degradation via a ubiquitin-proteasome pathway: it is required for recognition by an ubiquitin E3 ligase - anaphase-promoting complex (APC) complexed with its activator protein Cdh1 (Ke & Chang, 2004). The KEN box has been demonstrated to be targeting signal for APC-Cdh1 complex in several cell cycle-regulated proteins, including the R2 subunit of mouse ribonucleotide reductase (Pfleger & Kirschner, 2000; Chabes *et al.*, 2003). The C-terminal region of other proteins, for example of ornithine decarboxylase and of transcription factor IRF-1 (interferon regulatory factor) are also found to be critical for their turnover (Ghoda *et al.*, 1992; Nakagawa & Yokosawa, 2000; Zhang *et al.*, 2003).

[177]GADKYHSVCRLCYFKKASGQPAGPDNKENCPVPGKPGEAVAARKLFAPQQILQCSPAN
β5   α5
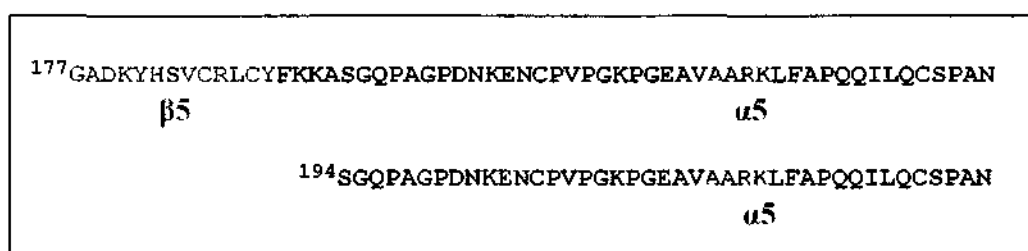
[194]SGQPAGPDNKENCPVPGKPGEAVAARKLFAPQQILQCSPAN
α5

Figure 8.1. Sequence of amino acid residues removed from TK1 upon construction of TK1Δ60 and TK1Δ40 respectively. The putative α- and β-forming residues according to JPred (se Fig. 4.3), and the KEN box (recognition sequence for proteolysis by anaphase promoting complex) are also shown.

Progressive deletion of three, six, ten, twenty and thirty C-terminal residues from mouse TK resulted in stepwise loss of its growth-dependent regulation: The TKΔ30 protein was constitutively expressed because it was resistant to mitosis-specific degradation (Sutterluety et al., 1996). Mouse TK C-terminally truncated by 37 residues, i.e. with Q196 as the last one, had full enzymatic activity, while if truncated by 44 residues, i.e. with Y189 as the last one, had no activity (Mikulits et al., 1997). Here, it is interesting to notice that thymidine kinase of vaccinia virus, highly homologous to human TK1, has no C-terminal domain, and its last amino acid corresponds to residue 192 in the human enzyme (Fig. 4.1). The absence of TK C-terminal domain in vaccinia virus (Fig. 4.1) parallels the absence of cell cycle regulation and a constantly high vaccinia TK activity in the infected cell, which might be of evolutionary advantage for replication of the virus. The residues up to position 192 in the human TK1 might then be important for the activity of the enzyme.

Two C-terminally truncated TK1 mutants, TK1Δ40 (Δ41) and TK1Δ60 (Δ58), were constructed by introduction of stop codons at positions 194 and 177 respectively into the amino acid sequence of TK1 (described in Materials and methods; Kristensen, 1996). The original purpose was to create TK1 mutants with lower molecular weight, and therefore susceptible to NMR analysis, but at the same time preserving the enzymatic activity of the wild type protein. Both TK1Δ40 and TK1Δ60 were expressed in E. coli as GST-fusion proteins and both had GST activity, while only TK1Δ40 showed thymidine kinase enzymatic activity (Kristensen, 1996). According to prediction of TK1 secondary structure in Fig. 4.3, deletion of 58 C-terminal amino acids removes the last β strand (β5) and the last α helix (α5), while removal of 41 amino acids from the C-terminal leaves the β5 strand intact, and only removes α5 helix (Fig. 8.1). Comparison to the 3D structure of HSV1 TK shows that the β5 strand and the few following residues are important for binding of ATP (among others Gln-331 that makes a single hydrogen bond to the adenosine moiety of ATP in HSV1 TK; Wild et al., 1997). The residues of TK1 removed in Δ58, but preserved in Δ41, are: $G^{177}ADKYHSVCRLCYFKKA^{193}$. Asp-179 might be the putative hydrogen acceptor corresponding to Gln-331 of HSV1 TK, and Arg-186 might supply hydrogen bonds to the phosphates of ATP.

When mouse thymidine kinase was synthesized in vitro in a coupled transcription/translation system, the enzymatic activity of a C-terminally truncated TKΔ30 mutant was 1.5 to 2-fold higher than the activity of the wild type (Posch et al., 2000). The

authors hypothesized that the C-terminus limits the access of substrates to the active sites, so that deletion of the C-terminal resulted in better accessibility for substrates and in higher specific activity of the TKΔ30 protein (Posch *et al.*, 2000).

In order to uncover the role of the C-terminal domain we have performed *in vitro* characterization of kinetic and oligomerization properties of TK1Δ40.

Results

Subunit and native molecular size

The subunit molecular size of TK1Δ40 was, as expected, about 20 kDa (Kristensen, 1996, and Fig. 8.2).
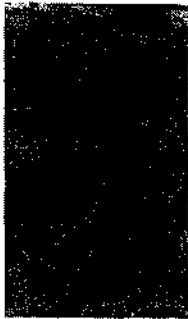
Figure 8.2. SDS gel electrophoresis of CM-Sepharose fraction of TK1Δ40 (left lane). Approximately 0.5 µg protein was applied to the gel. The standard proteins (in kDa) are from top to bottom: phosphorylase B (94), bovine serum albumine (67), ovalbumine (43), carbonic anhydrase (30), soybean trypsin inhibitor (20,1) and α-lactalbumin (14,4).

The apparent native molecular size of TK1Δ40 as determined by gel filtration on the Superdex column was about 45 kDa (Fig. 8.3A). According to the subunit size, TK1Δ40 eluted predominantly as a dimer, but with a pronounced shoulder of apparent molecular size of about 70 kDa, which corresponds to the tetrameric form of the enzyme.
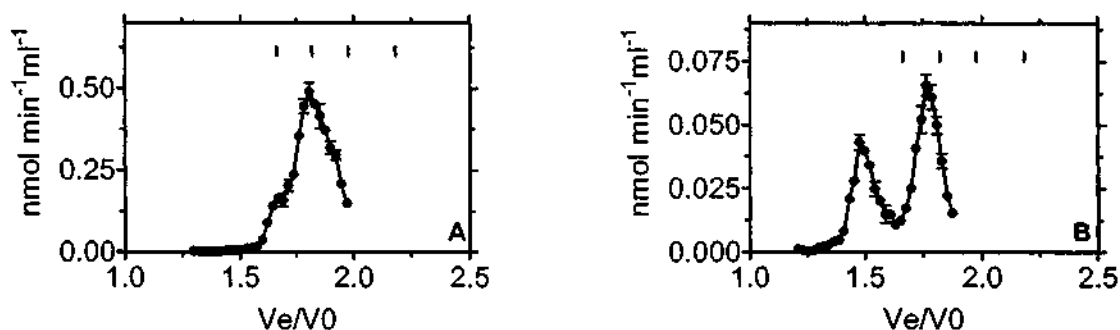
Figure 8.3. Estimation of native molecular size by gel filtration. About 1 µg protein (200 µl from solution of 5 µg/ml) of TK1Δ40 (A) and of TK1 wild type (B) was injected into a Superdex 200 column precalibrated with (from left to right) following standard proteins (|): bovine serum albumine (67 kDa), ovalbumine (45 kDa), carbonic anhydrase (29 kDa) and cytochrom C (12.4 kDa).

Absence of a separate tetrameric peak of TK1Δ40 (Fig. 8.3A) is surprising compared to the elution profile of TK1 wild type (Fig. 8.3B), where the dimeric and tetrameric peaks are well separated (apparent molecular sizes of about 50 kDa and 100 kDa). Also, the tetrameric form of TK1Δ40 elutes later than expected, and so seems smaller than its calculated size. A plausible explanation might be that the absence of the presumably disordered C-terminal makes the tetrameric form of TK1Δ40 much more compact than the corresponding tetrameric form of TK1 wild type.

Enzyme stability

As shown in Fig. 8.4, the stability of TK1Δ40 at 15°C is considerably higher than that of the wild type enzyme. The un-truncated, wild type TK1 lost 50% of its activity within 6.5 hours ($t_{1/2} \sim 6.5$ hours), while TK1Δ40 had barely lost any activity in this period ($t_{1/2} \sim 43$ hours). Recent experiments performed at 25°C and 37°C confirm the higher stability of C-terminally deleted TK1, and show that increase in stability is proportional to the length of the deletion. The order of several C-terminally truncated TK1 proteins from the most to the least stable was: TK1Δ40, TK1Δ30, TK1Δ20 and TK1 wild type (Chunying Zhu, personal communication). As it is very likely that the terminal 40 amino acids are disordered (se discussion), it is probable, that as the disordered stretch in a protein sequence increases, its thermodynamic stability decreases, and the quicker is the process of transition of the whole

protein from a folded, native state to an unfolded, denatured state. The compact structure of TK1Δ40 might increase its thermostability.
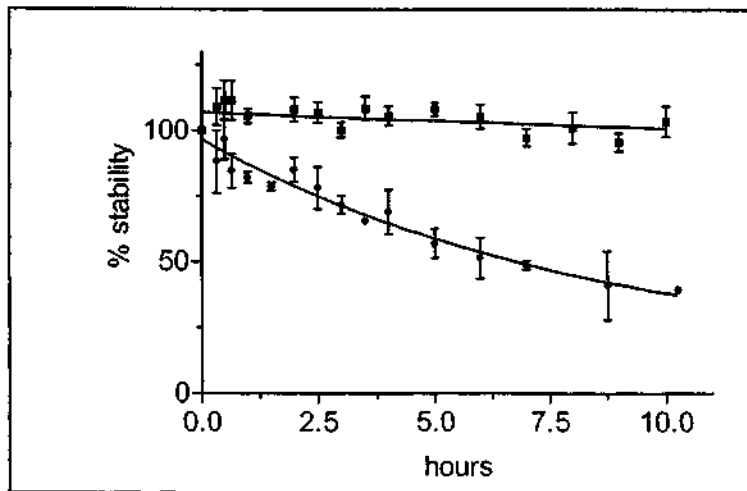
Figure 8.4. Stability of TK1Δ40 as compared to non-truncated TK1. TK1Δ40 (squares) and the wild type TK1 (circles) were incubated in phosphate buffer at 15°C. Each point is derived from at least two experiments, the bars are the standard error of the mean (SEM). A somewhat lower initial value of TK1Δ40 enzyme activity at time "0" was repeatedly seen. The curves were fitted by non-linear regression analysis with Graphpad Prism software.

Enzyme kinetics

Fig. 8.5 is an example of the thymidine substrate kinetics of TK1Δ40. Experiments were performed at saturating concentrations of ATP for both the TK1Δ40-ATP (incubated and stored without ATP) and TK1Δ40+ATP (incubated and stored with ATP) forms of the enzyme.
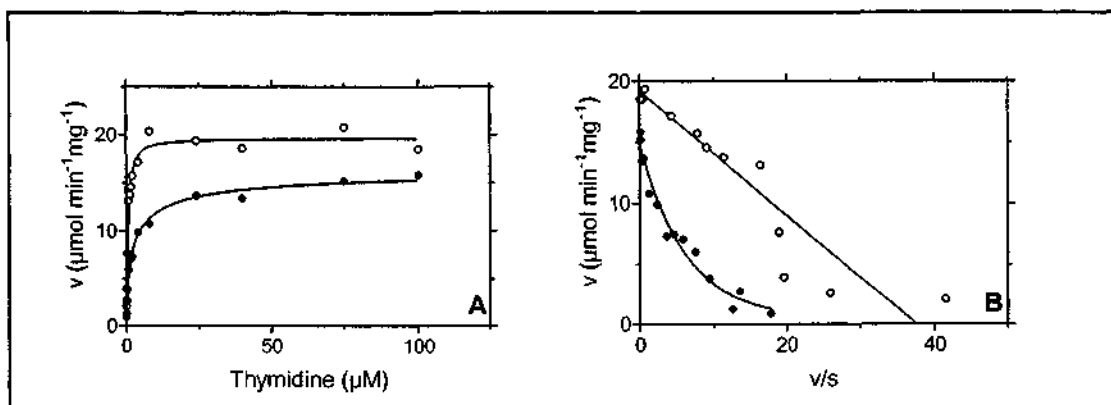
Figure. 8.5. Thymidine substrate kinetics of TK1Δ40 with enzyme concentration of 1 ng/ml in the assay mixture. A. The initial velocity of dTMP formation as a function of thymidine concentration. B. Hofstee plot of the same data. TK1Δ40-ATP form, closed symbols, TK1Δ40+ATP form, open symbols.

The kinetic mechanism of TK1Δ40 (Fig. 8.5) is the same as the mechanism of TK1 wild type enzyme described earlier (Munch-Petersen et al., 1993; Paper I, Berenstein et al., 2000). TK1Δ40-ATP displays a nonhyperbolic, "creeping" binding curve and a biphasic kinetic pattern in Hofstee plot (n value 0.64), while TK1Δ40+ATP gives a straight line indicating a hyperbolic rectangular reaction mechanism (n value 1.1). The $K_{0.5}$ values determined here were respectively 2.5 μM and 0.6 μM for the –ATP and +ATP form of TK1Δ40, and the V_{max} values were approximately 17 and 20 μmol/min/mg for the –ATP form for the +ATP form respectively. Independent estimation of the specific activity performed at 100 and 120 μM thymidine have given values of 22,1 ± 1,8 μmol/min/mg (n = 5) for the –ATP form, and 27,2 ± 2,3 μmol/min/mg (n = 8) for the +ATP form. Simultaneous estimation of the specific activity at 100 and 120 μM thymidine has given 20,9 ± 1,4 μmol/min/mg (n = 4) for TK1-WT-ATP and 20,0 ± 1,0 μmol/min/mg (n = 4) for TK1-WT+ATP. The kinetic values of TK1Δ40 are compared to TK1-WT in Table 8.1.

Table 8.1. Kinetic values of TK1-WT and of TK1Δ40

|	TK1 WT		TK1Δ40	
	-ATP	+ATP	-ATP	+ATP
V_{max} (μmol/min/mg)	20	20	22	27
K_{cat}[a] (sec^{-1})	8,5	8,5	7,8	9,6
$K_{0.5}$ (μM)	15	0,5	2,5	0,6
$K_{cat}/K_{0.5}$ (sec^{-1}μM^{-1})	0,57	17	3,1	16

[a]Calculations of K_{cat} were made using the calculated MW of TK1-WT equal to 25500 and MW of TK1Δ40 equal to 21400.

V_{max} of TK1Δ40 is somewhat higher than V_{max} of TK1-WT, but K_{cat} is practically the same, as is $K_{0.5}$ and $K_{cat}/K_{0.5}$ of the +ATP form. On the other hand, the –ATP form of TK1Δ40 has 5-6-fold lower $K_{0.5}$ and therefore 5-6-fold higher $K_{cat}/K_{0.5}$ (i.e. 5-6-fold higher enzymatic efficiency) than the TK1-WT enzyme. Thus, exposure of TK1Δ40 to ATP increases substrate affinity about 4-fold ($K_{0.5}$ values 2,5 and 0,6 respectively for the –ATP and +ATP form), whereas exposure of TK1 wild type to ATP increased substrate affinity 20-30-fold ($K_{0.5}$ values about 15 and 0.5 μM respectively for the –ATP and +ATP form (Munch-Petersen et al., 1993; Paper I, Berenstein et al., 2000).

Discussion

The C-terminal region in human and mouse TK1 has been shown to be dispensable for TK1 activity *in vivo*, and at the same time, to be essential for the specific degradation of the enzyme at mitosis and for the cell cycle regulated expression of TK1 (Kauffman & Kelly, 1991; Kauffman *et al.*, 1991; Sutterluety *et al.*, 1996; Ke & Chang, 2004). In order to get more insight into the role of the C-terminal region, I have determined kinetic, oligomerization and stability properties of TK1Δ40 *in vitro*, and analyzed amino acid composition of the 40 C-terminal residues.

Kinetic aspects

The about 6-fold decrease in $K_{0.5}$ for thymidine, as demonstrated here for the –ATP form of the C-terminally truncated TK1, is in agreement with the hypothesis of Posch *et al.* (2000) that the C-terminal limits access of substrate to the active site. This decrease in $K_{0.5}$ can also explain the finding that the C-terminal truncation increases kinetic activity of the mouse thymidine kinase (Posch *et al.*, 2000). In the wild type human TK1 enzyme, binding of ATP which induces formation of the tetrameric form (Munch-Petersen *et al.*, 1993; Paper I, Berenstein *et al.*, 2000) may at the same time induce a conformational change that pulls the C-terminal auto inhibitory domain away from the active site, thus permitting better access for the substrate. Removal of the C-terminal increases substrate accessibility considerably (6-fold lower $K_{0.5}$), but pre-assay ATP exposure of TK1Δ40 is necessary to reduce the $K_{0.5}$ value further to 0.5 µM.

Aspects of the *in vivo* and *in vitro* stability of TK1Δ40

The higher *in vivo* stability of the C-terminally truncated TK1 (Kauffman & Kelly, 1991; Kauffman *et al.*, 1991; Sutterluety *et al.*, 1996) can be explained by the recent finding of Ke & Chang (2004), that a KEN motif in the C-terminal targets TK1 for destruction by Cdh1-APC complex (se Chapter 2). Other research groups have demonstrated that the

KEN box is a recognition signal for proteolytic degradation of several cell cycle-regulated proteins (Pfleger & Kirschner, 2000) including ribonucleotide reductase (Chabes *et al.*, 2003). SCF ubiquitination complexes may also be utilized to control cell-cycle expression of TK1. The findings of Chang and collaborators showed that TK1 is hyperphosphorylated on serine (s) residues during mitosis in several human cell lines (Chang *et al.*, 1994), that the primary mitotic phosphorylation site is S^{13} (Chang *et al.*, 1998) and that phosphorylation of S^{13} contributes to the cell-cycle regulated degradation of TK1 via the SCF ubiquitine ligase and proteasome-mediated proteolytic pathway (Ke *et al.*, 2003). Still, at least in mouse cells, a proteasome-independent mechanism might as well be involved, because proteasome inhibitors failed to stabilize substrate-binding mutants of mouse TK in proliferating fibroblasts (Posch *et al.*, 2000). Here, it is interesting to note that in the mouse TK, the residue corresponding to human E204 in the ^{203}KEN box is deleted (Fig. 4.1).

KEN is not the only proteolytic sequence known. Another one is PEST - sequences enriched in proline (P), glutamate (E), serine (S) and threonine (T) (Rechsteiner & Rogers, 1996). PEST sequences are often present in unstructured regions of the protein, and there appears to be a statistical preference for location of these sequences in the C-terminal regions of proteins involved in cell cycle regulation (Wright & Dyson, 1999). An algorithm PEST-FIND that produces a score from –50 to +50 and defines sequences with a score above 0 as possible PEST sequence, defines residues K^{192}ASGQPAGPDNK^{203} as a poor PEST sequence with a score of –10. Many PEST sequences are conditional and phosphorylation of serines and threonines within such PEST regions are needed to activate a latent degradation signal (Rechsteiner & Rogers, 1996). Analysis of amino acid sequence of TK1 for potential phosphorylation sites by NetPhos 2.0 (at www.cbs.dtu.dk, Center for Biological Sequence Analysis) gives six serine residues high score values (position nr/score: 13/0.996, 30/0.908, 33/0.951, 62/0.886, 63/0.979 and 183/0.986), while S^{194} and S^{231} located in the C-terminal, get very low scores: 0.056 and 0.012 respectively. This very low score of S^{231} is in agreement with results of Chang *et al.* (1998) who found no difference in phosphorylation level between wild-type TK1 and TK1^{S231A} mutant in mitotically blocked cells. Besides serine residues, also T168 and Y181 get high prediction scores as possible phosphorylation sites in the C-terminal: 0.992 and 0.952 respectively. It seems reasonable to suppose that phosphorylation of residues outside the PEST region, for

example the above-mentioned S^{13}, might as well induce conformational changes in the protein and thus expose it to proteolytic enzymes.

Another proteolytic motif - AQP, enriched in alanine, glutamine and proline, is present in EBNA, Epstein-Barr virus nuclear antigen (Rechsteiner & Rogers, 1996). It has been observed that cloned EBNA proteins are expressed at higher amounts in lymphoma cell lines if their polyproline region is deleted (Peng et al., 2000). Among the last 41 C-terminal amino acids of TK1 are 7 prolines, 6 alanines and 4 glutamines, the AQP residues together stand for 17 of the 41 amino acids. Presence of the weak PEST motif and of AQP motif in the human TK1 may facilitate the interaction with the proteolytic machinery in vivo. Especially prolines have been proposed to be involved in protein-protein binding because of 1) the flat hydrophobic surface of the proline ring which binds well to other hydrophobic surfaces, and 2) the fact that proline is a good hydrogen bond acceptor, possibly thanks to the electron-donating potential of the methylene group attached to the amide nitrogen (Williamson, 1994). The proline-rich C-terminal of TK1 contains 3 lysines, K203, K211 and K220 which could serve as sites for binding and conjugation of ubiquitin to the enzyme.

The increased in vivo stability of C-terminally truncated TK1 does not provide a simple explanation for the finding that also in vitro stability of the truncated recombinant enzyme gets increased. A plausible explanation can be offered by a hypothesis that the C-terminal of TK1 is disordered[1]. The proportion of eucaryotic proteins with disordered regions has been estimated to be as high as 52-67%, with 35% of these proteins being completely disordered, 16% displaying disorder at their C-terminal ends, 30% at N-terminal end and 17% at internal regions (Vucetic et al., 2003). Examples of enzymes having disordered, function-dispensable termini are: human topoisomerase I with disordered N-terminal domain (Stewart et al., 1996), and the tumour suppressor phosphatase, PTEN with disordered C-terminal domain which decreased the catalytic activity (Vazques et al., 2000). The disordered regions have been found involved not in catalytic, but in different regulatory functions related to their disorder: molecular recognition and signalling through binding to proteins, DNA, different RNAs, and in mediation of posttranslational modification such as phosphorylation or proteolysis (Dunker et al., 2002). Phosphorylation and proteolysis are known to be involved in regulation of lifetime and enzymatic activity

of human TK1 (Ke *et al.*, 2003; Ke & Chang, 2004). The possibility that AQP residues and the KEN box of TK1 serve as a proteolytic signal has already been mentioned. This fits with the assumption that in the disordered regions the specific residues are not masked by the folded structure, and therefore make an easily accessible target for the proteolytic enzymes. Several reports describe these disordered regions to undergo *in vitro* protease digestion orders of magnitude faster than the ordered regions (Stewart *et al.*, 1996; Fontana *et al.*, 1997).

Some amino acids, like P, A, E, S and Q occur more often in disordered proteins than in ordered proteins (so called disorder-promoting amino acids), while W, F, Y, C, I, L, V and N are depleted in the disordered regions (called order-promoting amino acids) and T, N and D are ambivalent (Dunker *et al.*, 2001; Vucetic *et al.*, 2003). P, A, E, S and Q make together 21 out of the 41 C-terminal amino acids of TK1 (Fig. 8.1). High content of proline, which is the poorest α-helix- and β-strand-forming residue, and relatively high content of hydrophilic amino acids (41% in the C-terminal of TK1) that are prone for interactions with the solvent, can be imagined to result in lack of structure as opposed to folding into a compact globular domain.

A simple tool to identify regions of disorder is newly developed GlobPlot algorithm which is based on a sum of the propensity for amino acids to be in an ordered or a disordered state (Linding *et al.*, 2003).

