Universa Universi Patavina Libertas

Universitas Studii Paduani

Faculty of Mathematical, Physical and Natural Sciences

Socrates/Erasmus Exchange Project with

ROSKILDE UNIVERSITETSCENTER

Site directed mutagenesis and enzyme kinetics study of Thymidine Kinase 1 from *Caenorhabditis elegans*

Supervisors: **Prof. Birgitte Munch-Petersen**, Dept. of Chemistry and Life Sciences, RUC
Prof. Vera Bianchi, Dept. of Biology, University of Padua

Tutors: **Tine Skovgaard** and **Dvora Berenstein**

Student: **Andrea Telatin**

Academic Year 2005/2006

Hunc igitur terrorem animi tenebrasque necessest
non radii solis neque lucida tela diei
discutiant, sed naturae species ratioque.
So this terror of the soul and this darkness
are to be dispelled not by sun rays or by bright arrows
of the day, but by rational observation of Nature.

Lucretius, DE RERUM NATURA, Liber I

Gutta cavat lapidem, non vi sed saepe cadendo
The drop excavates the stone, not with force
but continuously falling.

Ovidius, EPISTULAE EX PONTO

Table of Contents

1. Introduction

1.1. Abstract

Thymidine Kinase 1, that is involved in nucleotide metabolism and therefore target for antiviral and antineoplastic drugs, has been object of intense studies. This project consisted in mutagenesis, purification and kinetic analysis of the recently cloned TK1 from the worm *Caenorhabditis elegans*.

The 42:Met→Ile was chosen to check the role of the methionine proximal to the active site in catalysis and substrate specificity, by substituting the highly conserved methionine with a branched aliphatic amino acid.

The vector (pGEX-2T) containing the ORF of *Ce*TK1 is designed to perform purification with the GSH-system. Mutagenesis was performed with the QuickChange™ protocol followed by transformation into *E. coli* BL21, purification and quantification with Bradford.

The mutant shows a decreased activity compared with wild type, with both the natural substrate (Thymidine) and the analogue AZT. Interestingly its activity with deoxycytidine is higher than wild type, and with a very low K_M.

The present report describes effect of M42I substitution in the worm TK1, comparing results with homologue substitution in Human orthologue.

1.2. Notes

Typographic conventions - Links to paragraphs are marked with a "§" symbol; images have the same number of the paragraph they are in; when needed the code is followed by an extra serial number (e.g.: fig. 3.2.2 is the second picture of the paragraph 3.2). Notes are indicated with roman lowercase number (i.e. "[iv]"), and are at the end of the report, while bibliographic references follows the same format, but with arabic numbers (like i.e. "[4]"). **Protein images** were made or with the PDB Viewer 3.7 from GlaxoSmithKline or with RasMol 2.7.3. Rendering was performed with POV-ray 3.6 and **data analysis** with Microsoft® Excel™, non linear curve fitting with Origin™. Custom scripts were written in perl (CPAN.org), and executed in an Intel/Win32 environment.

Nomenclature – Not to make the report dull reading, biochemical names will be used as follows: the word "nucleotide" is considered a generic name for both ribonucleotides and deoxyribonucleotides, and thus with "nucletide metabolism" we refer to metabolic pathways for their synthesis. Ribonucleotides and deoxyribonucleotides are instead used to specifically refer to one or the other class of nucleic acids precursors. Commonly used abbreviations for nucleotides are used as in [1] (i.e.: dCMP: deoxycytidine monophosphate...). Both thymidine and deoxythymidine are used as synonyms, as the ribonucleotide for this base doesn't exist.

Abbreviations: *Ce*: *Caenorhabditis elegans*, Hu: Human; TK1: Thymidine Kinase 1; mtDNA: mitochondrial DNA; ORF: open reading frame; GST: Glutatione-S-Transferase; GSH: Reduced Glutathione; AZT: 3'-Azidothymidine; ddH_2O: sterile and deionized water; AraC: arabinosilcytidine;

IMPORTANT: Due to use of Italian version of Excel, some plots has a comma as decimal separator. They are marked with [CSD], comma separated decimals. This issue shouldn't confuse the reader, as it is present only with number with two decimals.

1.3. Web Repository

Supplemental data (such as tables and extra or raw data) are provided for download at the Internet address: http://www.neofox.it/bio/tk1/

2. Theoretical foundations

2.1. Kinetic analysis as a clue on catalytic mechanism

Kinetics analysis was the first approach used to understand biological catalysis, and led to the fundamental idea of enzyme-substrate complex, inferred by Henry from the saturation curve obtained from "invertase" enzyme[4]. Even if a long time has passed since those first efforts, and structural studies solved several questions related to reaction mechanisms, kinetic analysis remains irreplaceable, as it – indirectly – studies the "live" features of a working enzyme, while X-ray diffraction shows – directly – structural features of a "chalky" crystallized enzyme. The aim of this project is to have a clue on the role of amino acid Met[42], and kinetic investigation should provide this.

> *"We are now in a better position to solve the old problem of the reaction equation of invertase in a rational manner and without help of more than one arbitrary constant.*
> **Michaelis and Menten, 1913[23]**

Michaelis-Menten analysis

Several complicating issues (i. e. feed-back inhibition, reverse reaction) made it difficult first studies. A powerful simplifying approach was proposed by Leonor Michaelis and Maud Menten (1913). Their fundamental innovation was to measure the **initial rate** (v_0) of the reaction, when the substrate concentration ([S]) is much greater than the one of the enzyme ([E]), so that can be considered constant. Even with their refinements[i] their results agreed with the previously proposed by Henry mechanism[4]: $E + S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E + P$.

Measuring the initial rate will cause the reverse reaction to be negligible, as [P]=0 in the beginning. Initial rate can be expressed as $v_0 = k_2[ES]$, resulting in the broadly used Michaelis-Menten equation[1]:

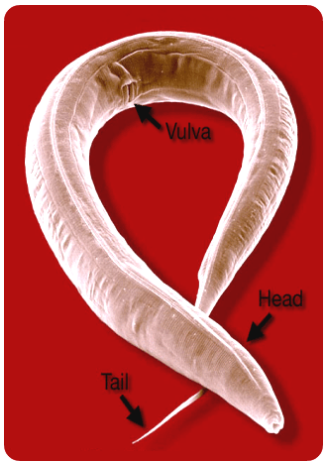
$$V = \frac{V_{max} \cdot [S]}{K_M + [S]}$$

Where V_{max} represents the asymptote of the curve, thus the maximum velocity reached in saturating conditions. The K_M is a function of different rate constants (just three in the case of the simplified mechanism shown above ($[k_{-1}+k_2]/k_1$), and is used as an indicator of the affinity of the enzyme for the substrate; it represents the concentration at which the reaction rate is one half of V_{max}. For comparison with different enzymes (or experiments) one should consider that $V_{max}=k_{cat}[E]$ [ii], in the above expression. This allows defining the *turnover number* (k_{cat}), as the maximum number of molecule of substrate that can be converted per second per active site.

Under physiological conditions, enzyme usually don't work in saturating conditions, and [S] is usually smaller than K_M. The *specificity constant* (k_{cat}/K_M) can be introduced, and it's useful to compare the relative activity with different compounds for the same enzyme, but also to compare differen enzymes[iii].

2.2. Model organism

C. elegans as a model organism



The studied enzyme was cloned from *Caenorhabditis elegans*. Unlike the other invertebrate model, *Drosophila melanogaster*, that can boast a long history in genetics and molecular biology, the flat worm has been used as a model organism only since 1963, when Sydney Brenner chose it for his studies of the role of genes in development and the nervous system.

This nematode is only 1 mm long containing less than 1000 somatic cells, and can be handled in a "microbiological fashion" within Petri dishes, feeding on bacteria, and nonetheless it's transparent throughout its life. Therefore, the behaviour of individual cells can be followed through development and gene expression patterns can be monitored very easily. A unique feature of this animal is that development is stereotypical, that means that cells divide and specialize in a totally characteristic way so that every normal worm has the same number and type of cells arranged in the same manner[3].

The *C. elegans* genome is 97 million base pairs in length and contains about 18.000 genes. *Many of these genes appear to have functional counterparts in humans, and whole pathways are often conserved*. The whole genome sequence was finished in 1998, being the first animal genome obtained[13].

Forty years later, and the wisdom of this choice has been consecrated not only by the impact of the worm on a broad range of research fields, but also by the award of the 2002 Nobel Prize for Physiology or Medicine to three worm biologists: Sydney Brenner, Sir John Sulston and Bob Horvitz.


Sydney Brenner


Robert Horvitz


John Sulston

The nematode is still of enormous interest because of the availability of molecular tools as RNA interference, that can be achieved simply by feeding the worm on *E. coli* transformed to produce antisense RNA of target gene[14].

Recently a genome-wide screening of *C. elegans* transcripts showed an abundant use of "operons"[12], polycistronic mRNAs, that differs from the common prokaryotic counterpart because they require a "trans-splicing" mechanism, and they also seems merge unrelated genes. TK1 gene from *C. elegans* is part of a bicistronic operon (see below).

Webrepository (§1.3) contains extra data on this topic.

2.3. Nucleotide metabolism in eukaryotic cells

Overview

Deoxynucleotides are the cellular precursor for both nuclear and mitochondrial DNA[9,10], thus they are required during S *phase* of the cell cycle for nuclear genome replication, but, even if in smaller amount, also during the whole life of the cell, for DNA repair and mtDNA replication. Nucleotides in general carry out several important cellular function, both as polymer units (for DNA and RNA) and as molecules themselves (energy carriers, intra- and extracellular messengers)[1].