---

[1] There is no general definition of disorder. It can be described as the lack of regular structure and a high degree of flexibility in the polypeptide chain.
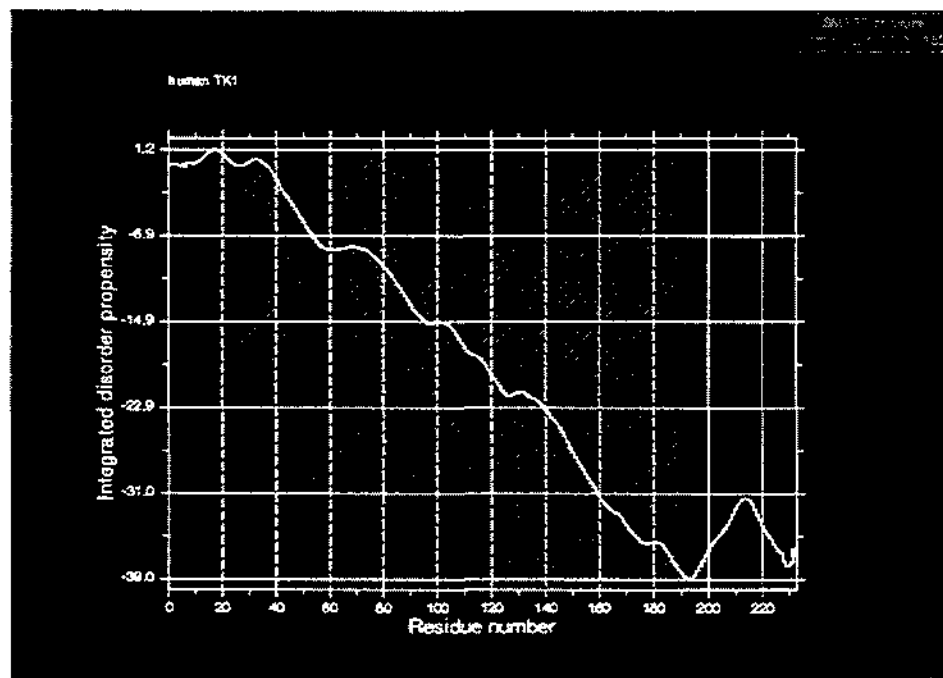
Figure 8.6. Prediction of disorder in TK1 according to GlobPlot predictor (http://globplot.embl.de). Curve going uphill indicates increased disorder/flexibility and downhill segments indicate globularity, domains or structured regions. The server sends by default the analysed sequence to SMART and Pfam databases (Bateman et al., 2002; Letunic et al., 2002), and displays any obtained homologous domains as coloured boxes (here blue) layered on the graph.

The GlobPlot predicts 11 consecutive residues ([5]NLPTVLPGSPS[16]) in the N-terminal and 20 consecutive residues in the C-terminal of TK1 ([193]ASGQPAGPDNKENCPVPGKP[212]) to be disordered. (Fig. 8.6). Altogether, 13.2% of TK1 protein is predicted to be disordered and 77.4% of the protein is predicted to be globular. For comparison, only the 10 C-terminal residues out of total 250 residues of dNK, or 4%, are predicted to be disordered, while 85.2% of the dNK sequence is predicted to be globular (http://globplot.embl.de). The increased amount of predicted disorder in TK1 might reflect the problems encountered during crystallization of this protein prior to X-ray structure determination.

Chapter 9
General discussion

The cytosolic thymidine kinase, TK1 is a cell-cycle regulated enzyme. Its activity fluctuates with DNA synthesis and is low or absent in quiescent cells, while being high in dividing cells. The control of TK1 activity is effected on transcriptional and several post-transcriptional levels, as described in Chapter 3. On the enzymatic level, native TK1 isolated from lymphocytes, has been shown to be regulated by pre-assay exposure to ATP: Incubation of the pure enzyme with ATP at 4°C induced enzyme concentration-dependent reversible transition from a dimer with low catalytic activity and non-hyperbolic thymidine substrate kinetics to a tetramer with a 20- to 30-fold higher catalytic activity and hyperbolic kinetics, but with the same maximal velocity (Munch-Petersen et al., 1993). To further investigate the effect of ATP, two expression plasmids were constructed: pET3a-TK1 which contained the amino acid coding region of TK1 cDNA from the pTK11 clone of Bradshaw & Deininger (1984) and pGEX-2T-LyTK1 which contained the amino acid coding region from normal human lymphocytes (Paper I, Berenstein et al., 2000). Surprisingly, the recombinant TK1 expressed from the pET3a-TK1 did not show the regulatory effect of ATP, which at that time was explained by the absence of post-translational modification of the enzyme in E. coli cells (Munch-Petersen et al., 1995). On the other hand, preliminary kinetic and gel filtration experiments with rLy-TK1 have shown that it behaved as native lymphocyte TK1 with respect to the regulatory effect of ATP. By comparison of the sequence of lymphocyte TK1 cDNA with that of pTK11 clone of Bradshaw & Deininger (1984) it was discovered that lymphocyte TK1 cDNA had a GTG codon for valine at amino acid position 106, while the TK1 cDNA in the pTK11 clone had a ATG codon for methionine at this position (Kristensen, 1996; Paper I, Berenstein et al., 2000).

This study was started in order to determine which amino acid is naturally present at position 106 in human TK1 and what is the effect of amino acid 106 on properties of the enzyme. Sequence analysis of altogether 22 TK1 genomic and cDNAs from healthy, leukemic or transformed cell lines revealed a GTG codon each time in both alleles. Thus, the experimental results presented in this thesis demonstrate that it is valine and not methionine that is the naturally occurring amino acid at position 106 in the human thymidine kinase. As described in Chapter 6 and Chapter 7 valine was found to be of

prominent importance for structure and enzymatic mechanism of TK1. Valine 106 is conserved among vertebrates and viruses from the pox family, and is positioned in a highly conserved area of TK1. This area, important for catalytic activity and regulation of function of TK1, is predicted to form an amphipathic helix facilitating subunit interaction, i.e. the interface (Fig. 7.9, and Eriksson et al., 2002). I have cloned, expressed, and purified the putative interface domain of rLy-TK1, rLy-TK1$^{66-136}$, with valine or methionine as amino acid 106. Investigation of oligomerization properties of rLy-TK1$^{66-136}$ confirmed the importance of amino acid 106 for the subunit arrangement of the enzyme, because in gel filtration experiments rLy-TK1$^{66-136(Val106)}$ interface fragment eluted as a mixture of a dimer and a tetramer, while rLy-TK1$^{66-136(Met106)}$ eluted as a tetramer only (Fig. 7.3).

The shift between a dimer and a tetramer, which according to the results described in this thesis, is possible for TK1 with Val-106, but not with Met-106, may have an important physiological role as proposed in a model of Munch-Petersen et al. (1995a): To correlate the thymidine kinase activity with DNA synthesis in the S phase, TK1 shifts from a low activity dimer with apparently negative cooperativity in G1, to a high activity tetramer with apparently hyperbolic reaction mechanism in S phase. (Fig. 9.1). This transition is dependent on ATP and enzyme concentration and can be considered as an additional regulation of TK1 activity, on the top of transcriptional and post-transcriptional regulation.

$K_m = 0.5\ \mu M$

$K_m = 14\ \mu M$

G1   S   G2/M   G1   S   G2/M
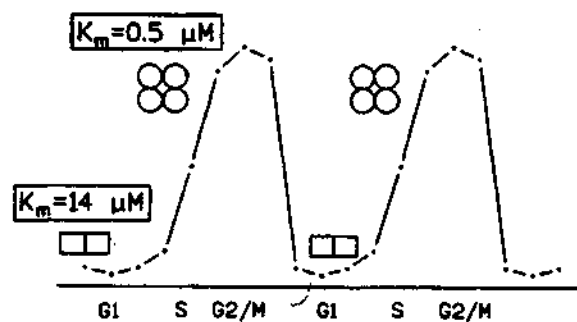
Figure 9. 1. Model for the fluctuation between the dimer and tetramer forms of thymidine kinase during the cell cycle. In the G1 phase with low TK1 concentration, the dimer TK1 dominates. With increasing concentrations during S-phase, more TK1 will be in the tetramer form. After Munch-Petersen et al. (1995a).

This model has recently been extended by Li C.L. et al. (2004) who used our recombinant TK1 with valine 106 for investigation of how serine 13 phosphorylation effects the activity of TK1. As described in Chapter 3, serine 13 has previously been shown to be the site of heavy mitotic phosphorylation (Chang et al., 1998). S13D substitution which mimics phosphorylation was found to have a perturbed pattern of ATP induced tetramerization: the equilibrium was shifted from a tetramer to a dimer and was paralleled by an 10-fold

increase in K_m (Li *et al.*, 2004). Because the concentration of TK1 remains high in G2 phase and in early mitosis (Ke & Chang, 2004) these results explain the previously observed downregulated activity of phosphorylated TK1 at G2/M phases in proliferating cells (Chang & Huang, 1993; Chang *et al.*, 1994; Chang *al.*, 1998). Fig. 9. 2 illustrates the role of phosphorylation in the regulation of TK1 activity during the cell cycle.
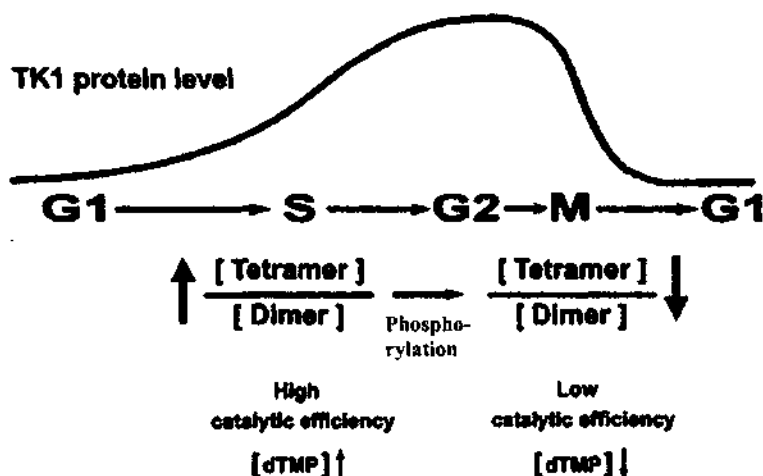
Figure 9. 2. Model of the role of phosphorylation in the quaternary structure and functional regulation of TK1 in the cell cycle. During progression from G1 to S phase expression of TK1 increases and TK1 protein is predominantly in its tetrameric form. In G2/M phase TK1 becomes phosphorylated and the oligomeric equilibrium shifts towards the dimeric form. After Li *et al.* (2004).

How to reconcile the effect of Ser-13 phosphorylation on dimer-tetramer transition with the proposal that the area around amino acid 106 in TK1 is the monomer-monomer and/or the dimer-dimer interface? A possible answer might be that formation of the tetramer from a pair of dimers requires the Ser-13s to be very close in space and, if phosphorylated, the negative charges repulse each other. In any case, phosphorylation of TK1 is involved in transition between a low activity dimer with apparently negative cooperativity and a high activity tetramer with apparently hyperbolic reaction mechanism, and therefore phosphorylation regulates TK1 activity as well as biosynthesis of dTTP. Recently, a H121N mutation in human mitochondrial TK2 that caused the loss of negative cooperativity with thymidine as substrate in the recombinant TK2, has been described (Wang *et al.*, 2003). The mutation was found in patients with the mitochondrial DNA depletion syndrome combined with fatal functional defects in muscles (Saada *et al.*, 2001). This finding underlined the importance of proper regulation mechanism of the enzymes involved in salvage of deoxynucleosides for integrity of the organism during lifetime.

It remains to be explained why methionine at position 106 was found in previous investigations of TK1 derived from SV40 transformed fibroblasts (Bradshaw & Deininger, 1984) and from HeLa cells (Flemington *et al.*, 1987). Because of the prominent difference between the Val-106 and the Met-106 forms of TK1 for the thymidine affinity and dimer-tetramer transition of human TK1, the earlier research performed with recombinant TK1 expressed from the pTK11 clone, *i.e.* with methionine at position 106, might not have yielded representative data. We have found several polymorphisms in the coding sequence of the TK1 gene, but none of them in the codon for amino acid 106. Our results were supported by sequencing results performed on TK1 gene from the human lymphoblastoid cell line, TK6 (Grosovsky *et al.*, 1993), from Burkitt lymphoma lymphocytes and from uterus tissue deposited in Genbank in 2001 by Strausberg for the NIH Mammalian Gene Collection (MGC) Project (gi:23503074), and recently in breast and colorectal cancer solid tumours (Gilles *et al.*, 2003). Most of the polymorphisms observed occurred in one allele only, and at the third base position in the codon, so they did not result in amino acid change. Besides the change from lysine to arginine at amino acid position 211, found in lymphocytes from the very first donor, only one other polymorphism resulted in amino acid change: a C to T transition in the cell line Reh changed the CAG codon for amino acid 20 in one allele to TAG, a stop codon. Thus, only a truncated TK1 polypeptide of 19 amino acids will be expressed from this allele, the cells will depend on TK1 synthesis from the other allele, and may become more vulnerable to DNA damage. Previous observation of Kristensen *et al.* (1994) in our laboratory, that in chronic lymphatic leukaemia lymphocytes, high level of TK1 mRNA is present without accompanying TK1 enzyme activity, may be explained by presence of a stop codon in one, or more probably, both alleles at the *tk* locus.

"Silent" polymorphisms that do not introduce any amino acid change may still be important for the expression and function of TK1 in the living cell. Silent mutations may change a "high usage codon" to a "low usage codon", which will result in less efficient translation of the gene. Kim *et al.* (1997) have shown that the human erythropoietin (EPO) gene with human high-frequency codons gave a considerably higher level of expression in mammalian cells than the gene with yeast high frequency codons. Opposite situation, where a "low usage codon" will be changed to a "high usage codon" with resulting higher expression of TK1, may be also detrimental to the cell. An example of importance of silent mutations for gene function comes from investigations on polymorphisms in XPD excision repair gene: psoriasis patients had a higher risk of early basal cell carcinoma if they were

AA homozygotes at a specific nucleotide position, than patients with AC or CC genotypes (Dybdahl *et al.*, 1999).

Val-106 might have been selected over Met-106 during natural evolution due to the physiological advantage of dimer – tetramer transition of TK1 protein, and may be also because of the higher thermal stability of the Val-106 enzyme. Though, the lower stability of TK1 (V106M) *in vitro* might be of no relevance *in vivo*. At enzyme assay conditions, where ATP/MgCl$_2$ (2.5 mM), Chaps (0.5 mM) and BSA at 3 mg/ml are present, both Val-106 and Met-106 enzymes are stable (Munch-Petersen, personal communication). With the total protein concentration of 200 - 300 mg/ml (Ellis, 2001), 2 mM ATP (Ferraro *et al.*, 1992) and approximately 1 μM thymidine (Holden *et al.*, 1980) in the cell cytoplasm, TK1 (V106M) might have the same stability as V106WT. The molecular crowding[1] of the living cell might bring still another regulatory mechanism of TK1: The putative flexible C-terminal of TK1 can be imagined to be occupied by/interact with other proteins, and in such case the catalytic activity may correspond to the higher catalytic activity of TK1Δ40.

The molecular mechanism behind the effect of valine/methionine substitution is not clear and must await the 3D structure of TK1 to be solved. Obtaining a monomeric and active TK1, preferably with full or nearly full activity, would have greatly facilitated structural studies of the molecule. The fact that no monomeric TK1 mutants have so far been created, in spite of several attempts, could be due to the fact that dimerization of TK1 subunits is essential for catalysis: Both subunits might be needed to keep the catalytic pocket in shape, or it might be the dimerization that brings the substrate binding site and the catalytic centre together. For example, residues from both subunits of thymidylate synthase contribute to each active site (Hardy *et al.*, 1987). The second scenario was suggested for Rnase T: the substrate-binding site and the catalytic centre of the same subunit are believed to be located far away from each other, but the centres from different subunits are brought together through homodimerization to form fully functional active sites (Zuo & Deutscher, 2002).

---

[1] When macromolecular species occupy a large fraction of the total volume in a given medium, the medium is referred to as "crowded" rather than "concentrated", as no individual macromolecular species may be present at high concentration (Minton, 2000).

TK1 should in principle be functional as a monomer, because each subunit contains the necessary information for substrate binding and catalytic activity. Based on structure designed, for example, by Swiss PDB Viewer program (also called "Deep View"), amino acids at chosen positions of the putative dimer interface could be substituted to introduce electrostatic repulsion into the monomer-monomer interface, but without destabilization of the monomer 3D structure. This strategy has been used successfully for creating several biologically active, monomeric insulins[2], where negative charges were inserted to oppose negative charges on the complementary interface, as *e.g.* in the B9 Ser to Asp mutant (Brange *et al.*, 1988).

If monomers of TK1 turn out to be genuinely devoid of enzymatic activity, drugs that inhibit dimerization of TK1 could be designed for possible use as an aid in anticancer therapy.

In conclusion:

Two regions of human recombinant TK1 have been the subject for structure-function studies in the present investigation: 1) the internal region surrounding amino acid 106 while stressing the role of the residue 106, and 2) the C-terminal stretch of the last 40 amino acids.

We have shown that valine is the naturally occurring amino acid at position 106 in the human thymidine kinase, that amino acid 106 and the sequence surrounding it are associated with reversible dimer-tetramer transition, and that the β-branching of amino acid at site 106 may be important for such conformational transitions.

Investigation of the C-terminally truncated TK1 (TK1Δ40) revealed that it is profoundly more stable *in vitro* than the non-truncated TK1. I propose that disorder (the lack of regular structure) and flexibility in the stretch of the last 40 amino acids can explain the higher *in vitro* stability.

---

[2] The 3D structure of insulin has been known since 1969.

Summary

Thymidine kinase 1 catalyses phosphorylation of thymidine to thymidylate (dTMP) using ATP as the phosphate donor. Due to the ubiquitous nucleotide kinases, intracellular dTMP is quickly phosphorylated to dTTP. dTTP in turn, is an important allosteric effector of ribonucleotide reductase, and a cell with defect TK1 activity has diminished ability to balance its supply of both purine and pyrimidine DNA precursors. TK1 is a cell cycle regulated enzyme, its activity parallels DNA synthesis rate, and its expression is carefully regulated on transcriptional, post-transcriptional, translational and post-translational levels. The co-substrate ATP has been shown to be a regulator of the catalytic activity of TK1, as well.

Because markedly different enzymatic properties were observed between endogenous TK1 isolated from human lymphocytes and a recombinant enzyme expressed from a TK1 cDNA clone present in the NCBI GenBank database (where mRNA from transformed human fibroblasts was used as cDNA source), and because DNA sequencing of lymphocyte TK1 cDNA showed a GTG codon for valine at amino acid position 106, and not ATG codon for methionine as in the GenBank database, we suspected the GenBank sequence to be wrong. Further, we wanted to investigate if amino acid 106 had any effect on the enzymatic properties. Therefore, the investigations described in the experimental part of this Ph.D. thesis were undertaken.

The naturally occurring amino acid at position 106 of human thymidine kinase 1 was identified by sequencing 22 independent isolations of genomic DNA and cDNA from healthy donors, leukemic cell lines, colon cancer cell lines, and transformed and nontransformed fibroblast cell lines. In each case a GTG codon for valine was found at this position in both alleles. Because amino acid 106 is located in a highly conserved area of the protein and is valine in other mammalian and pox-viral thymidine kinases, valine is most probably the naturally occurring amino acid at this position in human TK1.

The effect of valine or methionine at position 106 on the properties of the enzyme has been studied with the recombinant TK1 containing Val-106 (V106WT) or Met-106 (V106M). TK1(V106WT) and TK1(V106M) were cloned, expressed, purified and characterized with respect to enzymatic and oligomerization properties. The properties of recombinant TK1 (V106WT) were the same as those of the native TK1 isolated from human lymphocytes: pre-assay ATP exposure induced an enzyme concentration-dependent reversible transition from a dimer to tetramer with about 30-fold higher catalytic activity ($K_{0.5}$ towards thymidine was about 15 and 0.5 µM, respectively, V_{max} was the same). Substitution of Val-106 with methionine gave TK1(V106M), which was a permanent tetramer with high catalytic activity even without pre-exposure to ATP ($K_{0.5}$ towards thymidine was about 0.5 µM for both the ATP pre-exposed and the ATP unexposed enzyme forms). Furthermore, TK1(V106M) was considerably less stable than TK1(V106WT); $t_{1/2}$ values at 15°C were respectively about 40 and 390 minutes.

The role of size, conformation and polarity of amino acid 106 for the function and structure of human TK1 was examined by creation of nine recombinant mutant enzymes at amino acid site 106 and their characterization with regard to native size, kinetic properties and stability. Based on the obtained results the recombinant mutant enzymes could be divided into two groups: Group I with V106A, V106I and V106T behaved like V106WT, i.e. pre-assay exposure to ATP induced reversible transition from a dimer with low catalytic activity ($K_{0.5}$ towards thymidine from about 13 to 43 µM) to a tetramer with high catalytic activity ($K_{0.5}$ towards thymidine from about 0.3 to 0.9 µM). Group II with V106G, V106H, V106K, V106L and V106Q behaved like V106M in that they were permanently high activity tetramers ($K_{0.5}$ towards thymidine from 0.3 to 1.2 µM, both in the absence and presence of ATP). The stability of enzymes from Group II was much lower than the stability of Group I enzymes. The amino acids at site 106 in Group I enzymes are, except for alanine, of similar size and conformation, but different polarities, and the hydroxyl moiety of threonine does not seem to cause any disturbances. The common property of the side chains of valine, isoleucine and threonine is branching at the β-carbon atom considered to destabilize α-helices because of steric clashes. Destabilization of helix-helix interaction by β-branching amino acids adds conformational flexibility to the enzyme

protein and might be crucial for dimer-tetramer transition of TK1. The dimer-tetramer transition seems to play an important physiological role in the regulation of TK1 activity.

The role of amino acid 106 and the sequence surrounding it for dimer-tetramer transition was examined by cloning, expression and characterization of the putative interface fragment $TK1^{66-136(Val106)}$ and of its mutated form $TK1^{66-136(Met106)}$. The results from gel filtration experiments confirmed that amino acid 106 and the neighbouring residues are involved in dimer-tetramer transition, as the $TK1^{66-136(Val106)}$ fragment eluted as a mixture of a dimer and a tetramer, whereas $TK1^{66-136(Met106)}$ fragment eluted as a tetramer only.

The C-terminal of TK1 is known to have a regulatory role *in vivo* for cell cycle-regulated expression of this enzyme. In order to uncover the role of the last 40 C-terminal residues of TK1 for the kinetic, oligomerization and stability properties, the TK1Δ40 recombinant mutant enzyme was cloned, expressed and characterized *in vitro*. The properties of recombinant TK1Δ40 were similar to the native TK1 isolated from human lymphocytes and to the recombinant TK1(V106WT): pre-assay ATP exposure induced an enzyme concentration-dependent reversible transition from a dimer to tetramer with higher catalytic activity ($K_{0.5}$ towards thymidine was about 2.5 and 0.5 μM, respectively). The lower $K_{0.5}$ towards thymidine for the ATP unexposed TK1Δ40 as compared to the ATP unexposed TK1(V106WT) may be explained by a C-terminal limiting access of the substrate to the active site. The *in vitro* stability of TK1Δ40 was found to be considerably higher than the stability of TK1(V106WT), with $t_{1/2}$ values of about 43 hours and 6.5 hours respectively. It is very likely that this higher stability of TK1Δ40 is due to the absence of a disordered stretch of the C-terminal amino acids.

166

Dansk resumé

Thymidin kinase 1 katalyserer fosforylering af thymidin til thymidin monofosfat (dTMP) ved hjælp af ATP som fosfat donor. Intracellulær dTMP fosforyleres hurtigt videre til dTTP af allestedsnærværende nukleotid kinaser. dTTP har en nøgleposition i reguleringen af ribonukleotid reduktase, som er et vigtig enzym i både purin og pyrimidin DNA-prekursor metabolismen. En celle med en defekt thymidine kinase er derfor dårligere end den normale celle til at afbalancere sin dNTP-pool, og har større sandsynlighed for forstyrrelser under DNA syntesen og DNA repair.

Det eksperimentelle arbejde i denne afhandling udspringer fra den observation at TK1 isoleret fra humane lymfocytter har andre enzymatiske egenskaber end den rekombinante TK1, der er udtrykt fra en TK1 cDNA klon baseret på mRNA fra transformerede humane fibroblaster. Klonens DNA sekvens er den officielle TK1 sekvens som findes i NCBI GenBank databasen. Da DNA sekventering af en TK1 klonet fra lymfocyt DNA blev udført i Munch-Petersens laboratorium, fandtes overraskende en GTG kodon for valin i site 106, og ikke ATG kodon for methionin som i NCBI databasen. Arbejdet i denne afhandling handler om at klarlægge hvilken aminosyre der er på site 106 i human TK1, samt belyse denne aminosyres rolle for de enzymatiske egenskaber.

Den naturligt forekommende aminosyre i site 106 i den humane TK1 er blevet identificeret ved at sekventere 22 uafhængigt isolerede genomiske DNA og cDNA fra raske donorer, leukæmiske cellelinier, koloncancer cellelinier, samt transformerede og ikke transformerede fibroblast cellelinier. Resultaterne viste at kodon for valin forekom i site 106 på begge alleler i samtlige undersøgte donorer og cellelinier. Da site 106 er placeret i proteinets højt konserveret område, og er valin i andre mammale og pox-virale thymidin kinaser, er det mest sandsynligt at valin er den naturligt forekommende aminosyre i site 106 i den humane TK1.

Betydningen af valin og methionin i site 106 for egenskaberne af TK1 er blevet undersøgt med det rekombinante TK1 enzym. TK1 indeholdende Val-106 (V106WT) og Met-106 (V106M) blev klonet, udtrykt, oprenset og karakteriseret med hensyn til deres enzymatiske

egenskaber og den native molekylevægt (kvaternære struktur). Det blev påvist at TK1 (V106WT) opførte sig som det native TK1 enzym oprenset fra humane lymfocytter - uden tilstedeværelse af ATP var TK1(V106WT) en dimer og inkubationen med ATP inducerede en enzymkoncentrationsafhængig reversibel tetramerisering af denne dimer. Tetramerformens katalytiske aktivitet var omkring 30 gange højere end dimer-formens: $K_{0.5}$ værdier overfor thymidin var henholdsvis omkring 0.5 µM og 15 µM, mens V_{max} værdien var den samme. Ved at mutere Val-106 til methionin blev TK1(V106M) fremstillet, og enzymet viste sig at være en permanent tetramer med høj katalytisk aktivitet ($K_{0.5}$ omkring 0.5 µM uafhængigt af ATP-inkubationen inden den enzymatiske assay). Ydermere var TK1(V106M) væsentlig mindre stabil end TK1(V106WT); $t_{1/2}$ værdier var henholdsvis 40 og 390 minutter.