There are two complementary biosynthetic pathways for these precursors: the ***de novo* pathway**, that build up nucleotides from simpler precursors and is active in proliferating cells, and the **salvage pathway**, that rescue nucleobase and free bases, particularly important in differentiated cells. Deoxyribonucleotides (see below) are derived from ribonucleotides by reduction of 2'-OH group.

Nucleosides for salvage pathway come from degradation of endogenous nucleic acids (mainly ribonucleotides), and – in animals – from food (mainly deoxyribonucleotides).

Biosynthesis regulation is active at different levels: genetic expression, translation and directly on enzymes. This regulation is important not only to ensure an appropriate amount of precursor, but also to ensure a balance between the four nucleotides; a misbalance can indeed cause misincorporation of wrong base during DNA synthesis[9].

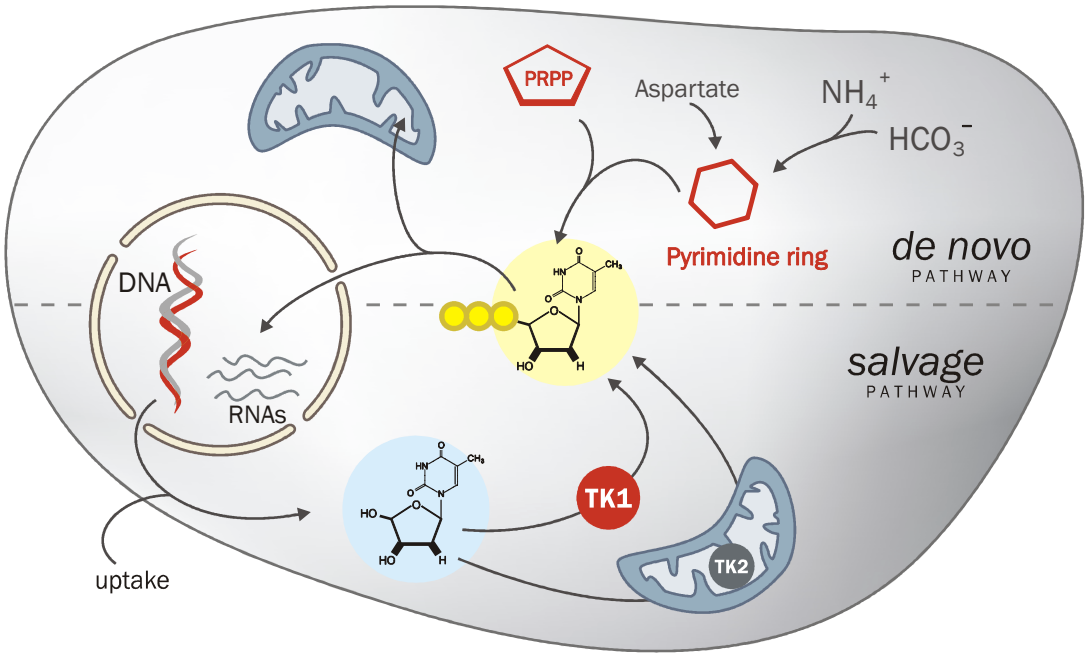
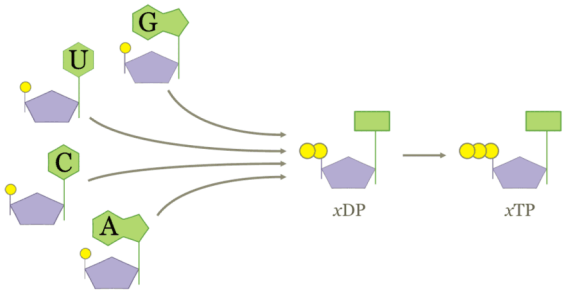


Fig. 2.3.1 – Schematic overview of *de novo* and salvage pathways for Thymidine synthesis. The salvage pathways is intended to rescue free bases or nucleosides originated from degradation of polynucleotides (mainly mRNAs) or cellular uptake of digestion products (mainly DNA).

De novo synthesis of pyrimidines

Synthesis of pyrimidine ring – Anabolic pathway for the synthesis of pyrimidine ribonucleotides[11] starts with the construction of the pyrimidine ring, as orotate, that is subsequently conjugated to the activated sugar phosphoribosylpyrophosphate (PRPP), while the purine ring is directly synthesized on the activated sugar, producing the inosin monophosphate nucleotide. Precursors for orotate are HCO_3^- ion and NH_3 taken from the amino acid glutamine, in a long and chemically expensive (thus tightly regulated) pathway.