For yderligere at undersøge betydningen af størrelsen, konformationen og polariteten af aminosyre 106 i den humane TK1, blev ni rekombinante site 106-mutant-enzymer fremstillet og karakteriseret med hensyn til enzymatiske egenskaber, nativ størrelse og stabilitet. Udfra de enzymatiske egenskaber kunne de rekombinante enzymer inddeles i to grupper: Gruppe I med TK1(V106WT) lignende egenskaber, og gruppe II med TK1(V106M) lignende egenskaber. Gruppe I indeholdt V106A, V106I og V106T. Ligesom TK1 (V106WT), var de overvejende i dimer-form uden tilstedeværelse af ATP, mens inkubationen med ATP inducerede reversibel tetramerisering med en samtidig overgang fra en form med lav katalytisk aktivitet ($K_{0.5}$ for thymidin fra omkring 13 til 43 µM) til en form med høj katalytisk aktivitet ($K_{0.5}$ fra omkring 0.3 til 0.9 µM). Gruppe II indeholdt V106G, V106H, V106K, V106L og V106Q. Gruppe II enzymer var permanente tetramerer med høj katalytisk aktivitet ($K_{0.5}$ fra omkring 0.3 til 1.2 µM, uafhængigt af forudgående inkubationsperiode med ATP). Stabiliteten af Gruppe II enzymer var meget lavere end Gruppe I enzymer. Aminosyrer i site 106 i Gruppe I enzymer, med undtagelsen af alanin, ligner hinanden med hensyn til størrelse og konformation, men har forskellig polaritet. Den polære hydroxyl kæde på threonin adskiller sig fra de hydrofobe sidekæder på valin og isoleucin, dog har disse tre aminosyrer en fælles egenskab: en forgrening på β-carbonatomet som vides at destabilisere α-helix strukturen og helix-helix interaktioner på grund af steriske forstyrrelser. Denne virkning af aminosyrer med β-carbonatomets forgrenede sidekæder kan gøre konformationen af enzymproteinet mere fleksibel, og kan

være afgørende for dimer-tetramer overgangen hos TK1. Den reversible dimer-tetramer overgang formodes at spille en vigtig fysiologisk rolle under reguleringen af TK1 aktivitet. Betydningen af aminosyre 106 og regionen omkring site 106 for dimer-tetramer overgangen blev undersøgt ved at klone, udtrykke og karakterisere det formodede interfacefragment TK1$^{66\text{-}136(\text{Val}106)}$ og dets mutant TK1$^{66\text{-}136(\text{Met}106)}$. Det fremgik udfra den native molekylevægtsbestemmelse ved gelfiltrering, at TK1$^{66\text{-}136(\text{Val}106)}$ optræder som en blanding af dimer og tetramer, mens TK1$^{66\text{-}136(\text{Met}106)}$ optræder udelukkende som tetramer. Resultatet bekræftede at aminosyre 106 og den omkringliggende region er vigtig for dimer-tetramer overgangen hos TK1.

Et andet område af TK1 som jeg undersøgte var C-terminalen. TK1 er et celle-cyklus reguleret enzym og mange rapporter har peget på at de C-terminale aminosyrer har en vigtig rolle for den mitotiske nedbrydning af TK1 *in vivo*. For at afdække betydningen af de sidste 40 C-terminale aminosyrer for egenskaber som stabilitet, nativ molekylestørrelse og enzym kinetik, blev TK1Δ40 rekombinant enzym (de sidste 40 aminosyrer er fjernet) karakteriseret *in vitro*. TK1Δ40 opførte sig som den native lymfocyt TK1 og den rekombinante TK1(V106WT) med hensyn til den reversible effekt af ATP inkubation forud for den enzymatiske assay: ATP inducerede en reversibel overgang fra en dimer-form til en tetramer-form med højere katalytisk aktivitet ($K_{0.5}$ overfor thymidin henholdsvis omkring 2.5 og 0.5 μM). Den lavere $K_{0.5}$ for den ikke ATP-eksponerede form af TK1Δ40 end for den ikke eksponerede TK1(V106WT) kan forklares ved at de C-terminale aminosyrer forhindrer substratet i at komme til enzymets aktive site. *In vitro* stabilitet af TK1Δ40 var væsentlig højere end stabiliteten af TK1(V106WT), med $t_{1/2}$ værdier på henholdsvis 43 timer og 6.5 timer. Det diskuteres i denne afhandling hvorvidt forskellen skyldes den formodede fleksibilitet og manglen på den ordnede struktur i det C-terminale område.

Appendix 1

Introduction to NMR principles

(Protein NMR without tears and physical principles)

In principle, all ^{1}H atoms in a protein can be observed by NMR. The basics of this technique as applied to protein structure can be best understood if one regards a single amino acid. Fig. Ap. 1A and Ap. 1B show a 1D and a 2D COSY spectrum of valine:

CH$_3$
H — C — CH$_3$
$^+$H$_3$N — C — COO$^-$
H
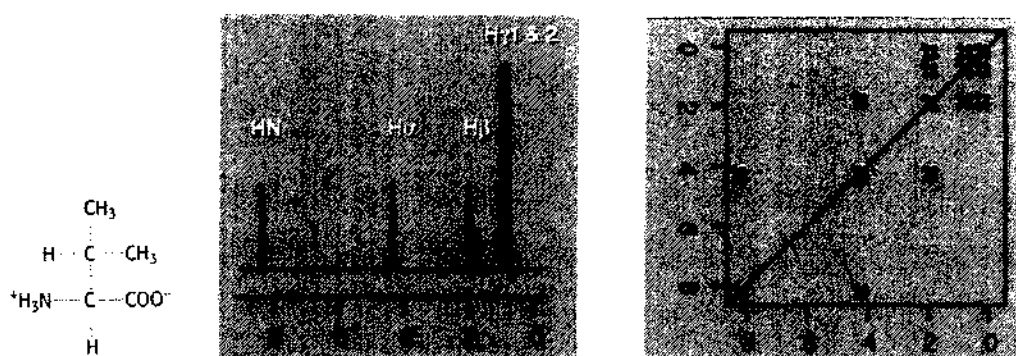
Figure Ap. 1A. 1D spectrum. of valine.
From Poulsen *et al.* (1997)

Figure Ap. 1B. 2D COSY spectrum of valine.

Fig. Ap. 1A shows how the different hydrogen atoms of valine can be resolved depending on the atom to which the ^{1}H is bonded (α, β and γ refer to the carbon atom of valine). The resolution is due to different resonance positions of the H atoms and is given by the *chemical shift* (ppm, *i.e.* in parts per million relative to a reference substance). The chemical shift is determined primarily by the chemical group in which the ^{1}H atom is found, and also in a smaller degree on environment of this group. For example, in an ideal random polypeptide coil all CH$_3$ group resonate at about 1ppm, NH groups at about 8 ppm, and multiple copies of a specific amino acid have identical chemical shifts and cannot be distinguished. In globular proteins chemical shift dispersion arises because interior peptide segments are shielded from the solvent and are nearest neighbours to other protein segments (Wüthrich, 1986). Still, less than 10% of the resonance lines are shifted from their random coil class into the chemical shift ranges of other classes of protons (Wüthrich, 1986).

When a second frequency axis is added to the 1D method, one gets a 2D NMR experiment - the resonances are spread in two dimensions and the spectral analysis is facilitated. In 2D (Fig. Ap. 1B) experiments the diagonal corresponds to the 1D spectrum, while the off-

diagonal peaks (so called cross-peaks) contain information about connections between resonances on the diagonal (Roberts, 1993). The connections observed depend on the kind of experiment that is being carried out and on the machine setting: in COSY (correlated spectroscopy, Fig. Ap. 1B) one can observe through-bond connections between resonances of nuclei separated by one, two or three up to four covalent bonds, in NOESY (nuclear Overhauser effect spectroscopy) one can observe through-space connections between the resonances of nuclei which are close together (*i.e.* below 5 Å) in space (Roberts, 1993).

All Val residues, for example, can be identified, but not distinguished from each other in COSY spectra and remain isolated entities, because COSY has no observable spin-spin coupling across the peptide bond which connects neighbouring residues (Redfield, 1993). These assignments are deduced from the through- space ^{1}H-^{1}H interactions of the nuclear Overhauser effect (NOE). Each cross-peak can be attributed to a distance between a pair of H atoms, irrespective if protons at distances up to 5 Å originate from neighbouring residues or from residues that are far apart in the amino acid sequence (Billeter *et al.*, 1982). The well-resolved resonances (ideally as in Fig. Ap.2) can be assigned to specific amino acids in the sequence.
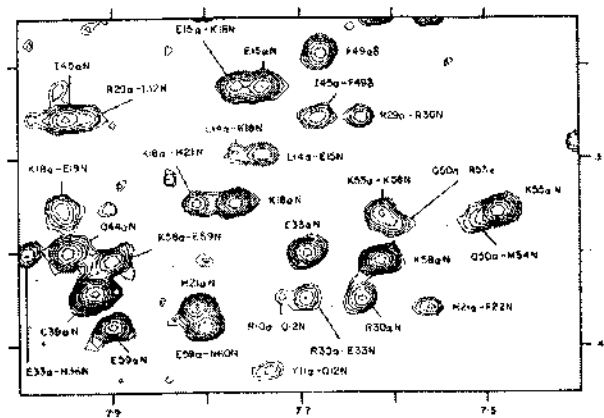
Figure Ap. 2. A small region from a ^{1}H-NOESY spectrum of a protein. The cross-peaks are identified by the amino acid position and the proton type (from Billeter *et al.*, 1990).

The sequential assignment procedure was developed by Wüthrich and collaborators (Wüthrich *et al.*, 1982; Billeter *et al.*, 1982) and consists of two steps. The first step utilizes COSY and involves identification of spin-spin coupled resonances that belong to a particular amino acid. Alanine residues for example are easily identified because Ala is the only amino acid whose α-proton and the β-methyl group make up a spin system that has a typical spin-coupling pattern. The second step utilizes NOESY to identify through-space

connections (d) between for example H_α, H_β and amide proton of one residue to the amide proton of the next residue ($d_\alpha N$, $d_\beta N$ and dNN, respectively, se Fig. Ap. 3). In the sequential assignment strategy unique segments of two or three amino acids are identified by NMR and then by comparison with the known amino acid sequence attributed to discrete positions in the protein.
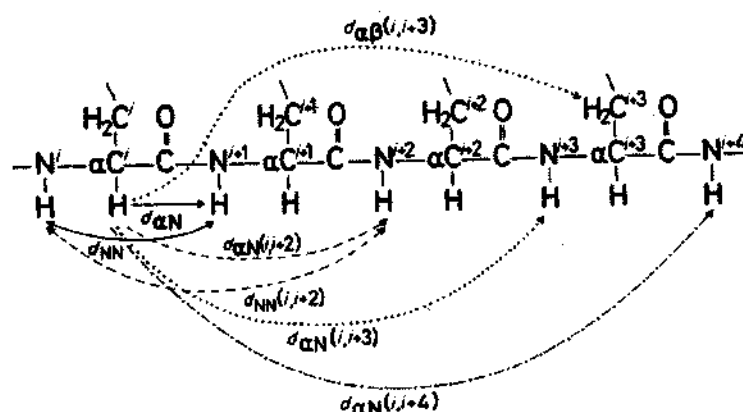
Figure Ap. 3. ^{1}H-^{1}H NOESY connectivities that are useful for identification of secondary structure (from Wüthrich et al., 1984).

For example if a NOESY cross-peak corresponding to $d_\alpha N(i,i$ +1) or $d_{NN}(i,i$ +1) indicates a dipeptide Val-Ala and this sequence occurs only once , the assignment problem is solved. If multiple sites containing Val-Ala are present, a tripeptide containing Val-Ala must be identified by NMR and then matched to the known sequence (Wüthrich, 1995). At the same time the sequential and medium range (up to five consecutive residues) NOE patterns give information about secondary structure (Fig. Ap. 3). In β-strands strong $d_\alpha N(i,i$ +1) NOEs are observed, while in helical structures strong $d_{NN}(i,i+1)$ are observed. Also $d_{\alpha\beta}(i,i+3)$ and $d_\alpha N(i,i$ +4) NOEs shorter than 4.5 Å are characteristic for helical conformation (Wüthrich et al., 1984). Following sequence-specific resonance assignment, the NOE peak intensities[1] are quantified and converted into distances and into backbone Φ and side chain χ_1 dihedral angles. In the next step structure calculations are performed to obtain atomic models consistent with the NMR-derived constrains, and finally, the structures are improved taking energy considerations into account (Kaptein et al., 1991; Wüthrich, 1995).

---

[1] The intensity of an NOE is related to r-6 where r is the distance between the interacting nuclei.

Abbreviations and glossary

Enzymes

ADA, adenosine deaminase
ADK, adenylate kinase
CAT, chloramphenicol acetyltransferase, often used as a reporter gene
CDK, cyclin dependent kinase
dCK, deoxycytidine kinase
dGK, deoxyguanosine kinase
dNK, (=*Dm*-dNK) the multisubstrate nucleoside kinase from *Drosophila melanogaster*
dNT, deoxynucleotidase
dUTPase, dUTP pyrophosphatase, produces the dTTP precursor dUMP and decreases cellular dUTP levels
GST, glutathione-S-transferase
HSV TK, Herpes simplex virus thymidine kinase
PNP, purine nucleoside phosphorylase
TK, thymidine kinase
TK1, the cell cycle regulated, cytoplasmic thymidine kinase
TK2, the constitutive, mitochondrial thymidine kinase
TP, thymidine phosphorylase, also called pyrimidine phosphorylase

Deoxyribonucleosides and deoxyribonucleotides

dAdo, dAMP, dADP, dATP, deoxyadenosine, mono-, di-, and triphosphate
dCyd, dCMP, dCDP, dCTP, deoxycytine, mono-, di-, and triphosphate
dGuo, dGMP, dGDP, dGTP, deoxyguanosine, mono-, di-, and triphosphate
dIno, dIMP, dIDP, dITP, deoxyinosine, mono-, di-, and triphosphate
dN, dNMP, dNDP, dNTP, any deoxyribonucleoside and its mono-, di-, and triphosphate
dThd, dTMP, dTDP, dTTP, deoxythymidine, mono-, di-, and triphosphate, the same as respectively **Thd, TMP, TDP** and **TTP**
dUrd, dUMP, dUDP, dUTP, deoxyuridine, mono-, di-, and triphosphate

Nucleoside analogues

AraA, 9-β-D-arabinofuranosyladenine
AraC, 1- β-D-arabinofuranosylcytosine
AraG, 9-β-D-arabinofuranosylguanine
AraT, 1-β-D-arabinofuranosyl-thymidine
AZT, 3'-azido-2',3'-dideoxythymidine
BvdU, E-5-bromovinyl-2'-deoxyuridine
CdA, 2-chloro-2'-deoxyadenosine
ddC, 2'-3'-dideoxycytidine
FdUrd, 5-Fluoro-2'-deoxyuridine
FIAU, 1-(2'-deoxy-2'fluoro 1-β-D-arabinofuranosyl)-5-iodouracil
FLT, 3'-fluoro-2',3'-dideoxythymidine

Other abbreviations

3D, three dimensional

AP-1, activating protein 1, a collective term referring to homo- or heterodimeric transcription factors composed of Jun, Fos or ATF (activating transcription factor) subunits that bind to a common DNA site, the AP-1 binding site TGACTCA

APC, anaphase promoting complex (cyclosome), a ubiquitin-conjugation protein complex

ARF, alternative reading frame

BLAST, Basic Local Alignment Search Tool

bZIP, basic region leucine zipper proteins, a family of transcription factors that bind to palindromic AP-1 regulatory elements and activate many genes involved in cell growth and differentiation. The bZIP domain consists of a basic region, contacting DNA, and a homo- or heterodimer-forming region called the leucine zipper.

CBP, CREB-binding protein, together with another co-activator - p300 promotes transcription by acetylating histones and recruiting basal transcription factors TFIIB, TBP (TATA-binding protein), and RNA polymerase II holoenzyme.

CDNB, 1-chloro-2,4-dinitrobenzene

CDP, CCAAT displacement protein

CHAPS, 3-(3-cholamidopropyl)-dimethylammoniol-1-propylsulfonate, a detergent

CKI, CDK inhibitor protein

CLL, chronic lymphatic (B-cell) leukaemia

COSY, NMR correlation spectroscopy

CREB, cAMP-response element-binding protein

DTT, dithiothreitol

E-box, a cis-element present in the promoter and enhancer regions, has sequence CANNTG and is recognized by a variety of helix-loop-helix transcription factors, eg. USF (upstream stimulatory factor)

E2F, a family of cellular transcription factors

EC, Enzyme Commission, appointed by the International Union of Biochemistry, works out systematic ways of classification and naming of enzymes, first report appeared in 1964

EGF, epidermal growth factor

h, human in for example hTK1, human TK1

hnRNA, heterogeneous nuclear RNA, a collective term referring to pre-mRNA and other nuclear RNAs of various sizes

hsp, heat-shock protein

HSV, Herpes simplex virus

IGF-1, insulin-like growth factor-1

IPTG, isopropyl-1-thio-β-D-galactopyranoside

G0 cells, quiescent, non-proliferating cells withdrawn from the cell cycle

G1 phase, The period in the cell cycle where cell prepares for DNA synthesis

G2 phase, the period in the cell cycle where cell prepares for mitosis

K_m, the Michaelis constant (= substrate concentration at ½ V_{max})

M-phase, mitosis phase

MDS, mitochondrial DNA depletion syndrome

MNGIE, mitochondrial neurogastrointestinal encephalomyopathy, a severe mitochondrial dysfunction

mtDNA, mitochondrial DNA

n, Hill constant

NCBI, National Center for Biotechnology Information, established in 1988 in USA as a national resource for molecular biology information

NMR, Nuclear magnetic resonance

NOE, nuclear Overhauser effect

NOESY, nuclear Overhauser enhanced spectroscopy

p53, tumor suppressor protein, has apparent molecular weight of 53 kDa. p53 gene is mutated in more than 50% of human cancers

PDGF, platelet-derived growth factor

PHA, phytohemagglutinin

PMSF, phenylmethylsulfonyl fluoride

PRPP, 5-Phosphoribosyl-1-pyrophosphate

Run-on assay (nascent chain assay) for transcription rate of a gene, isolated nuclei are briefly incubated with ^{32}P-labeled NTPs, the labelled RNA is hybridised to the specific gene, and radioactivity in total RNA and in this specific RNA is counted Thus, a fraction of total RNA produced from that gene can be measured

Rb, Retinoblastoma protein, tumor suppressor

SD, standard deviation

SE, standard error

Set, first fraction collected during protein purification on an affinity column. The volume of set fraction corresponds to the volume of the input which was applied to the column

S-phase, The period of the cell cycle where DNA is duplicated

SCF, ubiquitin ligase complex (from the name of the major compound proteins: Skp1-cullin/Cdc53-Fbox

SIC, S-phase cyclins inhibitor in budding yeast, corresponds to inhibitor p27 in mammalian cells

T lymphocytes (T cells), interact with infected host cells through receptors on T-cell surface (killer T cells), or interact with macrophages and produce interleukines that stimulate B and T cells to proliferate

UTR, untranslated region, noncoding regions at the ends of the primary RNA transcript

V_{max}, the maximal velocity

References

Abele, U. & Schulz, G. E. (1995). High-resolution Structures of Adenylate Kinase from Yeast Ligated with Inhibitor Ap$_5$A, Showing the Pathway of Phosphoryl Transfer. *Protein Science* **4**, 1262-1271.

Ackland, S. & Peters, G. J. (1999). Thymidine Phosphorylase: Its Role in Sensitivity and Resistance to Anticancer Drugs. *Drug Resistance Updates* **2**, 205-214.

Alber, F., Kuonen, O., Scapozza, L., Folkers, G. & Carloni, P. (1998). Density Functional Studies on Herpes Simplex Virus Type 1 Thymidine Kinase-Substrate Interactions: The Role of Tyr-172 and Met-128 in Thymine Fixation. *Proteins: Struct. Funct. Genet.* **31**, 453-459.

Allan, A. L., Albanese, C., Pestell, R. G. & LaMarre, J. (2001). Activating Transcription Factor 3 Induces DNA Synthesis and Expression of Cyclin D1 in Hepatocytes. *J. Biol. Chem.* **276**, 27272-27280.

Altschul, S. E., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic Local Alignment Search Tool. *J. Mol. Biol.* **215**, 403-410.

Altschul, S. E., Madden, T. L., Schaffer, A. A., Zhang, J., Miller, W. & Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **25**, 3389-3402.

Anderson, M. M., Chen, J., Cole, C. N. & Conrad, S. E. (1996). Activation of the Human Thymidine Kinase (TK) Promoter by Simian Virus 40 Large T Antigen Requires both the T Antigen pRB Family-Binding Domain and TK Promoter Sequences Resembling E2F-Binding Sites. *J. Virol.* **70**, 6304-6313.

Anfinsen, C. B. (1973). Principles that Govern the Folding of Protein Chains. *Science* **181**, 223-230.

Archakov, A. I., Govorun, V. M., Dubanov, A. V., Ivanov, Y. D., Veselovsky, A. V., Lewi, P. & Janssen, P. (2003). Protein-Protein Interaction as a Target for Drugs in Proteomics. *Proteomics* **3**, 380-391.

Arcot, S. S., Flemington, E. K. & Deininger, P. L. (1989). The Human Thymidine Kinase Gene Promoter. Deletion Analysis and Specific Protein Binding. *J. Biol. Chem.* **264**, 2343-2349.

Argos, P. (1988). An investigation of Protein Subunit and Domain Interfaces. *Protein Enginneering* **2**, 101-113.

Arnér, E. S. J., Spasokoukotskaja, T. & Eriksson, S. (1992). Selective Assays forThymidine Kinase 1 and 2 and Deoxycytidine Kinase and Their Activities in ExtractsFrom Human Cells and Tissues. *Biochem. Biophys. Res. Commun.* **188**, 712-718.

Arnér, E. S. J. & Eriksson, S. (1995). Mammalian Deoxyribonucleoside Kinases. *Pharmac. Ther.* **67**, 155-186.

Arpaia, E., Benveniste, P., Di Cristofano, A., Gu, Y., Dalal, I., Kelly, S., Hershfield, M.,Pandolfi, P. P., Roifman, C. M. & Cohen, A. (2000). Mitochondrial Basis for Immune Deficiency: Evidence from Purine Nucleoside Phosphorylase-deficient Mice. *J. Exp. Med.* **191**, 2197-2207.

Attardi, B. & Attardi, G. (1972). Persistence of Thymidine Kinase Activity in Mitochondria of a Thymidine Kinase-Deficient Derivative of Mouse L Cells. *Proc. Natl. Acad. Sci.* **69**, 2874-2878.

Ayabe, T., Takenaka, H., Takenaka, O., Sumida, M., Maruyama, H., Onitsuka, T., Shibata, K., Uesugi, S. & Hamada, M. (1997). Essential Lysine Residues in the N-Terminal and the C-Terminal Domain of Human Adenylate Kinase Interact with Adenine Nucleotides As Found by Site-Directed Mutagenesis. *Biochemistry* **36**, 4027-4033.

Balasubramaniam, N. K., Veerisetty, V. & Gentry, G. A. (1990). Herpesviral Deoxythymidine Kinases Contain a Site Analogous to the Phosphoryl-Binding Arginine-Rich Region of Porcine Adenylate Kinase; Comparison of Secondary Structure Predictions and Conservation. *J. Gen. Virol.* **71** (Pt12), 2979-2987.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M. & Sonnhammer, E. L. L. (2002). The Pfam Protein Families Database. *Nucleic Acids Res.* **30**, 276-280.

Bates, S., Phillips, A. C., Clark, P. A., Stott, F., Peters, G., Ludwig, R. L. & Vousden, K.H. (1998). P14^{ARF} Links the Tumour Suppressors RB and p53. *Nature* **395**, 124-125.

Behrends, H. W., Beck-Sickinger, A. G. & Folkers, G. (1996). Evaluation of the Secondary Structure of Vaccinia-virus Thymidine Kinase by Circular-Dichroism Spectroscopy of Overlapping Synthetic Peptides. *Eur. J. Biochem.* **241**, 126-132.

Beier, F., Lee, R. J., Taylor, A. C., Pestell, R. G. & LuValle, P. (1999). Identification of The Cyclin D1 Gene as a Target of Activating Transcription Factor 2 in Chondrocytes. *Proc. Natl. Acad. Sci. USA* **96**, 1433-1438.

Beier, R., Bürgin, A., Kiermaier, A., Fero, M., Karsunky, H., Saffrich, R., Möröy, T., Ansorge, W., Roberts, J. & Eilers, M. (2000a). Induction of Cyclin E-CDK2 KinaseActivity, E2F-Dependent Transcription and Cell Growth by Myc Are Genetically Separable Events. *EMBO J.* **19**, 5813-5823.