Synthesis of ribonucleotides – The first nucleotide formed is uridinmonophosphate (UMP), by the action of the bifunctional enzyme *UMP synthetase* combining the phosphorybosyl transferase and decarboxylase activities. Ribonucleotides has to be phosphorylated twice to become fully functional precursors, by the action of a *nucleoside monophosphate kinase* first, and a *nucleoside diphosphate kinase* then. In higher eukaryotes the former hasn't preference for the sugar, but only for the nitrogenbase (thus four are present to phosphorylate all occurring ribonucleotides), while the latter activity is present in one non-specific enzyme. UTP can be converted to CTP by the *cytidine triphosphate synthetase*.

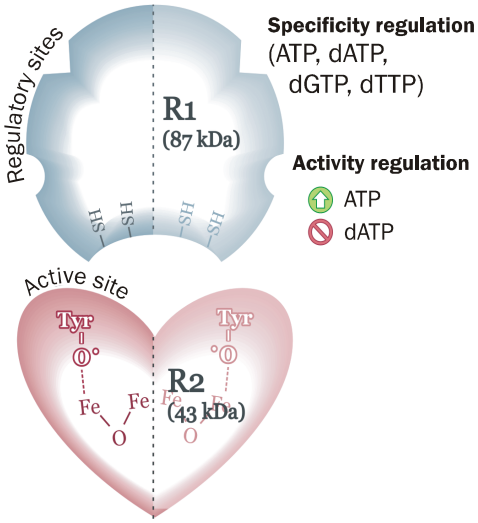


Deoxynucleotide biosynthesis

DNA precursors (deoxyribonucleoside triphosphate) are derived from ribonucleotides[17,18], with the reduction of 2'–OH group catalyzed by **Ribonucleotide reductase (RNR)**, an interesting enzyme[18], composed by two homodimeric subunits, acting with radicalic mechanism. Both activity and substrate specificity are allosterically regulated, as depicted in the scheme. The conversion involves the nucleosides diphosphate, thus yielding deoxyribonucleosides diphosphate.

Many evidences were accumulated in the past years supposing that RNA was the first biological molecule, with both the role of information storage, now acquired by DNA, and catalysis, now mainly performed by proteins. The fact that the *de novo* biosynthetic pathway of deoxyribonucleotides begins with the reduction of ribonucleotides, agrees with this broadly accepted hypothesis.

From a chemical point of view, the lack of the –OH group makes it more difficult to hydrolyze DNA rather RNA, resulting in a more stable molecule



(as the genetic material should be). Structurally it reduces the number of intramolecular hydrogen bonds, stabilizing the well-known double helix structure, instead of more complex RNA's tertiary structures.

Thymidine: the fourth musketeer

DNA sequences are encoded using deoxythymidine, that base-pairs with deoxyadenosine. This deoxynucleotide hasn't a corresponding ribonucleotide, and it's produced by dUMP methylation (*thymidilate synthase*). This compound, in turn, is mainly (over 80%) produced from deamination of dCMP (*dCMP deaminase*).

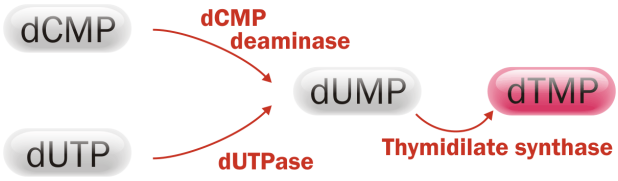


Fig. 2.3.2. - Deoxythymidine is produced partly from UDP (reduced to dUDP and), but mainly from dCMP.

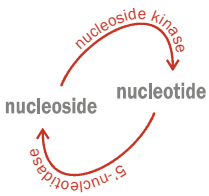
Salvage pathway

Salvage pathway rescues both nucleosides and free bases (derived from intracellular degradation or by cellular uptake) converting them into nucleotides, saving energy, carbon and nitrogen. Bases can be converted directly to nucleosides (by *nucleoside phosphorylases*) or interconverted. Nulcleosides are then phosphorlylated by *nucleoside kinases*. This step confers to the molecule a negative charge, trapping it into the cell.

In mammals four deoxyrobonucleoside kinases are present: TK1 (cell regulated and cytosolic), TK2 (mitochondrial and costitutive), dCK (cytosolic) and dGK (mitochondrial). Each kinase can accept different substrates, taking it name from the preferred one. Once the nucleoside monophosphate is formed, its subsequent phosphorylation are in common with enzymes of the *de novo* pathway.

5'-nucleotidases

Reaction catalyzed by *nucleoside kinases* is irreversible, but a class of cellular enzymes, 5'-nucleotidases[22], removes the phosphate from mononucleotides, giving back nucleosides.

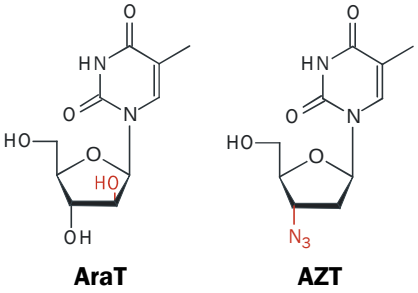


Nucleotidases and nucleoside kinases form a substrate cycle to maintain a balanced pool of precursor. This activity is fundamental, as a misbalance in nucleotide pool can cause severe errors during DNA replication[19].