Beier, F., Taylor, A. C. & LuValle, P. (2000b). Activating Transcription Factor 2 isNecessary for Maximal Activity and Serum Induction of the Cyclin A Promoter inChondrocytes. *J. Biol. Chem.* **275**, 12948-12953.

Bello, L. J. (1974). Regulation of Thymidine Kinase Synthesis in Human Cells. *Exp. Cell Res.* **89**, 263-74.

Belt, J. A., Marina, N. M., Phelps, D. A. & Crawford, C. R. (1993). NucleosideTransport in Normal and Neoplastic Cells. *Advan. Enzyme Regul.* **33**, 235-252.

Bennett , M. S., Wien, F., Champness, J. N., Batuwangala, T., Rutherford, T., Summers, W. C., Sun, H., Wright, G. & Sanderson, M. R. (1999). Structure to 1.9 Å Resolution of a Complex with Herpes Simplex Virus Type-1 Thymidine Kinase of a Novel, Non-substrate Inhibitor: X-Ray Crystallographic Comparison with Binding of Aciclovir. *FEBS Lett.* **443**, 121-125.

Berens, R. L., Krug, E. C. & Marr, J. J. (1995). Purine and Pyrimidine Metabolism. In: *Biochemistry and Molecular Biology of Parasites*. Eds.: Marr, J. J. & Müller, M. Academic Press.

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). *J. Biol. Chem.* **275**, 32187-32192.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F. Jr., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). The Protein Data Bank: A Computer-based Archival File for Macromolecular Structures. *J. Mol. Biol.* **112**, 535-542.

Berk, A. J. & Clayton, D. A. (1973). A Genetically Distinct Thymidine Kinase in Mammalian Mitochondria. Exclusive Labelling of Mitochondrial Deoxyribonucleic Acid. *J. Biol. Chem.* **248**, 2722-2729.

Best, C. J., McKenna, P. G. & McKelvey-Martin, V. J. (1997). Mutagen Sensitivity in Thymidine Kinase- and Methyltrasferase-Deficient Human Lymphoblastoid Cells. *Br. J.Biomed. Sci.* **54**, 267-272.

Bestwick, R. K., Moffett, G. L. & Mathews, C. K. (1982). Selective Expansion of Mitochondrial Nucleoside Triphosphate Pools in Antimetabolite-treated HeLa Cells. *J. Biol. Chem.* **257**, 9300-9304.

Bianchi, V., Pontis, E. & Reichard, P. (1986). Interrelations Between Substrate Cycles and *de novo* Synthesis of Pyrimidine Deoxyribonucleoside Triphosphates in 3T6 cells. *Proc. Natl. Acad. Sci. USA* **83**, 986-990.

Bianchi, V., Ferraro, P., Borella, S., Bonvini, P. & Reichard, P. (1994). Effects of Mutational Loss of Nucleoside Kinases on Deoxyadenosine 5'-Phosphate/Deoxyadenosine Substrate Cycle in Cultured CEM and V79 Cells. *J. Biol. Chem.* **269**, 16677-16683.

Bianchi, V. & Spychala, J. (2003). Mammalian 5'-Nucleotidases. *J. Biol. Chem.* **278**, 46195-46198.

Billeter, M., Braun, W. & Wüthrich, K. (1982). Sequential Resonance Assignments in Protein 1H Nuclear Magnetic Resonance Spectra. *J. Mol. Biol.* **155**, 321-346.

Billeter, M., Qian, Y., Otting, G., Müller, M., Gehring, W. J. & Wüthrich, K. (1990). Determination of the Two-dimensional Structure of the *Antennapedia* Homodomain from *Drosophila* in Solution by 1H Nuclear Magnetic Resonance. *J. Mol. Biol.* **214**, 183-197.

Bird, L. E., Ren, J., Wright, A., Leslie, K. D., Degréve, B., Balzarini, J. & Stammers, D.K. (2003). Crystal Structure of Varicella Zoster Virus Thymidine Kinase. *J. Biol. Chem.* **278**, 24680-24687.

Black, M. E. & Hruby, D. E. (1990). Identification of the ATP-binding Domain of Vaccinia Virus Thymidine Kinase. *J. Biol. Chem.* **265**, 17584-17592.

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**, 6801-6806.

Black, M. E. & Loeb, L. A. (1993). Identification of Important Residues within the Putative Nucleotide Binding Site of HSV-1 Thymidine Kinase by Random Sequence Selection: Analysis of Selected Mutants *in Vitro. Biochemistry* **32**, 11618-11626.

Black, M., Rechtin, T. M. & Drake, R. R. (1996). Effect on Substrate Binding of an Alteration at the Conserved Aspartic Acid-162 in Herpes Simplex Virus Type 1 Thymidine Kinase. *J. Gen. Virol.* **77**, 1521-1527.

Blaszczyk, J., Li, Y., Yan, H. & Ji, X. (2001). Crystal Structure of Unligated Guanylate Kinase from Yeast reveals GMP-Induced Conformational Changes. *J. Mol. Biol.* **307**, 247-257.

Bogan, A. A. & Thorn, K. S. (1998). Anatomy of Hot Spots in Protein Interfaces. *J. Mol. Biol.* **280**, 1-9.

Bogenhagen, D. & Clayton, D. A. (1976). Thymidylate Nucleotide Supply for Mitochondrial DNA Synthesis in Mouse L-Cells. Effect of 5-Fluorodeoxyuridine and Methotrexate in Thymidine Kinase Plus and Thymidine Kinase Minus Cells. *J. Biol. Chem.* **251**, 2938-2944.

Bohman, C. & Eriksson, S. (1988). Deoxycytidine Kinase from Human Leukemic Spleen: Preparation and Characterization of the Homogeneous Enzyme. *Biochemistry* **27**, 4258-4265.

Bouchard, C., Thieke, K., Maier, A., Saffrich, R., Hanley-Hyde, J., Ansorge, W., Reed, S., Sicinski, P., Bartek, J. & Eilers, M. (1999). Direct Induction of Cyclin D2 by Myc Contributes to Cell Cycle Progression and Sequestration of p27. *EMBO J.* **19**, 5321-5333.

Boyle, D. B., Coupar, B. E. H., Gibbs, A. J., Seigman, L. J. & Both, G. W. (1987). Fowlpox Virus Thymidine Kinase: Nucleotide Sequence and Relationship to Other Thymidine Kinases. *Virology* **156**, 355-365.

Bracken, A. P., Ciro, M., Cocito, A. & Helin, K. (2004). E2F Target Genes: Unraveling the Biology. *Trends Biochem. Sci.* **29**, 409-417.

Bradford, M. M. (1976). A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **72**, 248-254.

Bradshaw, H. D. Jr. (1983). Molecular Cloning and Cell Cycle Specific Regulation of a Functional Human Thymidine Kinase Gene. *Proc. Natl. Acad. Sci. USA* **80**, 5588-5591.

Bradshaw, H. D. Jr. & Deininger, P. L. (1984). Human Thymidine Kinase Gene: Molecular Cloning and Nucleotide Sequence of a cDNA Expressible in Mammalian Cells. *Mol. Cell. Biol.* **4**, 2316-2320.

Brange, J., Ribel, U., Hansen, J. F., Dodson, G., Hansen, M. T., Havelund, S., Melberg, S. G., Norris, F., Norris, K., Snel, L., Sørensen, A. R. & Voigt, H. O. (1988). Monomeric Insulins Obtained by Protein Engineering and Their Medical Implications. *Nature*, **333**, 679-682.

Brehm, A & Kouzarides, T. (1999). Retinoblastoma Protein Meets Chromatin. *Trends Biochem. Sci.* **24**, 142-145.

Bridges, E. G., Jiang, Z. & Cheng Y. (1999). Characterization of a dCTP Transport Activity Reconstituted from Human Mitochondria. *J. Biol. Chem.* **274**, 4620-4625.

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 Gangciclovir. *Nat. Struct. Biol.* **2**, 876-881.

Brugarolas, J., Moberg, K., Boyd, S.D., Taya, Y. & Jacks, T. (1999). Inhibition of Cyclin-Dependent Kinase 2 by p21 Is Necessary for Retinoblastoma Protein-Mediated G1 Arrest After γ-Irradiation. *Proc. Natl. Acad. Sci. USA* **96**, 1002-1007.

Byeon, I. L., Shi, Z. & Tsai, M. D. (1995). Mechanism of Adenylate Kinase. The "Essential Lysine" Helps to Orient the Phosphates and the Active Site Residues to Proper Conformation. *Biochemistry* **34**, 3172-3182.

Carozza, M. A. & Conrad, S. E. (1994). Regulation of Thymidine Kinase Protein Stability in Serum-stimulated Cells. *Cell Growth & Differ.* **5**, 901-908.

Carrozzo, R., Bornstein, B., Lucioli, S., Campos, Y., de la Pena, P., Petit, N., Dionisi-Vici, C., Vilarinho, L., Rizza, T., Bertini, E., Garesse, R., Santorelli, F. M. & Arenas, J. (2003). Mutation Analysis in 16 Patients with mtDNA Depletion. *Hum. Mutat.* **21** (4), 453-454.

Carson, D. A., Kaye, J. & Wasson, D. B. (1981). The Potential Importance of Soluble Deoxynucleotidase Activity in Mediating Deoxyadenosine Toxicity in Human Lymphoblasts. *J. Immunol.* **126**, 348-352.

Cass, C. E., Young, J. D., Baldwin, S. A., Cabrita, M. A., Graham, K. A., Griffiths, M., Jennings, L. L., Mackey, J. R., Ng A. M. L., Ritzel, M. W. L., Vickers, M. F. & Yao, S. Y. M. (1999). Nucleoside Transporters of Mammalian Cells. *Pharm. Biotechnol.* **12**, 313-352.

Chabes, A. L., Pfleger, C. M., Kirshner, M. W. & Thelander, L. (2003). Mouse Ribonucleotide Reductase R2 Protein: A New Target for Anaphase-Promoting Complex-Cdh1-Mediated Proteolysis. *Proc. Natl. Acad. Sci.* **100**, 3925-3929.

Champness, J. N., Bennett, M. S., Wien, F., Visse, R., Summers, W. C., Herdewijn, P., de Clercq, 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: Struct. Funct. Genet.* **32**, 350-361.

Chang, F. & McCubrey, J. A. (2001). P21[Cip1] Induced by Raf is Associated with Increased CDK4 Activity in Hematopoietic Cells. *Oncogene* **20**, 4354-4364.

Chang, L. & Karin, M. (2001). Mammalian MAP Kinases Signalling Cascades. *Nature* **410**, 37-40.

Chang, Z. F. & Chen, K. Y. (1988). Regulation of Ornithine Decarboxylase and Other Cell Cycle-Dependent Genes during Senescence of IMR-90 Human Diploid Fibroblasts. *J. Biol. Chem.* **263**, 11431-11435.

Chang, Z. F. & Cheng, S. M. (1993). Constitutive Overexpression of DNA Binding Activity to the Distal CCAAT Box of Human Thymidine Kinase Promoter in Human Tumor Cell Lines. *Cancer Res.* **53**, 3253-3256.

Chang, Z. F. & Huang, D. Y. (1993). The Regulation of Thymidine Kinase in HL-60 Human Promyeloleukemia Cells. *J. Biol. Chem.* **268**, 1266-1271.

Chang, Z. F., Huang, D. Y. & Hsue, N. C. (1994). Differential Phosphorylation of Human Thymidine Kinase in Proliferating and M Phase-arrested Human Cells. *J. Biol. Chem.* **269**, 21249-21254.

Chang, Z. F. & Liu, C. J. (1994). Human Thymidine Kinase CCAAT-binding Protein is NF-Y, Whose A subunit Expression Is Serum-dependent in Human IMR-90 Diploid Fibroblasts. *J. Biol. Chem.* **269**, 17893-17898.

Chang, Z. F., Huang, D. Y. & Lai, T. C. (1995). Different Regulation of the Human Thymidine Kinase Promoter in Normal Human Diploid IMR-90 Fibroblasts and HeLa Cells. *J. Biol. Chem.* **270**, 27374-27379.

Chang, Z. F., Huang, D. Y. & Chi, L. M. (1998). Serine 13 Is the Site of Mitotic Phosphorylation of Human Thymidine Kinase. *J. Biol. Chem.* **273**, 12095-12100.

Chang, Z. F., Huang, D. Y. & Hu, S. F. (1999). NF-Y- Mediated *Trans*-Activation of the Human Thymidine Kinase Promoter Is Closely Linked to Activation of Cyclin-Dependent Kinase. *J. Cell. Biochem.* **75**, 300-309.

Chang, Z. F. & Huang, D. Y. (2001). Regulation of Thymidine Kinase Expression during Cellular Senescence. *J. Biomed. Sci.* **8**, 176-183.

Chen, C. H. & Cheng, Y. C. (1992). The Role of Cytoplasmic Deoxycytidine Kinase in the Mitochondrial Effects of the Antihuman Immunodeficiency Virus Compound, 2',3'-Dideoxycytidine. *J. Biol. Chem.* **267**, 2856-2859.

Chen, E. H., Johnson, E. E. II, Vetter, S. M. & Mitchell, B. (1995). Characterization of the Deoxycytidine Kinase Promoter in Human Lymphoblast Cell Lines. *J. Clin. Invest.* **95**, 1660-1668.

Cheng, M., Olivier, P., Diehl, J. A., Fero, M., Roussel, M. F., Roberts, J. M. & Sherr, C. J. (1999). The p21 and p27 CDK 'Inhibitors' Are Essential Activators of Cyclin D-Dependent Kinases in Murine Fibroblasts. *EMBO J.* **18**, 1571-1583.

Chimploy, K., Tassotto, M. L. & Mathews, C. K. (2000). Ribonucleotide Reductase, a Possible Agent in Deoxyribonucleotide Pool Asymmetries Induced by Hypoxia. *J. Biol. Chem.* **275**, 39267-39271.

Chimploy, K. & Mathews, C. K. (2001). Mouse Ribonucleotide Reductase Control. Influence of Substrate Binding upon Interactions with Allosteric Effectors. *J. Biol. Chem.* **276**, 7093-7100.

Chothia, C. & Janin, J. (1975). Principles of Protein-Protein Recognition. *Nature* **256**, 705-708.

Chothia, C. & Lesk, A. M. (1986). The Relation Between the Divergence of Sequence and Structure in Proteins. *EMBO J.* **5**, 823-826.

Chottiner, E. G., Shewah, D. S., Datta, N. S., Ashcraft, E., Gribbin, D., Ginsburg, D., Fox, I. H. & Mitchell, B. S. (1991). Cloning and Expression of Human Deoxycytidine Kinase cDNA. *Proc. Natl. Acad. Sci. USA* **88**, 1531-1535.

Chou, P. Y. & Fasman, G. D. (1978). Prediction of the Secondary Structure of Proteins from Their Amino Acid Sequence. *Advan. Enzymol.* **47**, 45-148.

Chou, W. L. & Chang, Z. F. (2001). Cap-Independent Translation Conferred by the 5'-Untranslated Region of Human Thymidine Kinase mRNA. *Biochim. Biophys. Acta* **1519**, 209-215.

Christensen, J. F. (2000). Identifikation af Aminosyre 106 I Human Thymidin Kinase 1 (TK1) fra Normale Lymfocytter og Cancer Celle Linjer. Specialerapport i molekylærbiologi. Roskilde Universitetscenter.

Chuang, L. S. H., Ian, H. I., Koh, T. W., Ng, H. H., Xu, G. & Li, B. F. L. (1997). Human DNA-(Cytosine-5) Methyltransferase- PCNA Complex as a Target for p21^{WAF1}. *Science* **277**, 1996-2000.

Ciechanover, A. (1998). The Ubiquitin-Proteasome Pathway: On Protein Death and Cell Life. *EMBO J.* **17**, 7151-7160.

Ciechanover, A. & Ben-Saadon, R. (2004). N-Terminal Ubiquitination: More Protein Substrates Join in. *Trends Cell Biol.* **14**, 103-106.

Clurman, B. E. & Porter, P. (1998). New Insights Into the Tumor Suppression Function of p27. *Proc. Natl. Acad. Sci.* **95**, 15158-15160.

Coats, S., Flanagan, W. M., Nourse, J. & Roberts, J. M. (1996). Requirement of p27 for Restriction Point Control of the Fibroblast Cell Cycle. *Science* **272**, 877-880.

Coats, S., Whyte, P., Fero, M. L., Lacy, S., Chung, G., Randel, E., Firpo, E. & Roberts, J. M. (1999). A New Pathway for Mitogen-Dependent CDK2 Regulation Uncovered in p27-Deficient Cells. *Curr. Biol.* **9**, 163-173.

Cohen, A., Gudas, L. J., Ammann, A. J., Staal, G. E. J. & Martin, D. W. (1978). Deoxyguanosine Triphosphate as a Possible Toxic Metabolite in the Immunodeficiency Associated with Purine Nucleoside Phosphorylase Deficiency. *J. Clin. Invest.* **61**, 1405-1409.

Cohen, A., Barankiewicz, J., Lederman, H. M. & Gelfand,E. (1983). Purine and Pyrimidine Metabolism in Human T Lymphocytes. Regulation of Deoxyribonucleotide Metabolism. *J. Biol. Chem.* **258**, 12334-12340.

Coleman, M. S., Donofrio, J., Hutton, J. J., Hahn, L., Daoud, A., Lampkin, B. & Dyminski, J. (1978). Identification and Quantification of Adenine Deoxynucleotides in Erythrocytes of a Patient with Adenosine Deaminase Deficiency and Severe Combined Immunodeficiency. *J. Biol. Chem.* **253**, 1619-1626.

Cook, D. A. (1967). The Relation between Amino Acid Sequence and Protein Conformation. *J. Mol. Biol.* **29**, 167-171.

Coppock, D. L. & Pardee, A. B. (1985). Regulation of Thymidine Kinase Activity in the Cell Cycle by a Labile Protein. *J. Cell. Physiol.* **124**, 269-274.

Coppock, D. L. & Pardee, A. B. (1987). Control of Thymidine Kinase mRNA During the Cell Cycle. *Mol. Cell. Biol.* **7**, 2925-2932.

Coqueret, O. (2003). New Roles for p21 and p27 Cell Cycle Inhibitors: A Function for Each Cell Compartment? *Trends Cell Biol.* **13**, 65-70.

Cornish-Bowden, A. (ed.) (1995). Control of Enzyme Activity. In: *Fundamentals of Enzyme Kinetics*, pp.203-237. Portland Press Ltd, London.

Csapó, Z., Sasvari-Szekely, M., Spasokoukotskaja, T. & Staub, M. (2001a). Modulation of Human Deoxycytidine Kinase Activity as a Response to Cellular Stress Induced by NaF. *Acta Biochim. Polonica* **48**, 251-256.

Csapó, Z., Sasvari-Szekely, M., Spasokoukotskaja, T., Talianidis, I., Eriksson, S. & Staub, M. (2001b). Activation of Deoxycytidine Kinase by Inhibition of DNA Synthesis in Human Lymphocytes. *Biochem. Pharmacol.* **61**, 191-197.

Csapó, Z., Keszler, G., Safrany, G., Spasokoukotskaja, T., Talinidis, I., Staub, M. & Sasvari-Szekely, M. (2003). Activation of Deoxycytidine Kinase by Gamma-Irradiation and Inactivation by Hyperosmotic Shock in Human Lymphocytes. *Biochem. Pharmacol.* **65**, 2031-2039.

Cuff, J. A., Clamp, M. E., Siddiqui, A. S., Finlay, M. & Barton, G. J. (1998). JPred: a Consensus Secondary Structure Prediction Server. *Bioinformatics* **14**, 892-893.

Cuff, J. A. & Barton, G. J. (1999). Evaluation and Improvement of Multiple Sequence Methods for Protein Secondary Structure Prediction. *Proteins: Struct. Function Genet.* **34**, 508-519.

Cuff, J. A. & Barton, G. J. (2000). Application of Multiple Sequence Alignment Profiles to Improve Secondary Structure Prediction. *Proteins: Struct. Function Genet.* **40**, 502-511.

Dahnke, T. & Tsai, M. D. (1994). Mechanism of Adenylate Kinase. The Conserved Aspartates 140 and 141 Are Important for Transition State Stabilization Instead of Substrate-Induced Conformational Changes. *J. Biol. Chem.* **269**, 8075-8081.

Dang, C. V. (1999). c-Myc Target Genes Involved in Cell Growth, Apoptosis and Metabolism. *Mol. Cell. Biol.* **19**, 1-11.

Dao-pin, S., Baase, W. A. & Matthews, B. W. (1990). A Mutant T4 Lysozyme (Val131 to Ala) Designed to Increase Thermostability by the Reduction of Strain Within an Alpha-helix. *Proteins* **7**, 198-204.

Darby, G., Larder, B. A. & Inglis, M. M. (1986). Evidence that the "Active Centre" of the Herpes Simplex Virus Thymidine Kinase Involves an Interaction between Three Distinct Regions of the Polypeptide. *J. Gen. Virol.* **67**, 753-758.

Darè, E., Zhang, L. H., Jenssen, D. & Bianchi, V. (1995). Molecular Analysis of Mutations in the *hprt* Gene of V79 Hamster Fibroblasts: Effects of Imbalances in the dCTP, dGTP and dTTP pools. *J. Mol. Biol.* **252**, 514-521.

Datta, N. S., Shewach, D. S., Hurley, M. C., Mitchell, B. S. & Fox, I. R. (1989a). Human T-Lymphoblast Deoxycytidine Kinase: Purification and Properties. *Biochemistry* **28**, 114-123.

Datta, N. S., Shewach, D. S., Mitchell, B. S. & Fox, I. H. (1989b). Kinetic Properties and Inhibition of Human T Lymphoblast Deoxycytine Kinase. *J. Biol. Chem.* **264**, 9359-9364.

Deber, C. M., Li, Z., Joensson, C., Glibowicka, M. & Xu, G. Y. (1992). Transmembrane Region of Wild-type and Mutant M13 Coat Proteins. Conformational Role of β-branched Residues. *J. Biol. Chem.* **267**, 5296-5300.

Deber, C. M., Khan, A. R., Li, Z., Joensson, C., Glibowicka, M. & Wang, J . (1993). Val to Ala Mutations Selectively Alter Helix-Helix Packing in the Transmembrane Segment of Phage M13 Coat Protein. *Proc. Natl. Acad. Sci. USA* **90**, 11648-11652.

De Boer, H. A., Comstock, L. J. & Vasser, M. (1983). The *tac* Promoter: A Functional Hybrid Derived from the *trp* and *lac* promoters. *Proc. Natl. Acad. Sci. USA* **80**, 21-25.

DeGregori, J., Kowalik, T. & Nevins, J. R. (1995). Cellular Targets for Activation by the E2F1 Transcription Factor Include DNA Synthesis-and G1/S-Regulatory Genes. *Mol. Cell. Biol.* **15**, 4215-4224.

DeSalle, L. M. & Pagano, M. (2001). Regulation of the G1 to S Transition by the Ubiquitin Pathway. *FEBS Letters* **490**, 179-189.

De Winter, H. & Herdewijn, P. (1996). Understanding the Binding of 5-Substituted 2'-Deoxyuridine Substrates to Thymidine Kinase of Herpes Simplex Virus Type-1. *J. Med. Chem.* **39**, 4727-4737.

Diederichs, K. & Schulz, G. E. (1990). Three-Dimensional Structure of the Complex between the Mitochondrial Matrix Adenylate Kinase and Its Substrate AMP. *Biochemistry* **29**, 8138-8144.

Diederichs, S., Bäumer, N., Ji, P., Metzelder, S. K., Idos, G. E., Cauvet, T., Wang, W., Möller, M., Pierschalski, S., Gromoll, J., Schrader, M. G., Koeffler, H. P., Berdel, W. E., Serve, H. & Müller-Tidow, C. (2004). Identification of Interaction Partners and Substrates of the Cyclin A1-CDK2 Complex. *J. Biol. Chem.* **279**, 33727-33741.

Dobrovolsky, V. N., Bucci, T., Heflich, R. H., Desjardins, J. & Richardson, F. C. (2003). Mice Deficient for Cytosolic Thymidine Kinase Gene Develop Fatal Kidney Disease. *Mol. Genet. Metabol.* **78**, 1-10.

Dolce, V., Fiermonte, G., Runswick, M. J., Palmieri, F. & Walker, J. E. (2001). The Human Mitochondrial Deoxynucleotide Carrier and its Role in the Toxicity of Nucleoside Antivirals. *Proc. Natl. Acad. Sci. USA* **98**, 2284-2288.

Dou, Q. P., Fridovich-Keil, J. L. & Pardee, A. B. (1991). Inducible Proteins Binding to the Murine Thymidine Kinase Promoter in Late G1/S Phase. *Proc. Natl. Acad. Sci. USA* **88**, 1157-1161.

Dou, Q. P., Markell, P. J. & Pardee, A. B. (1992). Thymidine Kinase Transcription is Regulated at G1/S Phase by a Complex that Contains Retinoblastoma-like Protein and a Cdc2 Kinase. *Proc. Natl. Acad. Sci. USA* **89**, 3256-3260.