Medical implications of deoxynucleoside kinases

Development of pyrimidine and purine analogs as potential antineoplastic agents evolved from an early presumption that nucleic acids are involved in growth control[20]. Among the first analogs produced and tested for biologic activity were the 5-halogenated pyrimidines:

5-chloro-, 5-bromo-, and 5-iodo-Uracil, when G. B. Elion recognized that these compounds might be valuable in the treatment of cancer, which was correctly perceived as a disease of uncontrolled growth.



AraT **AZT**

Further development of these analogs was stimulated by the demonstration of *quantitative, but not qualitative*, differences in the activity of these pathways between normal versus neoplastic tissue.

This is the common approach of **chemotherapy**, still broadly used in cancer and AIDS cure. One famous drug, used in HIV therapy, is the thymidine analog AZT (5'-azido-thymidine); this and another common analogue (AraC, arabinosil-cytidine) were both tested for phosphorylation by *Ce*TK1-M42I mutant. Traditional drug design relies on enzymes present in patient's cells, either of its genome or encoded by the virus.

The new approach of **gene therapy**[8], on the other hand, aims at introducing new genes in the patient's body. The world's first human gene therapy experiment began on 14 September 1990, when a 4-year-old girl with adenosine deaminase (ADA) deficiency was injected with some of her own, genetically modified, bone marrow stem cells. This "traditional" gene therapy was intended to supply a missing gene (by chance, that case was also involved in nucleotide metabolism, even if the protocol can apply for virtually any missing gene).

Suicide gene therapy is another approach, that consists in introducing in the human body a new enzyme for drug activation. This new approach can take advatage of new drug activator, as nucleoside kinases with new specificity,

Use of analogs in biotechnology

Knowledge acquired in different fields of life sciences research often results in new tools for gene technology, resulting in an "auto-fueling" research. One of the most famous examples is probably the discovery of restriction enzymes. Studies on Thymidine Kinase also provided a very useful tool for genetic engineering. TK from Herpes Simplex Virus (tk^{HSV}) is a widely used negative selection gene for different applications[5], for instance when a double selection is required, combining the traditional negative selection (eliminating cells without the marker gene, for example drug resistance gene) and a positive selection, as using a thymidine kinase capable of activating analogs not phosphorylated by the host encoded gene. This cause cell death in cells containing the gene, when the drug is added to the medium.

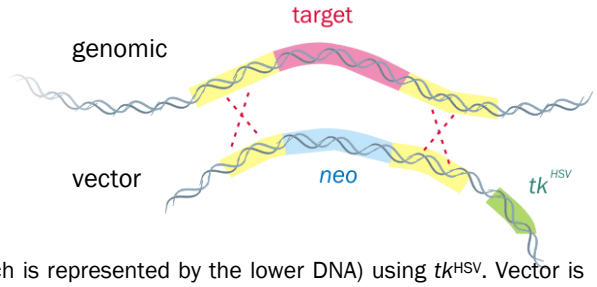
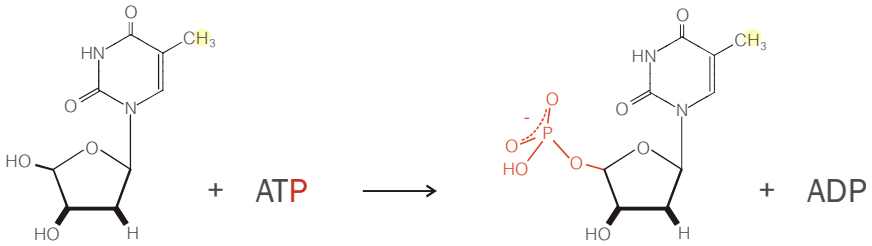


Fig. 2.3.3. - An example of vector (a sector of which is represented by the lower DNA) using tk^{HSV}. Vector is designed to disrupt a gene (target gene in pink, in genomic DNA), by homologous recombination mediated by the "homology tails" designed to be equal to target flanking regions (yellow). A double recombination is a rare event, and to prevent unwanted incorporation of the plasmid a double selection (neomicine + ganciclovir) is used. Double recombination, indeed, replace target gene with neo^R that confers neomicine resistance, but not the tk^{HSV} gene, that is outside the "homology tails".

2.4. Thymidine Kinase 1

Overview

In cellular metabolism, TK1 (UniProt **P**04183, cytosolic) and TK2 (UniProt **O**00142, mitochondrial) serve to maintain sufficient dTTP for DNA replication and repair, as involved in the first phosphorylation of thymidine nucleoside, using ATP as phosphate donor, as depicted in the scheme below (E.C. 2.7.1.21):



Reaction mechanism for HSV1-TK[16] involves two conserved residues: one Glutamic acid and an arginine; the former is proposed acting as a base in the deprotonation of the –OH group, while arginine should stabilize the transition state. Phosphates of the donor are stabilized by a Mg^{++} ion. According to Eriksson (2002) this mechanism should valid for other thymidine kinases.