Dou, Q. P., Molnar, G. & Pardee, A. B. (1994a). Cyclin D1/CDK2 Kinase Is Present in a G1 Phase-Specific Protein Complex Yi1 that Binds to the Mouse Thymidine Kinase Gene Promoter. *Biochem. Biophys. Res. Commun.* **205**, 1859-1868.

Dou, Q. P., Zhao, S., Levin, A. H., Wang, J., Helin, K. & Pardee, A. B. (1994b). G1/S- Regulated E2F-containing Protein Complexes Bind to the Mouse Thymidine Kinase Gene Promoter. *J. Biol. Chem.* **269**, 1306-1313.

Dou, Q. P. & Pardee, A. B. (1996). Transcriptional Activation of Thymidine Kinase, a Marker for Cell Cycle Control. *Progress Nucleic Acids Res.* **53**, 197-217.

Dreusicke, D. & Schulz, G. E. (1986). The Glycine-Rich Loop of Adenylate Kinase Forms a Giant Anion Hole. *FEBS Lett.* **208**, 301-304.

Dunker, A. K., Lawson, J. D., Brown, C. J., Williams, R. M., Romero, P., Oh, J. S., Oldfield, C. J., Campen, A. M., Ratliff, C. M., Hipps, K. W., Ausio, J., Nissen, M. S., Reeves, R., Kang, C.H., Kissinger, C. R., Bailey, R. W., Griswold, M. D., Chiu, W., Garner, E. C. & Obradovic, Z. (2001). Intrinsically Disordered Proteins. *J. Mol. Graphics Modelling* **19**, 26-59.

Dunker, A. K., Brown, C. J., Lawson, D., Iakoucheva, L. M. & Obradovic, Z. (2002). Intrinsic Disorder and Protein Function. *Biochemistry* **41**, 6573-6582.

Dybdahl, M., Vogel, U., Frentz, G., Wallin, H. & Nexø, B. (1999). Polymorphisms in the DNA Repair Gene *XPD*: Correlations with Risk and Age at Onset of Basal Cell Carcinoma. *Cancer Epidemiol. Biomarkers & Prevention* **8**, 77-81.

Dyson, N. (1998). The Regulation of E2F by pRB-Family Proteins. *Genes & Dev.* **12**, 2245-2262.

Egner, U., Tomasselli, A. G. & Schulz, G. E. (1987). Structure of the Complex of Yeast Adenylate Kinase with the Inhibitor P^1,P^5-di(adenosine-5'-)pentaphosphate at 2.6 Å Resolution. *J. Mol. Biol.* **195**, 649-658.

Elford, H. L., Freese, M., Passamani E. & Morris, H. P. (1970). Ribonucleotide Reductase and Cell Proliferation. I.Variations of Ribonucleotide Reductase Activity with Tumor Growth Rate in a Series of Rat Hepatomas. *J. Biol. Chem.* **245**, 5228-5233.

Ellis, R. J. (2001). Macromolecular Crowding: An Important But Neglected Aspect of the Intracellular Environment. *Current Opinion Struct. Biol.* **11**. 114-119.

Elpeleg, O., Saada, A. & Mandel. H. (2002). Depletion of the Other Genome-Mitochondrial DNA Depletion Syndromes in Humans. *J. Mol. Med.* **80**, 389-396.

Engström, Y., Rozell, B., Hansson, H. A., Stemme, S. & Thelander, L. (1984). Localization of Ribonucleotide Reductase in Mammalian Cells. *EMBO J.* **3**, 863-867.

Engström, Y., Eriksson, S., Jildevik, I., Skog, S., Thelander, L. & Tribukait, B. (1985). Cell Cycle-dependent Expression of Mammalian Ribonucleotide Reductase: Differential regulation of the two subunits. *J. Biol. Chem.* **260**, 9114-9116.

Engström, Y. & Rozell, B. (1988). Immunocytochemical Evidence for the Cytoplasmic Localization and Differential Expression during the Cell Cycle of the M1 and M2 Subunits of Mammalian Ribonucleotide Reductase. *EMBO J.* **7**, 1615-1620.

Eriksson, S., Gudas, L. J., Clift, S. M., Caras, I. W., Ullman, B. & Martin, D. W. Jr. (1981). Evidence for Genetically Independent Allosteric Regulatory Domains of the Protein M1 Subunit of Mouse Ribonucleotide Reductase. *J. Biol. Chem.* **256**, 10193-10197.

Eriksson, S., Gräslund A., Skog, S., Thelander, L. & Tribukait, B. (1984). Cell Cycle-dependent Regulation of Mammalian Ribonucleotide Reductase. The S phase-correlated Increase in Subunit M2 is Regulated by *de novo* Protein Synthesis. *J. Biol. Chem.* **259**, 11695-11700.

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, 586-592.

Eriksson, S., Amér, E., Spasokoukotskaja, T., Wang, L., Karlsson, A., Brosjö, O., Guvén, P., Julusson, G. & Liliemark, J. (1994). Properties and Levels of Deoxynucleoside Kinases in Normal and Tumor Cells; Implications for Chemotherapy. *Advan. Enzyme Regul.* 34, 13-25.

Eriksson, S. & Wang, L. (1997). Substrate Specificities, Expression and Primary Sequences of Deoxynucleoside Kinases; Implications for Chemotherapy. *Nucleosides Nucleotides* 16, 653-659.

Eriksson, S., Munch-Petersen, B., Johansson, K. & Eklund, H. (2002). Structure and Function of Cellular Deoxyribonucleoside Kinases. *CMLS, Cell. Mol. Life Sci.* 59, 1327-1346.

Evans, J. S., Lock, K. P., Levine, B. A., Champness, J. N., Sanderson, M. R., Summers, W. C., McLeish, P. J. & Buchan, A. (1998). Herpesviral Thymidine Kinases: Laxity and Resistance by Design. *J. Gen. Virol.* 79, 2083-2092.

Ewen, M. W. (2000). Relationship between Ras Pathways and Cell Cycle Control. *Progress Cell Cycle Res.* 4, 1-17.

Ferentz, A. E. & Wagner, G. (2000). NMR Spectroscopy: A Multifaceted Approach to Macromolecular Structure. *Quarterly Rev. Biophys.* 33, 29-65.

Ferraro, P., Bianchi, V., Biasin, M. R. & Celotti, L. (1992). Deoxynucleotide Pools and DNA Synthesis in Resting and PHA-Stimulated Human Lymphocytes Treated with Mutagens. *Exp. Cell Res.* 199, 349-354.

Fetzer, J. & Folkers, G. (1992). Temperature Dependent Formation of Inclusion Bodies during the Expression of Recombinant Viral Thymidine Kinase. *Pharm. Pharmacol. Lett.* 2, 112-114.

Flemington, E., Bradshaw, H. D. Jr., Traina-Dorge, V., Slagel, V. & Deininger, P. L. (1987). Sequence, Structure and Promoter Characterization of the Human Thymidine Kinase Gene. *Gene* 52, 267-277.

Folkers, G., Trumpp-Kallmeyer, S., Gutbrot, O., Krickl, S., Fetzer, J. & Keil, G. M. (1991). Computer-aided Active-site-directed Modeling of the Herpes Simplex Virus 1 and Human Thymidine Kinase. *J. Comp. Aided Mol. Des.* 5, 385-404.

Fontana, A., Zambonin, M., de Laureto, P. P., De Filippis, V., Clementi, A. & Scaramella, E. (1997). Probing the Conformational State of Apomyoglobin by Limited Proteolysis. *J. Mol. Biol.* 266, 223-230.

Fontecave, M., Nordlund, P., Eklund, H. & Reichard, P. (1992). The Redox Centers of Ribonucleotide Reductase of *Escherichia coli. Adv. Enzymol. Relat. Areas Mol. Biol.* 65, 147-83.

Fontecave, M. (1998). Ribonucleotide Reductases and Radical Reactions. *CMLS, Cell. Mol. Life Sci.* 54, 684-695.

Forrest, A. & Gabrielli, B. (2001). Cdc25B Activity is Regulated by 14-3-3. *Oncogene* 20, 4393-4401.

Frederiksen H. (1999). Betydningen af Aminosyre 106 for Struktur og Funktion af Human Thymidin Kinase I. Specialerapport i molekylærbiologi. Roskilde Universitetscenter.

Frederiksen, H., Berenstein, D. & Munch-Petersen, B. (2004). Effect of Valine 106 on Structure-Function Relation of Cytosolic Human Thymidine Kinase. Kinetic Properties and Oligomerization Pattern of Nine Substitution Mutants of V106. *Eur. J. Biochem.* **271**, 2248-2256.

Fry, D. C., Kuby, S. A. & Mildvan, A. S. (1986). ATP-binding Site of Adenylate Kinase: Mechanistic Implications of Its Homology with *ras*-encoded p21, F1-ATPase, and Other Nucleotide-binding Proteins. *Proc. Natl. Acad. Sci. USA* **83**, 907-911.

Fukuda, M., Gotoh, Y. & Nishida, E. (1997). Interaction of MAP Kinase with MAP Kinase Kinase: its Possible Role in the Control of Nucleocytoplasmic Transport of MAP kinase. *EMBO J.* **16**, 1901-1908.

Furuno, N., den Elzen, N. & Pines, J. (1999). Human Cyclin A Is Required for Mitosis until Mid Prophase. *J. Cell. Biol.* **147**, 295-306.

Gallinaro, L., Crovatto, K., Rampazzo, C., Pontarin, G., Ferraro, P., Milanesi, E., Reichard, P. & Bianchi, V. (2002). Human Mitochondrial 5'-Deoxyribonucleotidase. Overproduction in Cultured Cells and Functional Aspects. *J. Biol. Chem.* **277**, 35080-35087.

Garnier, J., Osguthorpe, D. J. & Robson, B. (1978). Analysis of the Accuracy and Implications of Simple Methods for Predicting the Secondary Structure of Globular Proteins. *J. Mol. Biol.* **120**, 97-120.

Gazziola, C., Ferraro, P., Moras, M., Reichard, P. & Bianchi, V. (2001). Cytosolic High K_m 5'-Nucleotidase and 5'(3')-Deoxyribonucleotidase in Substrate Cycles Involved in Nucleotide Metabolism. *J. Biol. Chem.* **276**, 6185-6190.

Ge, Y., Jensen, T. L., Matherly, L. H. & Taub, J. W. (2003). Physical and Functional Interactions between USF and Sp1 Proteins Regulate Human Deoxycytidine Kinase Promoter Activity. *J. Biol. Chem.* **278**, 49901-49910.

Geley, S., Kramer, E., Gieffers, C., Gannon, J., Peters, J. M. & Hunt, T. (2001). Anaphase-promoting Complex/Cyclosome-dependent Proteolysis of Human Cyclin A Starts at the Beginning of Mitosis and I Not Subject to the Spindle Assembly Checkpoint. *J. Cell Biol.* **153**, 137-147.

Gentry, G. A. (1992). Viral Thymidine Kinases and Their Relatives. *Pharmac. Ther.* **54**, 319-355.

Ghoda, L., Sidney, D., Macrae, M. & Coffino, P. (1992). Structural Elements of Ornithine Decarboxylase Required for Intracellular Degradation and Polyamine-dependent Regulation. *Mol. Cell. Biol.* **12**, 2178-2185.

Gilles, S. I., Romain, S., Casellas, P., Ouafik, I. H., Fina, F., Combes, T., Vuaroquaux, V., Seitz, J. F., Bonnier, P., Gallégue, S., Carayon, P. & Martin, P. M. (2003). Mutation Analysis in the Coding Sequence of Thymidine Kinase 1 in Breast and Colorectal Cancer. *Int. J. Biol. Markers* **18**, 1-6.

Glaser, F., Steinberg, D. M., Vakser, I. A. & Ben-Tal, N. (2001). Residue Frequencies and Pairing Preferences at Protein-Protein Interfaces. *Proteins: Struct. Funct. Genet.* **43**, 89-102.

Goncalves, O., Drobetsky, E. & Meuth, M. (1984). Structural Alterations of the *aprt* Locus Induced by Deoxyribonucleoside Triphosphate Pool Imbalances in Chinese Hamster Ovary Cells. *Mol. Cell. Biol.* **4**, 1792-1799.

Good, L., Chen, J. & Chen, K. Y. (1995). Analysis of Sequence-Specific Binding Activity of Cis-Elements in Human Thymidine Kinase Gene Promoter During G1/S Phase Transition. *J. Cell. Physiol.* **163**, 636-644.

Good, L. F. & Chen, K. Y. (1996). Cell Cycle- and Age-Dependent Transcriptional Regulation of Human Thymidine Kinase Gene: The Role of NF-Y in the CBP/tk Binding Complex. *Biol. Signals* **5**, 163-169.

Gower, W. R., Jr., Carr, M. & Ives D. (1979). Deoxyguanosine Kinase. Distinct Molecular forms in Mitochondria and Cytosol. *J. Biol. Chem.* **254**, 2180-2183.

Grim, J. E. & Clurman, B. E. (2003). Cycling Without CDK2? *Trends Cell Biol.* **13**, 396-399.

Grodberg, J. & Dunn, J. J. (1988). *ompT* Encodes the *Escherichia coli* Outer Membrane Protease That Cleaves T7 RNA Polymerase during Purification. *J. Bacteriol.* **170**, 1245-1253.

Gronenborn, A. M. & Clore, G. M. (1993). Identification of the Contact Surface of a Streptococcal Protein G Domain Complexed with a Human Fc Fragment. *J. Mol. Biol.* **233**, 331-335.

Grosovsky, A. J., Walter, B. N. & Giver, C. R. (1993). DNA-sequence Specificity of Mutations at the Human Thymidine Kinase Locus. *Mutat. Res.* **289**, 231-243.

Gross, M. K. & Merrill, G. F. (1988). Regulation of Thymidine Kinase Protein Levels During Myogenic Withdrawal from the Cell Cycle Is Independent of mRNA Regulation. *Nucleic Acids Res.* **16**, 11625-11643.

Gross, M. K. & Merrill, G. F. (1989). Thymidine Kinase Synthesis Is Repressed in Nonreplicating Muscle Cells by a Translational Mechanism that Does not Affect the Polysomal Distribution of Thymidine Kinase mRNA. *Proc. Natl. Acad. Sci. USA* **86**, 4987-4991.

Gudas, J. M., Knight, G. B. & Pardee, A. B. (1988). Nuclear Posttranscriptional Processing of Thymidine Kinase mRNA at the Onset of DNA Synthesis. *Proc. Natl. Acad. Sci. USA* **85**, 4705-4709.

Gudas, J. M., Fridovich-Keil, J. L. & Pardee, A. B. (1993). Posttranscriptional Control of Thymidine Kinase Messenger RNA Accumulation in Cells Released from G0-G1 Phase Blocks. *Cell Growth & Differ.* **4**, 421-430.

Haidweger, E., Novy, M. & Rotheneder, H. (2001). Modulation of Sp1 Activity by Cyclin A/CDK Complex. *J. Mol. Biol.* **306**, 201-212.

Hallek, M., Wanders, L., Strohmeyer, S. & Emmerich, B. (1992). Thymidine Kinase: A Tumor Marker with Prognostic Value for Non-Hodgkin's Lymphoma and a Broad Range of Potential Clinical Applications. *Annal. Hematol.* **65**, 1-5.

Hardy, L. W., Finer-Moore, J. S., Montfort, W. R., Jones, M. O., Santi, D. V. & Stroud, R. M. (1987). Atomic Structure of Thymidylate Synthase: Target for Rational Drug Design. *Science* **235**, 448-455.

Harper, J. W., Adami, G. R., Wel, N., Keyomarsi, K. & Elledge, S. J. (1993). The p21 Cdk-Interacting Protein Cip1 Is a Potent Inhibitor of G1 Cyclin-Dependent Kinases. *Cell* **75**, 805-816.

Harper, J. W. (2002). A Phosphorylation-Driven Ubiquitination Switch for Cell-Cycle Control. *Trends Cell Biol.* **12**, 104-107.

Harper, J. W., Burton, J. L. & Solomon, M. J. (2002). The Anaphase-Promoting Complex: It's Not Juast for Mitosis Any More. *Genes Dev.* **16**, 2179-2206.

Harrison, P. T., Thompson, R. & Davison, A. J. (1991). Evolution of Herpesvirus Thymidine Kinase from Cellular Deoxycytidine Kinase. *J. Gen. Virol.* **72**, 2583-2586.

Hartmann-Petersen, R., Seeger, M. & Gordon, C. (2003). Transferring Substrates to the 26S Proteasome. *Trends Biochem. Sci.* **28**, 26-31.

Hatamochi, A., Golumbek, P. T., Van Shaftingen, E. & de Crombrugghe, B. (1988). A CCAAT DNA Binding Factor Consisting of Two Different Components that Are Both Required for DNA Binding. *J. Biol. Chem.* **263**, 5940-5947.

Hatzis, P., Al-Madhoon, A. S., Jüllig, M., Petrakis, T. G., Eriksson, S. & Talianidis, I. (1998). The Intracellular Localization of Deoxycytidine Kinase. *J. Biol. Chem.* **273**, 30239-30243.

Hengstschläger, M., Denk, C. & Wawra, E. (1993). Cell Cycle Regulation of Deoxycytidine Kinase. Evidence for Post-transcriptional Control. *FEBS Lett.* **321**, 237-240.

Hengstschläger, M., Müllner, E. W. & Wawra, E. (1994a). Thymidine Kinase Is Expressed Differently in Transformed Versus Normal Cells: A Novel Test for Malignancy. *Int. J. Oncol.* **4**, 207-210.

Hengstschläger, M., Mudrak, I., Wintersberger, E. & Wawra, E. (1994b). A Common Regulation of Genes Encoding Enzymes of the Deoxynucleotide Metabolism Is Lost after Neoplastic Transformation. *Cell Growth & Differ.* **5**, 1389-1394.

Hengstschläger, M., Knöfler, M., Müllner, E. W., Ogris, E., Wintersberger, E. & Wawra, E. (1994c). Different Regulation of Thymidine Kinase During the Cell Cycle of Normal *Versus* DNA Tumor Virus-transformed Cells. *J. Biol. Chem.* **269**, 13836-13842.

Hengstschläger, M., Pusch, O., Henstschläger-Ottnad, E., Ambros, P. F., Bernaschek, G. & Wawra, E. (1996). Loss of the p16/MTS1 Tumor Suppressor Gene Causes E2F-Mediated Deregulation of Essential Enzymes of the DNA Precursor Metabolism. *DNA Cell Biol.* **15**, 41-51.

Higgins, D. G., Thompson, J. D. & Gibson, T. J. (1996). Using CLUSTAL for Multiple SequenceAlignments. *Methods Enzymol.* **266**, 383- 402.

Hiraga, S., Igarashi, K. & Yura, T. (1967). A Deoxythymidine Kinase-Deficient Mutant of *Escherichia coli*. I. Isolation and Some Properties. *Biochim. Biophys. Acta* **145**, 41-51.

Hirano, M., Silvestri, G., Blake, D. M., Lombes, A., Minetti, C., Bonilla, E., Hays, A. P., Lovelace, R. A., Butler, I., Bertorini, T. E., Threlkeld, A. B., Mitsumoto, H., Salberg, L., Rowland, L. P. & DiMauro, S. (1994). Mitochondrial Neurogastrointestinal Encephalomyopathy (MNGIE): Clinical,

Biochemical and Genetic Features of an Autosomal Recessive Mitochondrial Disorder. *Neurology* **44**, 721-727.

Hofbauer, R. & Denhardt, D. T. (1991). Cell Cycle-Regulated and Proliferation Stimulus-Responsive Genes. *Crit. Rev. Eukaryotic Gene Expression* **1**, 247-300.

Holden, L., Hoffbrand, A. V. & Tattersall, M. H. N. (1980). Thymidine concentration in human sera: Variations in Patients with Leukemia and Megaloblastic Anaemia. *Eur. J. Cancer* **16**, 115-121.

Honda, R. & Yasuda, H. (1999). Association of p19^{ARF} with Mdm2 Inhibits Ubiquitin Ligase Activity of Mdm2 for tumor suppressor p53. *EMBO J.* **18**, 22-27.

Höglund, L., Pontis, E. & Reichard, P. (1988). Effects of Deoxycytidine and Thymidine Kinase Deficiency on Substrate Cycles between Deoxyribonucleosides and Their 5'-Phosphates. *Cancer Res.* **48**, 3681-3687.

Huang, D. Y. & Chang, Z. F. (2001). Interaction of Human Thymidine Kinase 1 with p21^{Waf1}. *Biochem. J.* **356**, 829-834.

Hughes, T. L., Hahn, T. M., Reynolds, K. K. & Shewach, D. S. (1997). Kinetic Analysis of Human Deoxycytidine Kinase with the True Phosphate Donor Uridine Triphosphate. *Biochemistry* **36**, 7540-7547.

Hunsucker, S. A., Spychala, J. & Mitchell, B. (2001). Human Cytosolic 5'-Nucleotidase I. Characterization and Role in Nucleoside Analog Resistance. *J. Biol. Chem.* **276**, 10498-10504.

Hyland, P. L., Keegan, A. L., Curran, M. D., Middleton, D., McKenna, P. G. & Barnett, Y. A. (2000). Effect of a dCTP:dTTP Pool Imbalance on DNA Replication Fidelity in Friend Murine Erythroleukemia Cells. *Environ. Mol. Mutagen.* **36**, 87-96.

Igarashi, K., Hiraga, S. & Yura, T. (1967). A Deoxythymidine Kinase Deficient Mutant of *Escherichia coli*. II. Mapping and Transduction Studies with Phage Φ80. *Genetics* **57**, 643-654.

Ingemarson, R. & Thelander, L. (1996). A Kinetic Study on the Influence of Nucleoside Triphosphate Effectors on Subunit Interaction in Mouse Ribonucleotide Reductase. *Biochemistry* **35**, 8603-8609.

Ito, M. & Conrad, S. E. (1990). Independent Regulation of Thymidine Kinase mRNA and Enzyme Levels in Serum-stimulated Cells. *J. Biol. Chem.* **265**, 6954-6960.

Ives, D. H. & Durham, J. P. (1970). Deoxycytidine Kinase. III. Kinetics and Allosteric Regulation of the Calf Thymus Enzyme. *J. Biol. Chem.* **245**, 2285-2294.

Janin, J., Miller, S. & Chothia, C. (1988). Surface, Subunit Interfaces and Interior of Oligomeric Proteins. *J. Mol. Biol.* **204**, 155-164.

Janse, D., Crosas, B., Finley, D. & Church, G. M. (2004). Localization to the Proteasome Is Sufficient for Degradation. *J. Biol. Chem.* **279**, 21415-21420.

Jensen, H. K. (1994). Studies of the Cytosolic Thymidine Kinase in Human Cells and Comparison to the Recombinantly Expressed Enzyme. Ph.D. Thesis, Roskilde University.

Jensen, H. K. & Munch-Petersen, B. (1995). Altered Kinetic Properties of Recombinant Human Cytosolic Thymidine Kinase (TK1) as Compared with the Native Form. *Adv. Exp. Med. Biol.* **370**, 637-640.

Jinno, S., Suto, K., Nagata, A., Igarashi, M., Kanaoka, Y., Nojima, H. & Okayama, H. (1994). Cdc25A is a Novel Phosphatase Functioning Early in the Cell Cycle. *EMBO J.* **13**, 1549-1556.

Johansson, K., Ramaswamy, S., Ljungcrantz, C., Knecht, W., Piškur, J., Munch-Petersen, B., Eriksson, S. & Eklund, H. (2001). Structural Basis for Substrate Specificities of Cellular Deoxyribonucleoside Kinases. *Nature Struct. Biol.* **8**, 616-620.

Johansson, M. & Karlsson, A. (1996). Cloning and Expression of Human Deoxyguanosine Kinase cDNA. *Proc. Natl. Acad. Sci. USA* **93**, 7258-7262.

Johansson, M., Bajalica-Lagercrantz, S., Lagercrantz, J. & Karlsson, A. (1996). Localization of the Human Deoxyguanosine Kinase Gene (DGUOK) to Chromosome 2p13. *Genomics* **38**, 450-451.

Johansson, M. & Karlsson, A. (1997a). Cloning of the cDNA and Chromosome Localization of the Gene for Human Thymidine Kinase 2. *J. Biol. Chem.* **272**, 8454-8458.

Johansson, M., Brismar, S. & Karlsson, A. (1997b). Human Deoxycytidine Kinase is Located in the Cell Nucleus. *Proc. Natl. Acad. Sci. USA* **94**, 11941-11945.

Johansson, M., Norda, A. & Karlsson, A. (2000). Conserved Gene Structure and Transcription Factor Sites in the Human and Mouse Deoxycytidine Kinase Genes. *FEBS Lett.* **487**, 209-212.

Johnson, L. F., Rao, L. G. & Muench, A. J. (1982). Regulation of Thymidine Kinase Enzyme Level in Serum-Stimulated Mouse 3T6 Fibroblasts. *Exp. Cell Res.* **138**, 79-85.

Jones, D. H. & Howard B.H. (1991). A Rapid Method for Recombination and Site-Specific Mutagenesis by Placing Homologous Ends on DNA Using Polymerase Chain Reaction. *Biotechniques* **10**, 62-66.