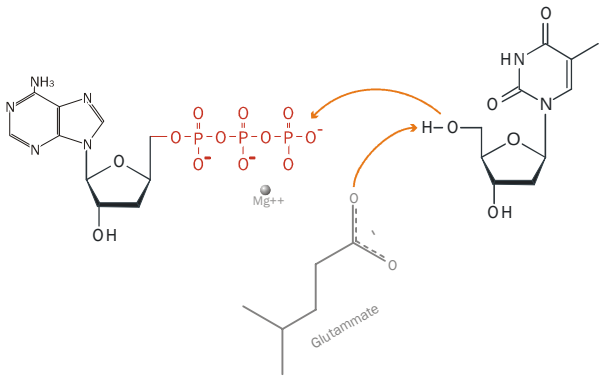
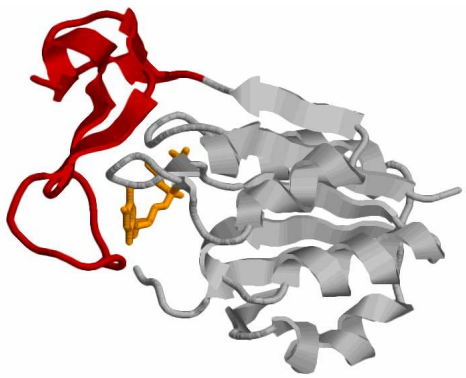


Fig. 2.4.1. – Mechanism of reaction showing Glu residue attaking 5'-OH of the substrate. Arginine is not showed.

Human cytosolic Thymidine Kinase 1 is a tetrameric cell-cycle-regulated enzyme[9]. Each subunit contains an α/β-domain (in gray in the fig. 2.4.1) and a domain containing a structural Zn^{++} ion (red in the picture), not related with other nucleoside kinases. The active site is between the two domains. Further details on structural properties of the Human orthologue are provided later in this paragraph. The cytosolic enzyme is shows a peack of activity during *S phase* of the cell cycle, when its presence is necessary. **Transcription** is promoted by E2F transcription factor, and activity is removed by mitotic degradation of the protein, that is polyubiquitinated by SCF-Cdh1 that recognize the signal sequence "KEN-box" at the C-terminus[10].



2.4.2. – Structure of a monomer of Human TK1 with its feedback inhibitor TTP in orange. PDB: 1XBT

The *Ce*-TK1 gene lies in chromosome IV and the complete mRNA is 1026 bp long, and It has 5 exons and encodes for a protein of 269 aa (M.W.: 29.6 kDa; pI: 6.0). Its product has been recently purified in Birgitte Munch-Petersen laboratory[21], where it was found to be a dimer. In the picture below the genomic organization of the gene is partly depicted (data from WormBase), while other genomic features from the UCSC GenomeBrowser are available on line (§1.3)[13]. It seems to be the only nucleoside kinase like gene in the whole genome[21].

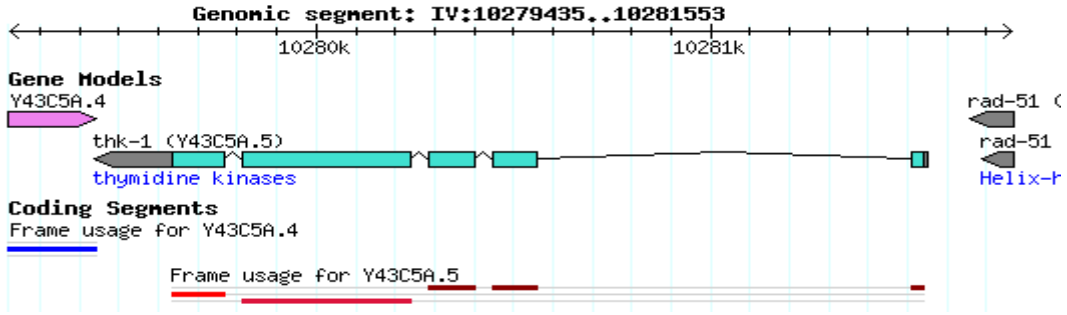


Fig. 2.4.3. – Genomic segment of *C. elegans* chromosome IV showing TK1's gene organization. Thk-1 gene is expressed in a common messenger with rad-51 (they are clustered as a dicistronic unit). From WormBase.

Biochemical properties of Human and *C. elegans* TK1

Human TK1 has been extensively studied[9], while characterization of the flat worm one is still in progress[21]. The human enzyme is more stable in a C-terminally truncated form[24]. Incubation with ATP of truncated or wild type Human TK1, affects both oligomerization[iv] and activity. As terminally truncated CeTK1 resulted less stable and with decreased activity, mutants were made on the wild type sequence, while HuTK1 mutants were made on TK1Δ40 sequence.

A table with a comparison of Human and *C. elegans* activity is shown below.

HuTK1[a]	**WT (25.5 kDa)**		**CΔ40 (21.1 kDa)**	
Storage conditions	**-ATP**	**+ATP**	**-ATP**	**+ATP**
k_{cat} (sec^{-1})	6.3±0.9	6.9±0.2	9.5±1.1	9.7±0.2
K_m (μM)	16±3.4	0.7±0.2	1.4±0.6	0.6±0.1
k_{cat}/K_m	0.4×10^6	9.9×10^6	0.4×10^6	9.9×10^6
CeTK1	**WT (29.7 kDa)**		**CΔ56 (23.6 kDa)**	
Storage conditions	**-ATP**	**+ATP**	**-ATP**	**+ATP**
k_{cat} (sec^{-1})	5.0±0.1	8.8±0.2	1.1±0.0	3.5±0.2
K_m (μM)	2.3±0.3	2.6±0.3	7.4±0.8	5.8±0.9
k_{cat}/K_m	2.2×10^6	3.4×10^6	0.1×10^6	0.6×10^6

Table 2.4. – Kinetic values with thymidine as substrate. After T. Skovgaard 2005 [21].

Sequence alignment

Several highly conserved regions are delimited in the alignment between TK1 from different organisms (from bacteria to higher eukaryotes). X-ray structure has been resolved for *H. sapiens* and *U. urealyticum*, and, being the former most related to the worm one, will be used as model. The two sequence are highly similar, with an evident extension of both C- and N-terminus in *C. elegans*.