Jones, D. H. & Winistorfer, S. C. (1992). Recombinant Circle PCR and Recombination PCR for Site-Specific Mutagenesis without PCR Product Purification. *Biotechniques* **12**, 528-535.

Jones, S. & Thornton, J. M. (1995). Protein-Protein Interactions: A Review of Protein Dimer Structures. *Progr. Biophys. Mol. Biol.* **63**, 31-65.

Jones, S. & Thornton, J. M. (1996). Principles of Protein-Protein Interactions. *Proc. Natl. Acad. Sci. USA* **93**, 13-20.

Jones, S. & Thornton, J. M. (1997). Analysis of Protein-Protein Interaction Sites Using Surface Patches. *J. Mol. Biol.* **272**, 121-132.

Jordan, A. & Reichard, P. (1998). Ribonucleotide Reductases. *Annu. Rev. Biochem.* **67**, 71-98.

Julias, J. G. & Pathak, V. K. (1998). Deoxyribonucleoside Triphosphate Pool Imbalances In Vivo Are Associated with an Increased Retroviral Mutation Rate. *J. Virol.* **72**, 7941-7949.

Jüllig, M. & Eriksson, S. (2000). Mitochondrial and Submitochondrial Localization of Human Deoxyguanosine Kinase. *Eur. J. Biochem.* **267**, 5466-5472.

Jüllig, M. & Eriksson, S. (2001). Apoptosis Induces Efflux of the Mitochondrial Matrix Enzyme Deoxyguanosine Kinase. *J. Biol. Chem.* **276**, 24000-24004.

Kabsch, W. & Sander C. (1983). How Good Are Predictions of Protein Secondary Structure? *FEBS Letters* **155**, 179-182.

Kadonaga, J. T., Carner, K. R., Masiarz, F. R. & Tjian, R. (1987). Isolation of cDNA Encoding Transcription Factor Sp1 and Functional Analysis of the DNA Binding Domain. *Cell* **51**, 1079-1090.

Kaelin, Jr., W. G. (1999). Functions of the Retinoblastoma Protein. *BioEssays* **21**, 950-958.

Kamijo, T., Zindy, F., Roussel, M. F., Quelle, D. E., Downing, J. R., Ashmun, R. A., Grosveld, G. & Sherr, C. J. (1997). Tumor Suppression at the Mouse *INK4a* Locus Mediated by the Alternative Reading Frame Product p19^{ARF}. *Cell* **91**, 649-659.

Kamijo, T., Weber, J. D., Zambetti, G., Zindy, F., Roussel. M. F. & Sherr, C. J. (1998). Functional and Physical Interactions of the ARF Tumor Suppressor with p53 and Mdm2. *Proc. Natl. Acad. Sci. USA* **95**, 8292-8297.

Kaptein, R., Boelens, R. & Koning, T. M. G. (1991). NMR Studies of Proteins, Nucleic Acids and Their Interactions. In: NMR and Biomolecular Structure. Edited by Bertani, I., Molinari, H. & Nicolai, N. VCH Weinheim.

Karayan, E. B., Seite, P., Brambilla, C., Brambilla, E., Larsen, C. J. & Gazzeri, S. (2001).Human ARF binds E2F1 and Inhibits Its Transcriptional Activity. *Oncogene* **10**, 1033-1041.

Karin, M., Liu, Z. G. & Zandi, E. (1997). AP-1 Function and Regulation. *Curr. Opin. Cell Biol.* **9**, 240-246.

Karlseder, J., Rotheneder, H. & Wintersberger, E. (1996). Interaction of Sp1 with the Growth- and Cell Cycle-Regulated Transcription Factor E2F. *Mol. Cell. Biol.* **16**, 1659-1667.

Kashlan, O. B., Scott, C. P., Lear, J. D. & Cooperman, B. S. (2002). A Comprehensive Model for the Allosteric Regulation of Mammalian Ribonucleotide Reductase. Functional Consequences of ATP- and dATP-Induced Oligomerization of the Large Subunit. *Biochemistry*, **41**, 462-474.

Kashlan, O. B. & Cooperman, B. S. (2003). Comprehensive Model for Allosteric Regulation of Mammalian Ribonucleotide Reductase: Refinements and Consequences. *Biochemistry*, **42**, 1696-1706.

Kato, J., Matsushime, H., Hiebert S., W., Ewen, M. & Sherr C. J. (1993). Direct Binding of Cyclin D to the Retinoblastoma Gene Product (pRb) and pRb Phosphorylation by the Cyclin D-dependent Kinase CDK4. *Genes Dev.* **7**, 331-342.

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**, 2538-2546.

Kauffman, M. G., Rose, P. A. & Kelly, T. J. (1991). Mutations in the Thymidine Kinase Gene that Allow Expression of the Enzyme in Quiescent (G0) Cells. *Oncogene* 6, 1427-1435.

Ke, P. Y., Yang, C. C., Tsai, I. C. & Chang, Z. F. (2003). Degradation of Human Thymidine Kinase Is Dependent on Serine-13 Phosphorylation: Involvement of the SCF-mediated Pathway. *Biochem. J.* 370, 265-273.

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. Cel. Biol.* 24, 514-526.

Kerkhoff, E. & Rapp, U. R. (1998). Cell Cycle Targets of Ras/Raf Signalling. *Oncogene* 17, 1457-1462.

Keszler, G., Szikla, K., Kazimierczuk, Z., Spasokoukotskaja, T., Sasvari-Szekely, M. & Staub, M. (2003). Selective Activation of Deoxycytidine Kinase by Thymidine-5′-thiosulphate and Release by Deoxycytidine in Human Lymphocytes. *Biochem. Pharmacol.* 65, 563-571.

Keszler, G., Spasokoukotskaja, T., Csapo, Z., Talianidis, I., Eriksson, S., Staub, M. & Sasvari-Szekely M. (2004). Activation of Deoxycytidine Kinase in Lymphocytes Is Calcium Dependent and Involves a Conformational Change Detectable by Native Immunostaining. *Biochem. Pharmacol.* 67, 947-955.

Khokhlatchev, A. V., Canagarajah, B., Wilsbacher, J., Robinson, M., Atkinson, M., Goldsmith, E. & Cobb, M. H. (1998). Phosphorylation of the MAP Kinase ERK2 Promotes its Homodimerization and Nuclear Translocation. *Cell* 93, 605-615.

Kim, C. H., Oh, Y. & Lee, T. H. (1997). Codon Optimization for High-level Expression of Human Erythropoietin (EPO) in Mammalian Cells. *Gene (Amst.)* 199, 293-301.

Kim, E. C., Rawlings, S. L., Li, L. J., Roy, B. & Lee, A. S. (1996). Identification of a Set of Protein Species Approximately 40 kDa as High-Affinity DNA Binding Factor(s) to the Cell Cycle Regulatory Region of the Human Thymidine Kinase Promoter. *Cell Growth Differentiation* 7, 1741-1749.

Kim, E. C., Lau, J. S., Rawlings, S. & Lee, A. S. (1997a). Positive and Negative Regulation of the Human Thymidine Kinase Promoter Mediated by CCAAT Binding Transcription Factors NF-Y/CBF, dbpA, and CDP/cut. *Cell Growth Differ.* 8, 1329-1338.

Kim, H. J., Nishikawa, S., Tokutomi, Y., Takenaka, H., Hamada, M., Kuby, S. A. & Uesugi, S. (1990). In Vitro Mutagenesis Studies at the Arginine Residues of Adenylate Kinase. A Revised Binding Site for AMP in the X-ray Deduced Model. *Biochemistry* 29, 1107-1111.

Kim, Y. K., Wells, S., Lau, Y. F. C. & Lee, A. S. (1988). Sequences Contained Within the Promoter of the Human Thymidine Kinase Gene Can Direct Cell-Cycle Regulation of Heterologous Fusion Genes. *Proc. Natl. Acad. Sci.* 85, 5894-5898.

Kim, Y. K. & Lee, A. S. (1991). Identification of a 70-Base-Pair Cell Cycle Regulatory Unit Within the Promoter of the Human Thymidine Kinase Gene and Its Interaction with Cellular Factors. *Mol. Cell. Biol.* 11, 2296-2302.

Kim, Y. K. & Lee, A. S. (1992). Identification of a Protein-binding Site in the Promoter of the Human Thymidine Kinase Gene Required for the G1-S-regulated Transcription. *J. Biol. Chem.* **267**, 2723-2727.

Kitagawa, M., Higashi, H., Jung, H. K., Suzuki-Takahashi, I., Ikeda, M., Tamai, K., (1996). The Consensus Motif for Phosphorylation by Cyclin D1-CDK4 is Different from that for Phosphorylation by Cyclin A/E-CDK2. *EMBO J.* **15**, 7060-7069.

Knecht, W., Petersen, G. E., Munch-Petersen, B. & Piškur, 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**, 529-540.

Knecht, W., Sandrini, M. P. B., Johansson, K., Eklund, H., Munch-Petersen, B. & Piškur, J. (2002b). A Few Amino Acid Substitutions Can Convert Deoxyribonucleoside Kinase Specificity from Pyrimidines to Purines. *EMBO J.* **21**, 1873-1880.

Knight, G. B., Gudas, J. M. & Pardee, A. B. (1987). Cell-Cycle-Specific Interaction of Nuclear DNA-Binding Proteins with a CCAAT Sequence from the Human Thymidine Kinase Gene. *Proc. Natl. Acad. Sci. USA* **84**, 8350-8354.

Kornberg, A. & Baker, T. A. (1992). *DNA Replication.* W.H. Freeman and Company: New York.

Koury, M. J. & Horne, D. W. (1994). Apoptosis Mediates and Thymidine Prevents Erythroblast Destruction in Folate Deficiency Anemia. *Proc. Natl. Acad. Sci. USA* **91**, 4076-4071.

Krawiec, K., Kierdaszuk, B., Eriksson, S., Munch-Petersen, B. & Shugar, D. (1995). Nucleoside Triphosphate Donors for Nucleoside Kinases: Donor Properties of UTP with Human Deoxycytidine Kinase. *Biochem Biophys. Res. Commun.* **216**, 42-48.

Krawiec, K., Kierdaszuk, B. & Shugar, D. (2003). Inorganic Tripolyphosphate (PPP$_i$) as a Phosphate Donor for Human Deoxyribonucleoside Kinases. *Biochem. Biophys. Res. Commun.* **301**, 192-197.

Kreidberg, J. A. & Kelly, T. J. (1986). Genetic Analysis of the Human Thymidine Kinase Gene Promoter. *Mol. Cell. Biol.* **6**, 2903-2909.

Kristensen, T., Jensen, H. K. & Munch-Petersen, B. (1994). Overexpression of Human Thymidine Kinase mRNA Without Corresponding Enzymatic Activity in Patients with Chronic Lymphatic Leukemia. *Leukemia Res.* **18**, 861-866.

Kristensen, T. (1996). Quantification of Thymidine Kinase (TK1) mRNA in Normal and Leukemic Cells and Investigation of Structure-Function Relationship of Recombinant TK1 Enzyme. Ph.D. Thesis, Roskilde University.

Kunz, B. A. (1988). Mutagenesis and Deoxyribonucleotide Pool Imbalance. *Mutat. Res.* **200**, 133-147.

Kunz, B. A. & Kohalmi, S. E. (1991). Modulation of Mutagenesis by Deoxyribonucleotide Levels. *Annu. Rev. Genet.* **25**, 339-359.

Kunz, B. A., Kohalmi, S. E., Kunkel, T. A., Mathews, C. K., McIntosh, E. M. & Reidy, J. A. (1994). *Mutation Res.* **318**, 1-64.

194

Kussmann-Gerber, S., Kuonen, O., Folkers, G., Pilger, B. D. & Scapozza, L. (1998). Drug Resistance of Herpes Simplex Virus Type 1. Structural Considerations at the Molecular Level of the Thymidine Kinase. *Eur. J. Biochem.* **255**, 472-481.

LaBaer, J., Garrett, M. D., Stevenson, L. F., Slingerland, J. M., Sandhu, C., Chou, H. S., Fattaey, A. & Harlow, E. (1997). New Functional Activities for the p21 Family of CDK Inhibitors. *Genes & Dev.* **11**, 847-862.

Lai, Y., Tse, C. M. & Unadkat, J. D. (2004). Mitochondrial Expression of the Human Equilibrative Nucleoside Transporter 1 (hENT1) Results in Enhanced Mitochondrial Toxicity of Antiviral Drugs. *J. Biol. Chem.* **279**, 4490-4497.

Lavie, A., Vetter, I. R., Konrad, M., Goody, R. S., Reinstein, J. & Schlichting, I. (1997). Structure of Thymidylate Kinase Reveals the Cause Behind the Limiting Step in AZT Activation. *Nature Struct. Biol.* **4**, 601-604.

Lavie, A., Konrad, M., Brundiers, R., Goody, R. S., Schlichting, I. & Reinstein, J. (1998). Crystal Structure of Yeast Thymidylate Kinase Complexed with the Bisubstrate Inhibitor P^{1}-(5'-Adenosyl) P^{5}-(5'-Thymidyl) Pentaphosphate (TP$_5$A) at 2.0 Å Resolution: Implications for Catalysis and AZT Activation. *Biochemistry* **37**, 3677-3686.

Lee, L. S. & Cheng, Y. C. (1976). Human Deoxythymidine Kinase. I. Purification and General Properties of the Cytoplasmic and Mitochondrial Isozymes Derived from Blast Cells of Acute Myelocytic Lekemia. *J. Biol. Chem.* **251**, 2600-2604.

Leeds, J. M., Slabaugh, M. B. & Mathews, C. K. (1985). DNA Precursor Pools and Ribonucleotide Reductase Activity: Distribution between the Nucleus and Cytoplasm of Mammalian Cells. *Mol. Cell. Biol.* **5**, 3443-3450.

Leeds, J. M. & Mathews, C. K. (1987). Cell Cycle- Dependent Effects on Deoxyribonucleotide and DNA Labeling by Nucleoside Precursors in Mammalian Cells. *Mol. Cell. Biol.* **7**, 532-534.

Lei, H. X., Li, Z. S., Xie, D. Y., Liu, B. C., Zuo, J. & Fang, F. D. (1998a). Role of dNTP in Mutagenesis. *Biomed. Environ. Sci.* **11**, 345-353.

Lei, H. X., Li, Z. S., Xie, D. Y., Liu, B. C. & Fang, F. D. (1998b). Role of Deoxyribonucleoside Triphosphates (dNTPs) in Cell Transformation. *Biomed. Environ. Sci.* **11**, 354-362.

Leone, G., De Gregori, J., Sears, R., Jakoi, L. & Nevins, J. R. (1997). Myc and Ras Collaborate in Inducing Accumulation of Active Cyclin E/CDK2 and E2F. *Nature* **387**, 422-426.

Leoni, L. M., Chao, Q., Cottam, H. B., Genini, D., Rosenbach, M., Carrera, C. J., Budihardjo, I., Wang, X. & Carson, D. A. (1998). Induction of an Apoptotic Program in Cell-Free Extracts by 2-Chloro-2'-deoxyadenosine 5'-triphosphate and cytochrome c. *Proc. Natl. Acad. Sci. USA* **95**, 9567-9571.

Lesk, A. M. & Chothia, C. (1980). How Different Amino Acid Sequences Determine Similar Protein Structures: The Structure and Evolutionary Dynamics of the Globins. *J. Mol. Biol.* **136**, 225-270.

Letunic, I., Goodstadt, L., Dickens, N. J., Doerks, T., Schultz, J., Mott, R., Ciccarelli, F., Copley, R. R., Ponting, C. P. & Bork, P. (2002). Recent Improvements to the SMART Domain-Based Sequence Annotation Resource. *Nucleic Acids Res.* **30**, 242-244.

Levine, A. J. (1997). p53, the Cellular Gatekeeper for Growth and Division. *Cell* **88**, 323-331.

Li, C. L., Lu, C. Y., Ke, P. Y. & Chang, Z. F. (2004). Perturbation of ATP-induced Tetramerization of Human Cytosolic Thymidine Kinase by Substitution of Serine-13 with Aspartic Acid at the Mitotic Phosphorylation Site. *Biochem. Biophys. Res. Commun.* **313**, 587-593.

Li, L. J., Naeve, G. S. & Lee, A. S. (1993). Temporal Regulation of Cyclin A-p107 and p33^{CDK2} Complexes Binding to a Human Thymidine Kinase Promoter Element Important for G1-S Phase Transcriptional Regulation. *Proc. Natl. Acad. Sci. USA* **90**, 3554-3558.

Li, P., Li, C., Zhao, X., Zhang, X., Nicosia, S. V. & Bai, W. (2004a). p27^{Kip1} Stabilization and G1 Arrest by 1,25-Dihydroxyvitamin D$_3$ in Ovarian Cancer Cells Mediated through Down-regulation of Cyclin E/Cyclin-dependent Kinase 2 and Skp1-CullinF-box Protein/Skp2 Ubiquitin Ligase. *J. Biol. Chem.* **279**, 25260-25267.

Li, S. & Deber, C. M. (1994). A Measure of Helical Propensity for Amino Acids in Membrane Environments. *Nature Struct. Biol.* **1**, 368-373.

Lim, V. I. (1974a). Structural Principles of the Globular Organization of Protein Chains. A Stereochemical Theory of Globular Protein Secondary Structure. *J. Mol. Biol.* **88**, 857-872.

Lim, V. I. (1974b). Algorithms for Prediction of alpha-helical and beta-structural Regions in Globular Proteins. *J. Mol. Biol.* **88**, 873-94.

Lin, S. L., Tsai, C. J. & Nussinov, R. (1995). A Study of Four-helix Bundle: Investigating Protein Folding *via* Similar Architectural Motifs in Protein Cores and Subunit Interfaces. *J. Mol. Biol.* **248**, 151-161.

Linding, R., Russell, R. B., Neduva, V. & Gibson, T. J. (2003). GlobPlot: Exploring Protein Sequences for Globularity and Disorder. *Nucl. Acids Res.* **31**, 3701-3708.

Lipson, K. E. & Baserga, R. (1989). Transcriptional Activity of the Human Thymidine Kinase Gene Determined by a Method Using the Polymerase Chain Reaction and an Intron-Specific Probe. *Proc. Natl. Acad. Sci. USA* **86**, 9774-9777.

Liu, Q. & Summers, W. C. (1988). Site-Directed Mutagenesis of a Nucleotide-Binding Domain in HSV-1 Thymidine Kinase: Effects on Catalytic Activity. *Virology* **163**, 638-642.

Lozano, G. & Elledge, S. J. (2000). p53 Sends Nucleotides to Repair DNA. *Nature* **404**, 24-25.

Lukas, C., Sørensen, C. S., Kramer, E., Santoni-Rugiu, E., Lindeneg, C., Peters, J. M., Bartek, J. & Lukas, J. (1999). Accumulation of Cyclin B1 Requires E2F and Cyclin-A-Dependent Rearrangement of the Anaphase-Promoting Complex. *Nature* **401**, 815-818.

Luo, R. X., Postigo, A. A. & Dean, D. C. (1998). Rb Interacts with Histone Deacetylase to Repress Transcription. *Cell* **92**, 463-473.

Mailly, F., Bérubé, G., Harada, R., Mao, P. L., Phillips, S., Nepveu, A. (1996). The Human Cut Homeodomain Protein Can Repress Gene Expression by Two Distinct Mechanisms: Active Repression and Competition for Binding Site Occupancy. *Mol. Cell. Biol.* **16**, 5346-5357.

Maity, S. N., de Crombrugghe, B. (1998). Role of the CCAAT-Binding Protein CBF/NF-Y in Transcription. *Trends Biochem. Sci.* **23**, 174-178.

Mancuso, M., Salviati, L., Sacconi, S., Otaegui, D., Camano, P., Marina A., Bacman, S., Moraes C. T., Carlo, J. R., Garcia, M., Garcia-Alvarez, M., Monzo-Naini, A. B., Hirano, M., Bonilla, E., Taratuto, A. L., DiMauro, S. & Vu, T. H. (2002). Mitochondrial DNA Depletion: Mutations in Thymidine Kinase Gene with Myopathy and SMA. *Neurology* **59**, 1197-1202.

Mancuso, M., Filosto, M., Tsujino, S., Lamperti, C., Shanske, S., Coquet, M., Desnuelle, C. & DiMauro, S. (2003). Muscle Glycogenosis and Mitochondrial Hematopathy in an Infant with Mutations in both the Myophosphorylase and Deoxyguanosine Kinase Genes. *Arch. Neurol.* **60**, 1445-1447. (Erratum in *Arch. Neurol.* 61, 29).

Mandel, H., Szargel, R., Labay, V., Elpeleg, O., Saada, A., Shalata, A., Anbinder, Y., Berkowitz, D., Hartman, C., Barak, M., Eriksson, S. & Cohen, N. (2001). The Deoxyguanosine Kinase Gene Is Mutated in Individuals with Depleted Hepatocerebral Mitochondrial DNA. *Nature Genet.* **29**, 337-341.

Mann, G. J., Musgrove, E. A., Fox, R. M. & Thelander L. (1988). Ribonucleotide Reductase M1 Subunit in Cellular Proliferation, Quiescence, and Differentiation. *Cancer Research* **48**, 5151-5156.

Mantovani, R., Pessara, U., Tronche, F., Li, X. Y., Knapp, A. M., Pasquali, J. L., Benoist, C. & Mathis D. (1992). Monoclonal Antibodies to NF-Y Define Its Function in MHC Class II and Albumin Gene Transcription. *EMBO J.* **11**, 3315-3322.

Mao, X., Xia, L., Liang, G., Gai, X., Huang, D. Y., Prystowski, M. B. & Lipson, K. E. (1995). CCAAT-Box Contributions to Human Thymidine Kinase mRNA Expression. *J. Cell. Biochem.* **57**, 701-710.

Mao, Y., Wu, J., Wang, N., He, L., Wu, C., He, Q. & Skog S. (2002). A Comparative Study: Immunohistochemical Detection of CytosolicThymidine Kinase and Proliferating Cell Nuclear Antigen in Breast Cancer. *Cancer Invest.* **20**, 922-931.

Marston, F. A. O. (1986). The Purification of Eukaryotic Polypeptides Synthesized in *Escherichia coli. Biochem. J.* **240**, 1-12.

Martelli, F. & Livingston, D. M. (1999). Regulation of Endogenous E2F1 Stability by the Retinoblastoma Family Proteins. *Proc. Natl. Acad. Sci. USA* **96**, 2858-2863.

Marti, R., Nishigaki, Y. & Hirano, M. (2003). Elevated Plasma Deoxyuridine in Patients with Thymidine Phosphorylase Deficiency. *Biochem. Biophys. Res. Commun.* **303**, 14-18.

Martinez, A. M., Afshar, M., Martin, F., Cavadore, J. C., Labbé, J. C. Dorée, M. (1997). Dual Phosphorylation of the T-Loop in CDK7: Its Role in Controlling Cyclin H binding and CAK Activity. *EMBO J.* **16**, 343-354.

Mathews, C. K. & Ji, J. (1992). DNA Precursor Asymmetries, Replication Fidelity, and Variable Genome Evolution. *BioEssays* **14**, 295-301.

Matthews, B. W. & Rossmann, M. G. (1985). Comparison of Protein Structure. *Methods Enzymol.* **115**, 397-420.

May, P. & May, E. (1999). Twenty Years of p53 Research: Structural and Functional Aspects of the p53 Protein. *Oncogene* **18**, 7621-7636.

McDougall, J. K., Kucherlapati, , R. & Ruddle, F.H. (1973). Localization and Induction of the Human Thymidine Kinase by Adenovirus 12. *Nature New Biol.* **245**, 172-175.

McKenna, P. G., Wilkinson, Y. A. & McKelvey, V. J. (1991). The Significance of Thymidine Kinase in Determining Resistance to Mutagenesis in Friend Mouse Leukemia Cells. *Anticarcinogen. Radiat. Protec.* **2**, 73-77.

Meuth, M. (1989). The Molecular Basis of Mutations Induced by Deoxyribonucleoside Triphosphate Pool Imbalances in Mammalian Cells. *Exp. Cell Res.* **181**, 305-316.

Michael, M., Fetzer, J. & Folkers, G. (1994). Site-directed Mutagenesis Clarifies the Substrate Position Within the Three-dimensional Model of the Active Site of Herpes Simplex Virus Type-1 Thymidine Kinase. *Eur. J. Biochem.* **226**, 219-226.

Michael, M., Fetzer, J. & Folkers, G. (1995). Site-directed Mutagenesis of Herpes Simplex Virus Type 1 Thymidine Kinase Opposes the Importance of Amino Acid Positions 251, 321 and 348 for Selective Recognition of Substrate Analogs. *Biochem. Biophys. Res. Commun.* **209**, 966-973.

Mikkelsen, N. E., Johansson, K., Karlsson, A., Knecht, W., Andersen, G., Piškur, 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**, 5706-5712.

Mikulits, W. & Müllner, E. W. (1994). Post-Transcriptional Repression of Thymidine Kinase Expression During Cell Cycle and Growth Stimulation. *FEBS Letters* **345**, 172-176.

Mikulits, W., Hengstschläger, M., Sauer, T., Wintersberger, E. & Müllner, E. W. (1996). Overexpression of Thymidine Kinase mRNA Eliminates Cell Cycle Regulation of Thymidine Kinase Enzyme Activity. *J. Biol. Chem.* **271**, 853-860.

Mikulits, W., Knöfler, M., Stiegler, P., Dolznig, H., Wintersberger, E. & Müllner, E. W. (1997). Mouse Thymidine Kinase Stability in Vivo and After in Vitro Translation. *Biochim. Biophys. Acta* **1338**, 267-274.

Miller, O. J., Allderdice, P. W., Miller, D. A., Breg, W. R. & Migeon, B. R. (1971). Human Thymidine Kinase Gene Locus: Assignment to Chromosome 17 in a Hybrid of Man and Mouse Cells. *Science* **173**, 244-245.

Miller, S., Lesk, A. M., Janin J. & Chothia, C. (1987). The Accessible Surface Area and Stability of Oligomeric Proteins. *Nature* **328**, 834-836.

Miller, S. (1989). The Structure of Interfaces Between Subunits of Dimeric and Tetrameric Proteins. *Protein Engineering* **3**, 77-83.

Minton, A. P. (2000). Implications of Macromolecular Crowding for Protein Assembly. *Curr Opin. Struct. Biol.* **10**, 34-39.

Mitchell, B. S., Song, J. J., Johnson, E. E. II, Chen, E. Dayton, J. S. (1993). Regulation of Human Deoxycytidine Kinase Expression. *Advan. Enzyme Regul.* **33**, 61-68.

Mizutani, Y., Wada, H., Yoshida, O., Fukushima, M., Nakao, M. & Miki, T. (2003). Significance of Thymidine Kinase Activity in Renal Cell Carcinoma. *J. Urol.* **169**, 706-709.

Mohapatra, S., Agrawal, D. & Pledger, W. J. (2001). P27 Regulates T Cell Proliferation. *J. Biol. Chem.* **276**, 21976-21983.

Mole, M. L., Hunter, D. H., Gao, P. & Lau, P. (1998). Sample Preparation and High-Performance Liquid Chromatographic Analysis of Deoxyribonucleoside Triphosphates in Individual Rat Embryos. *Anal. Biochem.* **259**, 245-252.

Morgan, D. O. (1997). Cyclin-Dependent Kinases: Engines, Clocks, and Microprocessors. *Annu. Rev. Cell Dev. Biol.* **13**, 261-91.

Morgan, D. O. (1999). Regulation of the APC and the Exit from Mitosis. *Nature Cell Biol.* **1**, E47-E53.

Munch-Petersen, B., Tyrsted, G. & Dupont, B. (1973). The Deoxyribonucleoside 5'-Triphosphate (dATP and dTTP) Pool in Phytohemagglutinin-stimulated and Nonstimulated Human Lymphocytes. *Exptl. Cell Res.* **79**, 249-256.

Munch-Petersen, B. & Tyrsted G. (1977). Induction of Thymidine Kinases in Phytohaemagglutinin-stimulated Human Lymphocytes. *Biochim. Biophys. Acta* **478**, 364-365.

Munch-Petersen, B. (1984). Differences in the Kinetic Properties of Thymidine Kinase Isoenzymes in Unstimulated and Phytohemagglutinin-Stimulated Human Lymphocytes. *Mol. Cell. Biochem.* **64**, 173-185.

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**, 9032-9038.

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**, 15621-15625.

Munch-Petersen, B., Cloos, L., Jensen, H. K. & Tyrsted, G. (1995a). Human Thymidine Kinase 1. Regulation in Normal and Malignant Cells. *Advan. Enzyme Regul.* **35**, 69-89.

Munch-Petersen, B., Tyrsted, G. & Cloos, L., Beck, R. A. & Eger, K. (1995b). Different Affinity of the Two Forms of Human Cytosolic Thymidine Kinase Towards Pyrimidine Analogs. *Biochim. Biophys. Acta* **1250**, 158-162.

Munch-Petersen, B. (1996). DNA Precursor Pool Balance and Thymidine Kinase Isoenzymes in Normal and Malignant Cells. Enzymatic Regulation Mechanisms and Clinical Therapeutic Applications. (Dissertation. Institute of Life Sciences and Chemistry, Roskilde University, Denmark). ISBN 87-7349-325-2

Munch-Petersen, B., Piškur, 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**, 3926-3931.

Munir, K. M., French, D. C., Dube, D. K. & Loeb, L. A. (1992). Permissible Amino Acid Substitutions within the Putative Nucleoside-binding Site of Herpes Simplex Virus Type 1 Established by Random Sequence Mutagenesis. *J. Biol. Chem.* **267**, 6584-6589.

Munir, K. M., French, D. C., Dube, D. K. & Loeb, L. A. (1994). Herpes Thymidine Kinase Mutants with Altered Catalytic Efficiencies Obtained by Random Sequence Selection. *Protein Eng.* **7**, 83-89.

Müller, C. W. & Schulz, G. E. (1992). Structure of the Complex Between Adenylate Kinase from *Escherichia coli* and the Inhibitor Ap_5A Refined at 1.9Å Resolution. A Model for a Catalytic Transition State. *J. Mol. Biol.* **224**, 159-177.

Müller-Dieckmann, H. J. & Schulz, G. E. (1994). The Structure of Uridylate Kinase with Its Substrates, Showing the Transition State Geometry. *J. Mol. Biol.* **236**, 361-367.

Müller-Dieckmann, H. J. & Schulz, G. E. (1995). Substrate Specificity and Assembly of the Catalytic Center Derived from two Structures of Ligated Uridylate Kinases. *J. Mol. Biol.* **246**, 522-530.

Myles, G. M., Hearst, J. E. & Sancar, A. (1991). Site-Specific Mutagenesis of Conserved Residues Within Walker A and B Sequences of *Escherichia coli* UvrA Protein. *Biochemistry* **30**, 324-3834.

Mørch, U. & Petersen P. L. (1997). Betydningen af Methionine som Aminosyre 106 i Thymidin Kinase 1. 2-en og 3-ie modul projekt ved Roskilde Universitetscenter.

Nakagawa K. & Yokosawa, H. (2000). Degradation of Transcription Factor IRF-1 by the Ubiquitin-Proteasome Pathway. *Eur. J. Biochem.* **267**, 1680-1686.

Nevins, J. R. (1992). E2F: A Link Between the Rb Tumor Suppressor Protein and Viral Oncoproteins. *Science* **258**, 424-429.

Nicander, B. & Reichard, P. (1985). Evidence for the Involvement of Substrate Cycles in the Regulation of Deoxyribonucleoside Triphosphate Pools in 3T6 Cells. *J. Biol.Chem.* **260**, 9216-9222.

Nigg, E. A. (1997). Nucleocytoplasmic Transport: Signals, Mechanisms and Regulation. *Nature* **386**, 779-787.

Nishigaki, Y., Marti, R., Copeland, W. C. & Hirano, M. (2003). Site-specific Somatic Mitochondrial DNA Point Mutations in Patients with Thymidine Phosphorylase Deficiency. *J. Clin. Invest.* **111**, 1913-1921.

Nishino, I., Spinazzola, A. & Hirano, M. (1999). Thymidine Phosphorylase Gene Mutations in MNGIE, a Human Mitochondrial Disorder. *Science* **283**, 689-692.

Noble, M. E. M., Endicott, J. A., Brown, N. R. & Johnson, L. N. (1997). The Cyclin Box Fold: Protein Recognition in the Cell-Cycle and Transcription Control. *Trends Biochem. Sci.* **22**, 482-487.

Nordenskjöld, B. A., Skoog,L., Brown, N. C. & Reichard, P. (1970). Deoxyribonucleotide Pools and Deoxyribonucleic Acid Synthesis in Cultured Mouse Embryo Cells. *J. Biol. Chem.* **245**, 5360-5368.

Nordhoff, A., Bücheler, U. S., Wener, D. & Schirmer, R. H. (1993). Folding of the Four Domains and Dimerization Are Impaired by the Gly446 to Glu Exchange in Human Glutathione Reductase. Implications for the Design of Antiparasitic Drugs. *Biochemistry* **32**, 4060-4066.

Obaya, A. J., Mateyak, M. K. & Sedivy, J. M. (1999). Mysterious Liaisons: the Relationship Between c-Myc and the Cell Cycle. *Oncogene* **18**, 2934-2941.

Obaya, A. J., Kotenko, I., Cole, M. D. & Sedivy, J, M. (2002). The Proto-oncogene *c-myc* Acts through the Cyclin-dependent Kinase (Cdk) Inhibitor p27^{KIP1} to Facilitate the Activation of Cdk4/6 and Early G1 Phase Progression. *J. Biol. Chem.* **277**, 31263-31269.

Obaya, A. J. & Sedivy, J. M. (2002). Regulation of Cyclin-Cdk Activity in Mammalian Cells. *CMLC Cell. Mol. Life Sci.* **59**, 126-142.

O'Connor, R. J., Schaley, J. E., Feeney, G. & Hearing P. (2001). The p107 Tumor Suppressor Induces Stable E2F DNA Binding to Repress Target Promoters. *Oncogene* **20**, 1882-91.

Ofran, Y. & Rost, B. (2003). Analysing Six Types of Protein-Protein Interfaces. *J. Mol. Biol.* **325**, 377-387.

Okajima, T., Tanizawa, K. & Fukui, T. (1993a). Site –directed Mutagenesis of AMP-binding Residues in Adenylate Kinase. Alteration of Substrate Specificity. *FEBS Lett.* **334**, 86-88.

Okajima, T., Tanizawa, K. & Fukui, T. (1993b). Site-Directed Mutagenesis of AMP-Binding Residues in Adenylate Kinase. *J. Biochem.* **114**, 627-633.

Okayama, H. & Berg, P. (1983). A cDNA Cloning Vector that Permits Expression of cDNA Inserts in Mammalian Cells. *Mol. Cell Biol.* **3**, 280-289.

Oliver, F. J., Collins, M. K. L. & López-Rivas, A. (1996). Regulation of the Salvage Pathway of Deoxynucleotides Synthesis in Apoptosis Induced by Growth Factor Deprivation. *Biochem J.* **316**, 421-425.

Oliver, F. J., Collins, M. K. L. & López-Rivas, A. (1997). Overexpression of a Heterologous Thymidine Kinase Delays Apoptosis Induced by Factor Deprivation and Inhibitors of Deoxynucleotide Metabolism. *J. Biol. Chem.* **272**, 10624-10630.

O'Neill, K. L., Grigsby, R. V. & Fairbairn, D. W. (1995). Thymidine Kinase: the Future in Breast Cancer Prognosis. *The Breast* **4**, 79-83.

Ostermann, N., Schlichting, I., Brundiers, R., Konrad, M., Reinstein, J., Veit, T., Goody, R. S. & Lavie, A. (2000). Insights into the Phosphoryltransfer Mechanism of Human Thymidylate Kinase Gained from Crystal Structures of Enzyme Complexes Along the Reaction Coordinate. *Structure* **8**, 629-642.

Pai, E. F., Krengel, U., Petsko, G. A., Goody, R. S., Kabsch, W., Wittinghofer, A. (1990). Refined Crystal Structure of the Triphosphate Conformation of H-*ras* p21 at 1.35 Å Resolution: Implications for the Mechanism of GTP Hydrolysis. *EMBO J.* **9**, 2351-2359.

Pai, E. F., Sachsenheimer, W., Schirmer, R. H. & Schulz, G. E. (1977). Substrate Positions and Induced-fit in Crystalline Adenylate Kinase. *J. Mol. Biol.* **114**, 37-45.

Pang, J. H. & Chen, K. Y. (1993). A Specific CCAAT-binding Protein, CBP/tk, May Be Involved in the Regulation of Thymidine Kinase Gene Expression in Human IMR-90 Diploid Fibroblasts During Senescence. *J. Biol. Chem.* **268**, 2909-2916.

Pardee, A. B. (1974). A Restriction Point for Control of Normal Animal Cell Proliferation. *Proc. Natl. Acad. Sci. USA* **71**, 1286-1290.

Pardee, A. B. (1989). G1 Events and Regulation of Cell Proliferation. *Science* **246**, 603-608.

Peng, R., Gordadze, A., V., Panana, E. M. F., Wang, F., Zong, J., Hayward, G. S., Tan, J. & Ling, P. D. (2000). Sequence and Functional Analysis of EBNA-LP and EBNA2 Proteins from Nonhuman Primate Lymphocryptoviruses. *J. Virol.* **74**, 379-389.

Perez-Roger, I., Kim, S. H., Griffiths, B., Sewing, A & Land H. (1999). Cyclins D1 and D2 Mediate Myc-Induced Proliferation via Sequestration of p21 and p27. *EMBO J.* **19**, 5310-5320.

Perozzo, R., Jelesarov, I., Bosshard, H. R., Folkers, G. & Scapozza, L. (2000). Compulsory Order of Substrate Binding to Herpes Simplex Virus Type 1 Thymidine Kinase. A Calorimetric Study. *J. Biol. Chem.* **275**, 16139-16145.

Peter, M. & Herskowitz, I. (1994). Joining the Complex: Cyclin-Dependent Kinase Inhibitory Proteins and the Cell Cycle. *Cell* **79**, 181-184.

Peters, J. M. (1998). SCF and APC: the Yin and Yang of Cell Cycle Regulated Proteolysis. *Curr. Opinion Cell Biol.* **10**, 759-768.

Petersen, T. N., Lundegaard, C., Nielsen M., Bohr, H., Bohr, J., Brunak, S., Gippert, G. P. & Lund, O. (2000). Prediction of Protein Secondary Structure at 80% Accuracy. *Proteins: Struct. Function Genet.* **41**, 17-20.

Pfleger, C. M. & Kirschner, M. W. (2000). The KEN Box: An APC Recognition Signal Distinct from the D Box Targeted by Cdh1. *Genes Dev.* **14**, 655-665.

Pfleger, C. M., Lee, E. & Kirschner, M. W. (2001). Substrate Recognition by the Cdc20 and Cdh1 Components of the Anaphase-Promoting Complex. *Genes Develop.* **15**, 2396-2407.

Phear, G. & Meuth, M. (1989). The Genetic Consequences of DNA Precursor Pool Imbalance: Sequence Analysis of Mutations Induced by Excess Thymidine at the Hamster *aprt* locus. *Mutation Res.* **214**, 201-206.

Pilger, B. D., Perozzo, R., Alber, F., Wurth, C., Folkers, G. & Scapozza, L. (1999). Substrate Diversity of Herpes Simplex Virus Thymidine Kinase. Impact of the Kinematics of the Enzyme. *J. Biol. Chem.* **274**, 31967-31973.

Pines, J. (1999). Checkpoint on the Nuclear Frontier. *Nature* **397**, 104-105.

Plagemann, P. G. W., Wohlhueter, R. M. & Woffendin, C. (1988). Nucleoside and Nucleobase Transport in Animal Cells. *Biochim. Biophys. Acta* **947**, 405-443.

Pontarin, G., Gallinaro, L., Ferraro, P., Reichard, P. & Bianchi, V. (2003). Origins of Mitochondrial Thymidine Triphosphate: Dynamic Relations to Cytosolic Pools. *Proc. Natl. Acad. Sci. USA* **100**, 12159-12164.

Ponting, C. P. & Russell, R. R. (2002). The Natural History of Protein Domains. *Annu. Rev. Biophys. Biomol. Struct.* **31**, 45-71.

Posch, M., Hauser, C. & Seiser, C. (2000). Substrate Binding is a Prerequisite for Stabilization of Mouse Thymidine Kinase in Proliferating Fibroblasts. *J. Mol. Biol.* **300**, 493-502.

Poulsen, F. M., Kjær, M., Thomsen, N. K:, Lerche, M. & Jensen, P. H. (1997). Form og Bevægelse – Orden og Uorden – Struktur og Funktion – i Proteiner. NMR Undersøgelser af Proteiners Rumlige Struktur i Deres Arbejdsmiljø – den Vandige Opløsning. *Naturens Verden* **9**, 321-330.

Prota, A., Vogt, J., Pilger, B., Perozzo, R., Wurth, C., Marquez, V. E., Russ, P., Schulz, G. E., Folkers, G. &' Scapozza, L. (2000). Kinetics and Crystal Structure of the Wild-Type and the Engineered Y101F Mutant of *Herpes simplex* Virus Type 1 Thymidine Kinase Interacting with (North)-methanocarba-thymidine. *Biochemistry* **39**, 9597-9603.

Pugmire, M. J. & Ealick, S. E. (2002). Structural Analyses Reveal Two Distinct Families of Nucleoside Phosphorylases. *Biochem. J.* **361** (Pt 1), 1-25.

Purring-Koch, C. & McLendon, G. (2000). Cytochrome *c* Binding to Apaf-1: The Effects of dATP and Ionic Strength. *Proc. Natl. Acad. Sci. USA* **97**, 11928-11931.

Ramirez-Carrozzi, V. R. & Kerppola, T. K. (2001). Control of the Orientation of Fos-Jun Binding and the Transcriptional Cooperativity of Fos-Jun-NFAT1 Complexes. *J. Biol. Chem.* **276**, 21797-21808.

Rampazzo, C., Johansson, M., Gallinaro, L., Ferraro, P., Hellman, U., Karlsson, A., Reichard, P. & Bianchi, V. (2000a). Mammalian 5'(3')-Deoxyribonucleotidase, cDNA Cloning, and Overexpression of the Enzyme in *Escherichia coli* and Mammalian Cells. *J. Biol. Chem.* **275**, 5409-5415.

Rampazzo, C., Gallinaro, L., Milanesi, E., Frigimelica, E., Reichard, P. & Bianchi, V. (2000b). A Deoxyribonucleotidase in Mitochondria: Involvement in Regulation of dNTP Pools and Possible Link to Genetic Disease. *Proc. Natl. Acad. Sci. USA* **97**, 8239-8244.

Rampazzo, C., Ferraro, P., Pontarin, G., Fabris, S., Reichard, P. & Bianchi, V. (2004). Mitochondrial Deoxyribonucleotides, Pool Sizes, Synthesis, and Regulation. *J. Biol. Chem.* **279**, 17019-17026.

Rao, , S. T. & Rossmann, M. G. (1973). Comparison of Super-secondary Structure in Proteins. *J. Mol. Biol.* **76**, 241-256.

Rechsteiner, M. & Rogers, S. W. (1996). PEST Sequences and Regulation by Proteolysis. *TIBS* **21**, 267-271.

Redfield, C. (1993). Resonance Assignment Strategies for Small Proteins. In "NMR of Macromolecules. A Practical Approach" Edited by Roberts, G. C. K., Oxford University Press.

Reichard, P. (1988). Interactions between Deoxyribonucleotide and DNA Synthesis. *Annu. Rev. Biochem.* **57**, 349-74.

Reichard, P., Eliasson, R., Ingemarson, R. & Thelander, L. (2000). Cross-talk between the Allosteric Effector-binding Sites in Mouse Ribonucleotide Reductase. *J. Biol. Chem.* **275**, 33021-33026.

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**, 7451-7459.

Roberts, G. C. K. (1993). Introduction to NMR of Macromolecules. A Practical Approach. Edited by Roberts, G. C. K. IRL Press.

Roehl, H. H. & Conrad, S. E. (1990). Identification of a G1/S-Phase-Regulated Region in the Human Thymidine Kinase Gene Promoter. *Mol. Cell. Biol.* **10**, 3834-3837.

Roehl, H. H., Anderson, M. M., Mehigh, C. S. & Conrad, S. E. (1993). Regulation of the Cellular Thymidine Kinase Gene Promoter in Simian Virus 40-Infected Cells. *J. Virol.* **67**, 4964-4971.

Romain, S., Spyratos, F., Descotes, F., Daver, A., Rostaing-Puissant, B., Bougnoux, P., Colonna, M., Bolla, M. & Martin, P. M. (2000). Prognostic of DNA-Synthesizing Enzyme Activities (Thymidine Kinase and Thymidylate Synthase) in 908 T1-T2, N0-N1, M0 Breast Cancers: A Retrospective Multicenter Study. *Int. J. Cancer* **87**, 860-868.

Rosenberg, M. J., Agarwala, R., Bouffard, G., Davis, J., Fiermonte, G., Hilliard, M. S., Koch, T., Kalikin, L., Mikalowska, I., Morton, D. H., Petty, E. M., Weber, J. I., Palmieri, F., Kelley, R. I., Schäffer, A. A. & Biesecker, L. G. (2002). Mutant Deoxynucleotide Carrier Is Associated with Congenital Microcephaly. *Nature Genetics* **32**, 175-179.

Rossmann, M. G., Moras, D. & Olsen, K. W. (1974). Chemical and Biological Evolution of a Nucleotide-binding Protein. *Nature* **250**, 194-199.

Rost, B. & Sander, C. (1993). Prediction of Protein Secondary Structure at Better than 70% Accuracy. *J. Mol. Biol.* **232**, 584-599.

Rost, B. & Sander, C. (1994). Combining Evolutionary Information and Neural Networks to Predict Protein Secondary Structure. *Proteins: Struct. Function & Genet.* **19**, 55-72.

Rost, B. & Sander C. (1996). Bridging the Protein Sequence-Structure Gap by Structure Predictions. *Annu. Rev. Biophys. Biomol. Struct.* **25**, 113-136.

Rotheneder, H., Grabner, M. & Wintersberger, E. (1991). Presence of Regulatory Sequences Within Intron 2 of the Mouse Thymidine Kinase Gene. *Nucleic Acids Res.* **19**, 6805-6809.

Rotheneder, H., Geymayer, S. & Haidweger, E. (1999). Transcription Factors of the Sp1 Family: Interaction with E2F and Regulation of the Murine Thymidine Kinase Promoter. *J. Mol. Biol.* **293**, 1005-1015.

Rout, M. P. & Aitchison, J. D. (2001). The Nuclear Pore Complex as a Transport Machine. *J. Biol. Chem.* **276**, 16593-16596.

Rubin, E., Tamrakar, S. & Ludlow, J. W. (1998). Protein Phosphatase Type 1, the Product of the Retinoblastoma Susceptibility Gene, and Cell Cycle Control. *Front. Biosci.* **3**, d1209-1219.

Rudolph., B., Saffrich, R., Zwicker, J., Henglein, B., Müller, R., Ansorge, W. & Eilers, M. (1996). Activation of Cyclin-dependent Kinases by Myc Mediates induction of Cyclin A, but no Apoptosis. *EMBO J.* **15**, 3065-3076.

Saada, A., Shaag, A., Mandel, H., Nevo, Y., Eriksson, S. & Elpeleg, O. (2001). Mutant Mitochondrial Thymidine Kinase in Mitochondrial DNA Depletion Myopathy. *Nature Genet.* **29**, 342-344.

Saada, A., Ben-Shalom, E., Zyslin, R., Miller, C., Mandel, H. & Elpeleg, O. (2003). Mitochondrial Deoxyribonucleoside Triphosphate Pools in Thymidine Kinase 2 Deficiency. *Biochem. Biophys. Res. Commun.* **310**, 963-966.

Sabini, E., Ort, S., Monnerjahn, C., Konrad, M. & Lavie A. (2003). Structure of Human dCK Suggests Strategies to Improve Anticancer and Antiviral Therapy. *Nature Struct. Biol.* **10**, 513-519.

Salviati, L., Sacconi, S., Mancuso, M., Otaegui, D., Camaño, P., Marina, A., Rabonowitz, S., Shiffman, R., Thompson, K., Wilson, C. M., Feigenbaum, A., Naini, A. B., Hirano, M., Bonilla, E., DiMauro, S. & Vu, T. H. (2002). Mitochondrial DNA Depletion and dGK Gene Mutations. *Ann. Neurol.* **52**, 311-316.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual.* 2nd Edn. Cold Spring Harbor Laboratory Press. New York.

Saraste, M., Sibbald, P. R. & Wittinghofer A. (1990). The P-loop – a Common Motif in ATP- and GTP-Binding Proteins. *TIBS* **15**, 430-434.

Sasvari-Szekely, M., Spasokoukotskaja, T., Szoke, M., Csapo, Z., Turi, A., Szanto, I., Eriksson, S. & Staub, M. (1998). Activation of Deoxycytidine Kinase During Inhibition of DNA Synthesis by 2-chloro-2'-deoxyadenosine (Cladribine) in Human Lymphocytes. *Biochem. Pharmacol.* **56**, 1175-1179.

Saurin, A. J., Borden, K. L. B., Boddy, M. N. & Freemont, P. S. (1996). Does This Have a Familiar RING? *Trends Biochim. Sci.* **21**, 208-214.

Schaeffer, H. J. & Weber, M. J. (1999). Mitogen-Activated Protein Kinases: Specific Messages from Ubiquitous Messengers. *Mol. Cell. Biol.* **19**, 2435-2444.

Schlauderer, G. J., Proba, K. & Schulz, G. E. (1996). Structure of a Mutant Adenylate Kinase Ligated with an ATP-analogue Showing Domain Closure Over ATP. *J. Mol. Biol.* **256**, 223-227.

Schägger, H. & von Jagow, G. (1987). Tricine-Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis for the Separation of Proteins in the Range from 1 to 100 kDa. *Anal. Biochem.* **166**, 368-379.

Schleif, R. F. & Wensink, P. C. Practical Methods in Molecular Biology. Springer- Verlag. New York, Heidelberg, Berlin. 1987.

Schlichting, I. & Reinstein, J. (1997). Structures of Active Conformations of UMP Kinase from *Dictyostelium discoideum* Suggest Phosphoryl Transfer Is Associative. *Biochemistry* **36**, 9290-9296.

Schulz, G. E., Müller, C. W. & Diederichs K. (1990). Induced–Fit Movements in Adenylate Kinases. *J. Mol. Biol.* **213**, 627-630.

Schulz, G. E., Schiltz, E., Tomasselli, A. G., Frank, R., Brune, M., Wittinghofer, A. & Schirmer, R. H. (1986). Structural Relationships in the Adenylate Kinase Family. *Eur. J. Biochem.* **161**, 127-132..

Sekulic, N., Shuvalova, L., Spangenberg, O. & Konrad, M. (2002). Structural Characterization of the Closed Conformation of Mouse Guanylate Kinase. *J. Biol. Chem.* **277**, 30236-30243.

Serber, Z. & Dötsch, V. (2001). In-Cell NMR Spectroscopy. *Biochemistry* **40**, 14317-14323.

Shaulian, E. & Karin, M. (2001). Ap-1 in Cell Proliferation and Survival. *Oncogene* **20**, 2390-2400.

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**, 375-382.

Sherley, J. L. & Kelly, T. J. (1988b). Regulation of Human Thymidine Kinase During the Cell Cycle. *J. Biol. Chem.* **263**, 8350-8358.

Sherr, C. J. (1994). G1 Phase Progression: Cycling on Cue. *Cell* **79**, 551-555.

Sherr, C. J. (1996). Cancer Cell Cycles. *Science* **274**, 1672-1677.

Sherr, C. & Roberts, J. (1999). CDK Inhibitors: Positive and Negative Regulators of G1 –Phase Progression. *Genes Dev.* **13**, 1501-1512.

Shewach, D. S., Reynolds, K. K. & Hertel, L. (1992). Nucleotide Specificity of Human Deoxycytidine Kinase. *Mol. Pharmacol.* **42**, 518-524.

Sjöberg, A. H., Wang, L. & Eriksson, S. (1998). Substrate Specificity of Human Recombinant Mitochondrial Deoxyguanosine Kinase with Cytostatic and Antiviral Purine and Pyrimidine Analogs. *Mol. Pharmacol.* **53**, 270-273.

Sjöberg, A. H., Wang, L. & Eriksson, S. (2001). Antiviral Guanosine Analogs as Substrates for Deoxyguanosine Kinase: Implications for Chemotherapy. *Antimicrob. Agents Chemother.* **45**, 739-742.

Skoog, K. L., Nordenskjöld, B. A. & Bjursell, K. G. (1973). Deoxyribonucleoside-Triphosphate Pools and DNA Synthesis in Synchronized Hamster Cells. *Eur. J. Biochem.* **33**, 428-432.

Skoog, K. L., Bjursell, K. G. & Nordenskjöld, B. A. (1974). Cellular Deoxyribonucleotide Triphosphate Pools Levels and DNA Synthesis. *Adv. Enzyme Regul.* **12**, 345-354.

Slagel, V., Flemington, E., Traina-Dorge, V., Bradshaw, H. & Deininger, P. (1987). Clustering and Subfamily Relationships of the Alu Family in the Human Genome. *Mol. Biol. Evol.* **4**, 19-29.

Smith, D. B. & Johnson, K. S. (1988). Single-Step Purification of Polypeptides Expressed in *Escherichia coli* as Fusions with Glutathione S-Transferase. *Gene* **67**, 31-40.

Solomon, E., Ledbetter, D. & Fain, P. (1993). Report of the Committee on the Genetic Constitution of Chromosome 17. *Genome Prior. Rep.* **1**, 473-504.

Song, J. J., Walker, S., Chen, E., Johnson, E. E. II, Spychala, J., Gribbin, T. & Mitchell, B. S. (1993). Genomic Structure and Chromosomal Localization of the Human Deoxycytidine Kinase Gene. *Proc. Natl. Acad. Sci. USA* **90**, 431-434.

Song, S., Wheeler, L. J. & Mathews, C. K. (2003). Deoxyribonucleotide Pool Imbalance Stimulates Deletions in HeLa Cell Mitochondrial DNA. *J. Biol. Chem.* **278**, 43893-43896.

Sorensen, P. & Wintersberger, E. (1999). Sp1 and NF-Y Are Necessary and Sufficient for Growth-dependent Regulation of the Hamster Thymidine Kinase Promoter. *J. Biol. Chem.* **274**, 30943-30949.

Spasokoukotskaja, T., Taljanidisz, J., Sasvári-Székely, M. & Staub, M. (1991). Deoxycytidine is Salvaged not Only into DNA but also to Phospholipid Precursors. III. dCDP-diacylglycerol Formation in Tonsillar Lymphocytes. *Biochem. Biophys. Res. Commun.* **174**, 680-687.

Spasokoukotskaja, T., Amér, E. S. J., Brosjö, O., Gunvén, P., Juliusson, G., Liliemark, J. & Eriksson, J. (1995). Expression of Deoxycytidine Kinase and Phosphorylation of 2-Chlorodeoxyadenosine in Human Normal and Tumour Cells and Tissues. *Eur. J. Cancer* **31A**, 202-208.

Spasokoukotskaja, T., Sasvári-Székely, M., Keszler, G., Albertioni, F., Eriksson, S. & Staub, M. (1999). Treatment of Normal and Malignant Cells with Nucleoside Analogues and Etoposide Enhances Deoxycytidine Kinase Activity. *Eur. J. Cancer* **35**, 1862-1867.

Spinazzola, A., Marti, R., Nishino, I., Andreu, A. L., Naini, A., Tadesse, S., Pela, I., Zammarchi, E., Donati, M. A., Oliver, J. A. & Hirano, M. (2002). Altered Thymidine Metabolism Due to Defects of Thymidine Phosphorylase. *J. Biol. Chem.* **277**, 4128-4133.

Spyrou, G. & Reichard, P. (1987). Compartmentation of dCTP Pools. Evidence from Deoxyliponucleotide Synthesis. *J. Biol. Chem.* **262**, 16425-16432.

Stegmann, A. P. A., Honders, M. W., Bolk, M. W. J., Wessels, J., Willemze, R. & Landegent, J. E. (1993). Assignment of the Human Deoxycytidine Kinase (DCK) Gene to Chromosome 4 Band q13.3-q21.1. *Genomics* **17**, 528-529.

Sternberg, M. J. E., Bates, P. A., Kelley, L. A. & MacCallum, R. M. (1999). Progress in Protein Structure Prediction: Assessment of CASP3. *Current Opin. Struct. Biol.* **9**, 368-373.

Stewart, C. J., Ito, M. & Conrad, S. E. (1987). Evidence for Transcriptional and Post-Transcriptional Control of the Cellular Thymidine Kinase Gene. *Mol. Cell. Biol.* **7**, 1156-1163.

Stewart, L., Ireton, G. C., Parker, L. H., Madden, K. R. & Champoux, J. J. (1996). Biochemical and Biophysical Analyses of Recombinant Forms of Human Topoisomerase I. *J. Biol. Chem.* **271**, 7593-7601.

Stuart, P., Ito, M., Stewart, C. & Conrad, S. E. (1985). Induction of Cellular Thymidine Kinase Occurs at the mRNA Level. *Mol. Cell. Biol.* **5**, 1490-1497.

Sulpizi, M., Schelling, P., Folkers, G., Carloni, P. & Scapozza, L. (2001). The Rational of Catalytic Activity of Herpes Simplex Virus Thymidine Kinase. A Combined Biochemical and Quantum Chemical Study. *J. Biol. Chem.* **276**, 21692-21697.

Sutterluety, H., Barti, S., Karlseder, J., Wintersberger, E. & Seiser, C. (1996). Carboxy-Terminal Residues of Mouse Thymidine Kinase are Essential for Rapid Degradation in Quiescent Cells. *J. Mol. Biol.* **259**, 383-392.

Sutterluety, H. & Seiser, C. (1997). Thymidine Inhibits the Growth-arrest-specific Degradation of Thymidine Kinase Protein in Transfected L Fibroblasts. *J. Mol. Biol.* **265**, 153-160.

Sutterluety, H., Barti, S., Doetzlhofer, A., Khier, H., Wintersberger, E. & Seiser, C. (1998). Growth-Regulated Antisense Transcription of the Mouse Thymidine Kinase Gene. *Nucleic Acids Res.* **26**, 4989-4995.

Swank, R. T. & Munkres, K. D. (1971). Molecular Weight Analysis of Oligopeptides by Electrophoresis in Polyacrylamide Gel with Sodium Dodecyl Sulfate. *Analyt. Biochem.* **39**, 462-477.

Taanman, J. W., Kateeb, I., Muntau, A. C., Jaksch, M., Cohen, N, & Mandel, H. (2002). A Novel Mutation in the Deoxyguanosine Kinase Gene Causing Depletion of Mitochondrial DNA. *Ann. Neurol.* **52**, 237-239.

Taanman, J. W., Muddle, J. R. & Muntau, A. C. (2003). Mitochondrial DNA Depletion Can Be Prevented by dGMP and dAMP Supplementation in a Resting Culture of Deoxyguanosine Kinase-Deficient Fibroblasts. *Hum. Mol. Genet.* **12**, 1839-1845.

Tamrakar, S., Rubin, E. & Ludlow, J. W. (2000). Role of pRB Dephosphorylation in Cell Cycle Regulation. *Front. Biosci.* **5**, d121-137.

Tanaka, H., Arakawa, H., Yamaguchi, T., Shiraishi, K., Fukuda, S., Matsui, K., Takei, Y. & Nakamura, Y. (2000). A Ribonucleotide Reductase Gene Involved in a p53-dependent Cell-cycle Checkpoint for DNA Damage. *Nature* **404**, 42-49.

Tetsu, O. & McCormick, F. (2003). Proliferation of Cancer Cells Despite CDK2 Inhibition. *Cancer Cell* **3**, 233-245.

Thelander, L. & Reichard, P. (1979). Reduction of Ribonucleotides. *Annu. Rev. Biochem.* **48**, 133-58.

Tommasi, S. & Pfeifer, G. P. (1997). Constitutive Protection of E2F Recognition Sequences in the Human Thymidine Kinase Promoter during Cell Cycle Progression. *J. Biol. Chem.* **272**, 30483-30490.

Traut, T. W. (1994a). Physiological Concentrations of Purines and Pyrimidines. *Mol. Cell. Biochem.* **140**, 1-22.

Traut, T. W. (1994b). The Functions and Consensus Motifs of Nine Types of Peptide Segments that Form Different Types of Nucleotide-binding Sites. *Eur. J. Biochem.* **222**, 9-19.

Travali, S., Lipson, K. E., Jaskulski, D., Lauret, E. & Baserga, R. (1988). Role of the Promoter in the Regulation of the Thymidine Kinase Gene. *Mol. Cell. Biol.* **8**, 1551-1557.

Tsai, C. J., Lin, S. L., Wolfson, H. J. & Nussinov, R. (1996). Protein-Protein Interfaces: Architectures and Interactions in Protein-Protein Interfaces and in Protein Cores. Their Similarities and Differences. *Crit. Rev. Biochem. Mol. Biol.* **31**, 127-152.

Tsvetkov, L. M., Yeh, K. H., Lee, S. J., Sun, H. & Zhang, H. (1999). p27^{Kip1} Ubiquitination and Degradation Is Regulated by the SCF^{Skp2} Complex Through Phosphorylated Thr187 in p27. *Curr. Biol.* **9**, 661-654.

Turk, B., Awad, R., Usova, E. V., Björk, I. & Eriksson, S. (1999). A Pre-Steady-State Kinetics Analysis of Substrate Binding to Human Recombinant Deoxycytidine Kinase: A Model for Nucleoside Kinase Action. *Biochemistry* **38**, 8555-8561.

Tyers, M. & Jorgensen, P. (2000). Proteolysis and the Cell Cycle: With this RING I Do Thee Destroy. *Curr. Opinion Genet. Dev.* **10**, 54-64.

Tyrsted, G. (1975). The Pool Size of Deoxyguanosine 5'-Triphosphate and Deoxycytidine 5'-Triphosphate in Phytohemagglutinin-stimulated and Non-stimulated Human Lymphocytes. *Exptl. Cell Res.* **91**, 429-440.

Tyrsted, G. & Munch-Petersen, B. (1977). Early Effects of Phytohemagglutinin on Induction of DNA Polymerase, Thymidine Kinase, Deoxyribonucleoside Triphosphate Pools and DNA Synthesis in Human Lymphocytes. *Nucleic Acids Res.* **4**, 2713-2723.

Vazques, F., Ramaswamy, S., Nakamura, N. & Sellers, W. (2000). Phosphorylation of the PTEN Tail Regulates Protein Stability and Function. *Mol. Cell. Biol.* **20**, 5010-5018.

Vojtek, A. B. & Der, C. J. (1998). Increasing Complexity of the Ras Signaling Pathway. *J. Biol. Chem.* **273**, 19925-19928.

Vogt, J., Perozzo, R., Pautsch, A., Prota, A., Schelling, P., Pilger, B., Folkers, G., Scapozza, L. & Schulz, G. E. (2000). Nucleoside Binding Site of Herpes Simplex Type 1 Thymidine Kinase Analysed by X-Ray Crystallography. *Proteins: Struct. Funct. Genet.* **41**, 545-553.

Vonrhein, C., Schlauderer, G. J. & Schulz, G. E. (1995). Movie of the Structural Changes During a Catalytic Cycle of Nucleoside Monophosphate Kinases. *Structure* **3**, 483-490.

Vucetic, S., Brown, C. J., Dunker, A. K. & Obradovic, Z. (2003). Flavors of Protein Disorder. *Proteins: Struct. Function Genet.* **52**, 573-584.

Wakazono, Y., Kubota, M., Furusho, K., Liu, L. & Gerson, S. L. (1996). Thymidine Kinase Deficient Cells with Decreased TTP Pools Are Hypersensitive to DNA Alkylating Agents. *Mutation Res.* **362**, 119-125.

Walker, J. E., Saraste, M., Runswick, M. J. & Gay, N. J. (1982). Distantly Related Sequences in the α- and β-subunits of ATP Synthase, Myosin, Kinases and Other ATP-requiring Enzymes and a Common Nucleotide Binding Fold. *EMBO J.* **1**, 945-51.

Wang, J. & Eriksson, S. (1996). Phosphorylation of the Anti-Hepatitis B Nucleoside Analog 1-(2'-Deoxy-2'-Fluoro-1-β-D-Arabinofuranosyl)-5-Iodouracil (FIAU) by Human Cytosolic and Mitochondrial Thymidin Kinase and Implications for Cytotoxicity. *Antimicrob. Agents Chemother.* **40**, 1555-1557.

Wang, J., Choudhury, D., Chattopadhyaya J. & Eriksson, S. (1999b). Stereoisomeric Selectivity of Human Deoxyribonucleoside Kinases. *Biochemistry* **38**, 16993-16999.

Wang, L., Karlsson, A. Arnér, E. S. J. & Eriksson, S. (1993). Substrate Specificity of Mitochondrial 2'-Deoxyguanosine Kinase. Efficient Phosphorylation of 2-Chlorodeoxyadenosine. *J. Biol. Chem.* **268**, 22847-22852.

Wang, L., Hellman, U. & Eriksson, S. (1996). Cloning and Expression of Human Mitochondrial Deoxyguanosine Kinase cDNA. *FEBS Lett.* **390**, 39-43.

Wang, L., Munch-Petersen, B., Sjöberg, A. H., Hellman, U., Bergman, T., Jörnvall, H. & Eriksson, S. (1999a). Human Thymidine Kinase 2: Molecular Cloning and Characterization of the Enzyme Activity with Antiviral and Cytostatic Nucleoside Substrates. *FEBS Lett.* **443**, 170-174.

Wang, L. & Eriksson, S. (2000). Cloning and Characterization of Full-Length Mouse Thymidine Kinase 2: the N-terminal Sequence Directs Import of the Precursor Protein into Mitochondria. *Biochem. J.* **351**, 469-476.

Wang, L. & Eriksson, S. (2003). Mitochondrial Deoxyguanosine Kinase Mutations and Mitochondrial DNA Depletion Syndrome. *FEBS Lett.* **554**, 319-322.

Wang, L., Saada, A, & Eriksson, S. (2003). Kinetic Properties of Mutant Human Thymidine Kinase 2 Suggest a Mechanism for Mitochondrial DNA Depletion Myopathy. *J. Biol. Chem.* **278**, 6963-6968.

Wawra, E. (1988). Microinjection of Deoxynucleotides into Mouse Cells. No Evidence that Precursors for DNA Synthesis Are Channelled. *J. Biol. Chem.* **263**, 9908-9912.

Weinberg, G., Ullman, B. & Martin, D. W. Jr. (1981). Mutator Phenotypes in Mammalian Cell Mutants with Distinct Biochemical Defects and Abnormal Deoxyribonucleoside Triphosphate Pools. *Proc. Natl. Acad. Sci. USA* **78**, 2447-2451.

Wente, S. R. (2000). Gatekeepers of the Nucleus. *Science* **288**, 1374-1377.

Westbrook, J., Feng, Z., Jain, S., Bhat, T. N., Thanki, N., Ravichandran, V., Gilliland, G. L., Bluhm, W. F., Weissig, H., Greer, D. S., Bourne, P. E. & Berman, H. M. (2002). The Protein Data Bank: Unifying the Archive. *Nucleic Acids Res.* **30**, 245-248.

White, J. C. & Capizzi, R. L. (1991). A Critical Role for Uridine Nucleotides in the Regulation of Deoxycytidine Kinase and the Concentration Dependence of 1-β-D-Arabinofuranosylcytosine Phosphorylation in Human Leukemia Cells. (1991). *Cancer Res.* **51**, 2559-2565.

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**, 289-292.

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. *Proteins Sci.* **6**, 2097-2106.

Wilkinson, Y. A. & McKenna, P. G. (1989). The effects of Thymidine on Deoxyribonucleotide Pool Levels, Cytotoxicity and Mutation Induction in Friend Mouse Erythroleukaemia Cells. *Leukemia Res.* **13**, 615-620.

Willecke, K., Reber, T., Kucherlapati, R. S. & Ruddle, F. H. (1977). Human Mitochondrial Thymidine Kinase is Coded for by a Gene on Chromosome 16 of the Nucleus. *Somat. Cell Genet.* **3**, 237-245.

Williamson, M. P. (1994). The Structure and Function of Proline-Rich Regions in Proteins. *Biochem. J.* **297**, 249-260.

Wintersberger, E. (1997). Regulation and Biological Function of Thymidine Kinase. *Biochem Soc. Transactions* **25**, 303-308.

Wishart, D. S., Sykes, B. D. & Richards, F. M. (1991). Relationship between Nuclear Magnetic Resonance Chemical Shift and Protein Secondary Structure. *J. Mol. Biol.* **222**, 311-333.

Wlodawer, A., Miller M., Jaskolski, M., Sathyanarayana, B. K., Baldwin, E., Weber, I. T., Selk, L. M., Clawson, L., Schneider, J. & Kent, S. B. (1989). Conserved Folding in Retroviral Proteases: Crystal Structure of a Synthetic HIV-1 Protease. *Science* **245**, 616-621.

Wright, P. E. (1989). What Can Two-Dimensional NMR Tell Us About Proteins? *TIBS* **14**, 255-260.

Wright, P. E. & Dyson, H. J. (1999). Intrinsically Unstructured Proteins: Re-assessing the Protein Structure-Function Paradigm. *J. Mol. Biol.* **293**, 321-331.

Wurth, C., Kessler, U., Vogt, J., Schulz, G. E., Folkers, G. & Scapozza, L. (2001). The Effect of Substrate Binding on the Conformation and Structural Stability of Herpes simplex Virus Type 1 Thymidine Kinase. *Protein Sci.* **10**, 63-73.

Wüthrich, K., Wider, G., Wagner, G. & Braun, W. (1982). Sequential Resonance Assignments as a Basis for Determination of Spatial Protein Structures by High Resolution Proton Nuclear Magnetic Resonance. *J. Mol. Biol.* **155**, 311-319.

Wüthrich, K., Billeter, M. & Braun, W. (1984). Polypeptide Secondary Structure Determination by Nuclear Magnetic Resonance Observation of Short Proton-Proton Distances. *J. Mol. Biol.* **180**, 715-740.

Wüthrich, K. (1986). NMR of Proteins and Nucleic Acids. John Wiley & Sons, Inc.

Wüthrich, K. (1995). NMR – This Other Method for Protein and Nucleic Acid Structure Determination. *Acta Cryst.* **D51**, 249-270.

Xu, D., Lin, S. L. & Nussinov, R. (1997). Protein Binding *versus* Protein Folding: The Role of Hydrophobic Bridges in Protein Associations. *J. Mol. Biol.* **265**, 68-84.

Yam, C. H., Fung, T. K. & Poon, R. Y. C. (2002). Cyclin A in Cell Cycle Control and Cancer. *CMLS Cell. Mol. Life Sci.* **59**, 1317-1326.

Yan, H. & Tsai, M. D. (1991). Mechanism of Adenylate Kinase. Demonstration of a Functional Relationship Between Aspartate 93 and Mg^{2+} by Site-directed Mutagenesis and Proton, Phosphorus-31, and Magnesium-25 NMR. *Biochemistry* **30**, 5539-5546.

Yan, H. & Tsai, M. D. (1999). Nucleoside Monophosphate Kinases: Structure, Mechanism, and Substrate Specificity. *Advances Enzymol. Rel. Areas Mol. Biol.* **73A**, 103-134.

Yoshioka, A., Tanaka, S., Hiraoka, O., Koyama, Y., Hirota, Y. & Wataya Y. (1987). Deoxyribonucleoside-triphosphate Imbalance Death: Deoxyadenosine-induced dNTP Imbalance and DNA Double Strand Breaks in Mouse FM3A Cells and the Mechanism of Cell Death. *Biochem. Biophys. Res. Commun.* **146**, 258-264.

Young, P., Leeds, J. M., Slabaugh, M. B. & Mathews, C. K. (1994a). Ribonucleotide Reductase: Evidence for Specific Association with HeLa Cell Mitochondria. *Biochem. Biophys. Res. Commun.* **203**, 46-52.

Young, L., Jernigan, R. L. & Covell, D. G. (1994b). A role for Surface Hydrophobicity in Protein-Protein Recognition. *Protein Sci.* **3**, 717-729.

Zachariae, W. (1999). Progression Into and Out of Mitosis. *Curr. Opin. Cell Biol.* **11**, 708-716.

Zaharevitz, D. W., Anderson, L. W., Malinowski, N. M., Hyman, R., Strong, J. M. & Cysyk, R. L. (1992). Contribution of *De novo* and Salvage Synthesis to the Uracil Nucleotide Pool in Mouse Tissues and Tumors *in vivo*. *Eur. J. Biochem.* **210**, 293-296.

Zarkowska, T. & Mittnacht, S. (1997). Differential Phosphorylation of the Retinoblastoma Protein by G1/S Cyclin-Dependent Kinases. *J. Biol. Chem.* **272**, 12738-12746.

Zhang, C. T. (1997). Relations of the Numbers of Protein Sequences, Families and Folds. *Protein Eng.* **10**, 757-761.

Zhang, H., Hannon, G. J. & Beach, D. (1994). P21-Containing Cyclin Kinases Exist in Both Active and Inactive States. *Genes & Dev.* **8**, 1750-1758.

Zhang, M., Pickart, C. M. & Coffino, P. (2003). Determinants of Proteasome Recognition of Ornithine Decarboxylase, a Ubiquitine-independent Substrate. *EMBO J.* **22**, 1488-1496.

Zhang, Z., Wang, H., Li, M., Agrawal,. S., Chen, X. & Zhang, R. (2004). MDM2 Is a Negative Regulator of p21$^{WAF/CIP1}$, Independent of p53. *J. Biol. Chem.* **279**, 16000-16006.

Zhu, C., Johansson, M., Permert, J. & Karlsson, A. (1998). Phosphorylation of Anticancer Nucleoside Analogs by Human Mitochondrial Deoxyguanosine Kinase. *Biochem. Pharmacol.* **56**, 1035-1040.

Zhu, C., Johansson, J. & Karlsson, A. (2000). Incorporation of Nucleoside Analogs into Nuclear or Mitochondrial DNA is Determined by the Intracellular Phosphorylation Site. *J. Biol. Chem.* **275**, 26727-26731.

Zuiderweg, E. R. P. (2002). Mapping Protein-Protein Interactins in Solution by NMR Spectroscopy. *Biochemistry* **41**, 1-7.

Zuo, Y. & Deutscher, M. P. (2002). Mechanism of Action of Rnase T. *J. Biol. Chem.* **277**, 50160-50164.

Zutshi, R., Brickner, M. & Chmielewski, J. (1998). Inhibiting the Assembly of Protein-Protein Interfaces. *Curr. Opin. Chem. Biol.* **2**, 62-66.

Zvelebil, M. J., Barton, G. J., Taylor, W. R. & Sternberg, M. J. E. (1987). Prediction of Protein Secondary Structure and Active Sites using the Alignment of Homologous Sequences. *J. Mol. Biol.* **195**, 957-961.