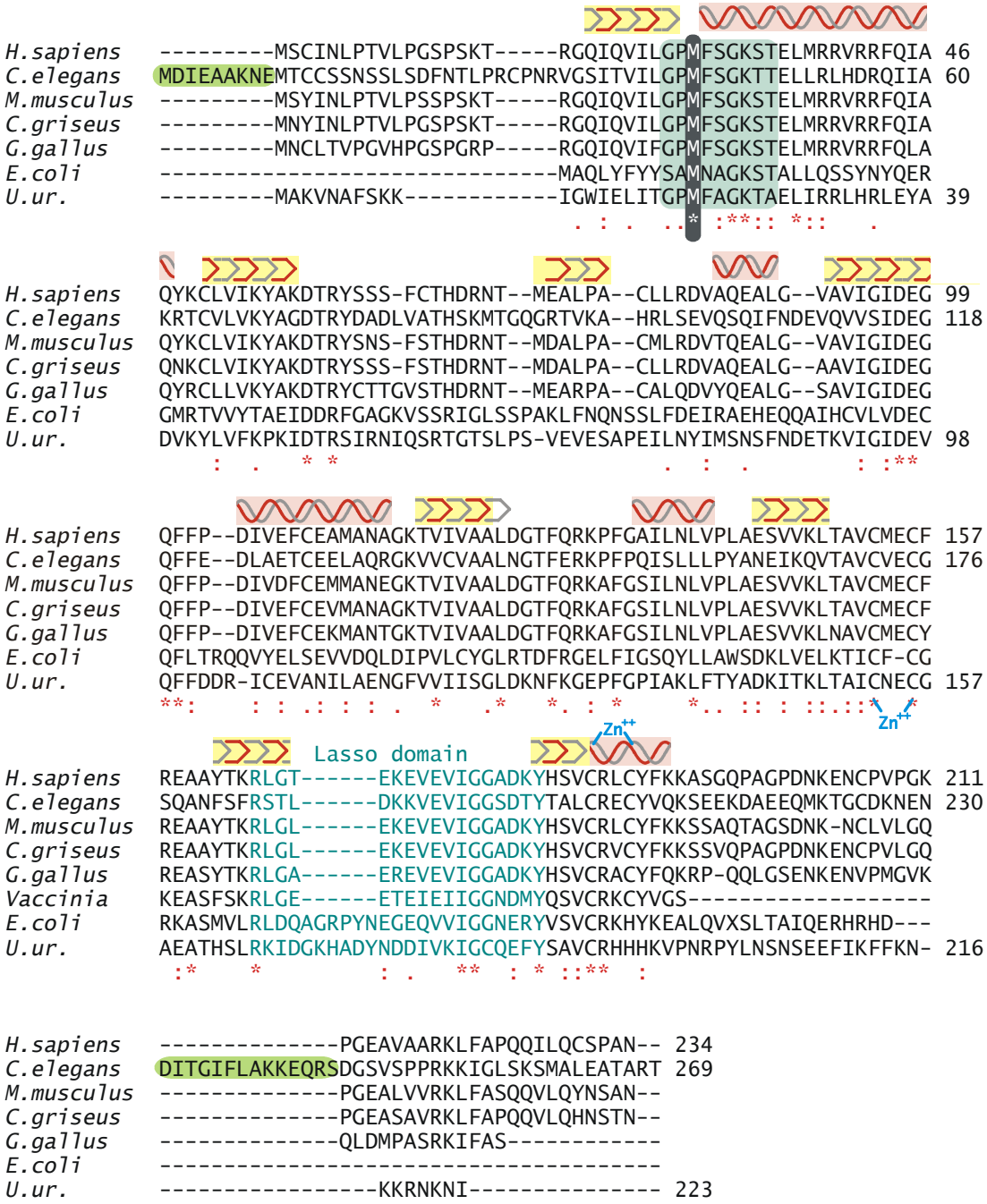


Fig. 2.4.4. – Multiple alignment between different TK1 (performed with ClustalX for Windows, Blosum64 matrix). Secondary structure as annotated in PDB file (1XBT) is also aligned, the mutation site is in black, an the P-loop containing it is highligted in light cyan. Zinc-coordinating cysteines are annotated, and aminoacids of the "lasso domain" are writtein in light cyan.

Structural properties of the Human orthologue

The structure most related[v] to studied *Ce*TK1 available in PDB is the Human TK1, resolved on 2004[15]. This enzyme showed unexpected features just when it was diffracted in Grenoble: first a Zinc ion tightly coordinated by four cysteines, and second its striking affinity for the feed-back inhibitor dTTP, that was found in the structure even when never added to a medium, meaning that it can pass several purification steps after extraction from the expression host (*E. coli*). Human TK1 can associate in tetrameres in presence of ATP (picture on the left), while from studies being currently carried out the *C. elegans* TK1 seems to aggregate only in dimers[20].

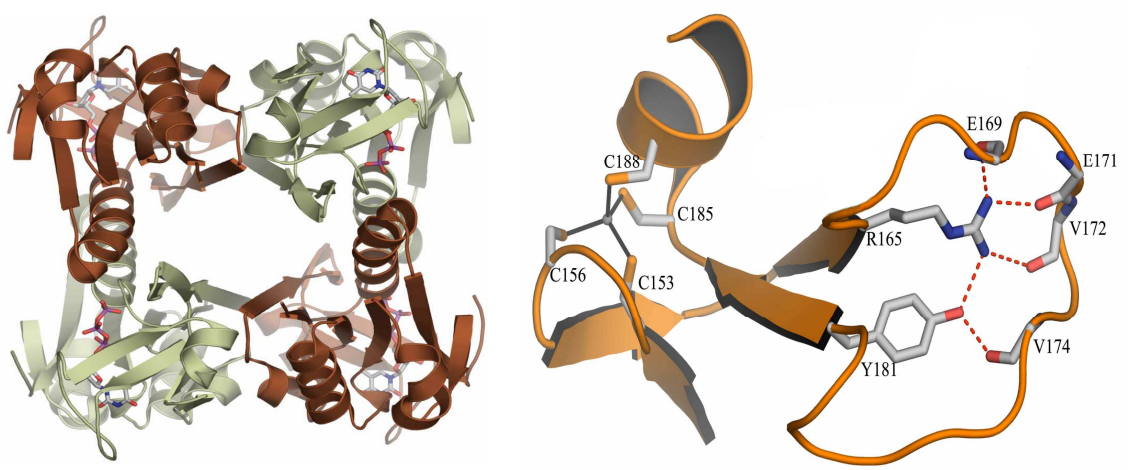


Fig. 2.4.5. – On the left the human tetramere. On the right part of the Zinc-containing domain, with the Lasso domain (right), and residues involved in loop stabilization are shown, and the Zinc-coordinating cysteines (left). From Welin et. al 2004 [15].

The monomer has an α/β-domain linked to the Zinc containing one by a single-loop. The first has its β-sheet lying on a plane and covered by α-elices. The other domain is the most peculiar of TK1 family for the presence of a Zinc ion, and contains both the **P-loop** motiv, involved in phosphate-donor binding, and the "**Lasso domain**", well structurated thanks to several hydrogen bonds occuring between its residues, expecially a Arg–Tyr couple placed at the extremities of the lasso (picture on the right). Lasso domain is an essential component of the active site.

After resolving the structure several aminoacids where found eligible for mutagenesis and kinetics analysis, remembering that the fixed structure from the X-ray diffraction is a model that is partly distant from reality, where enzymes can "breathe" as result of atomic vibrations, changing thus the steric conformation of the active site itself.

Figure 2.4.5. shows a section of the outer electrostatic surface, where a pocket is clearly visible near the active site. This could be the entrance channel for both substrates and feedback inhibitors. On the right the sidechain of Met[28] (emerging from P-loop) in the context of the whole molecule. Lasso domain is evident on the top of Zinc-domain (left), and also the "beta sandwich" on the right. TTP molecule is shown.

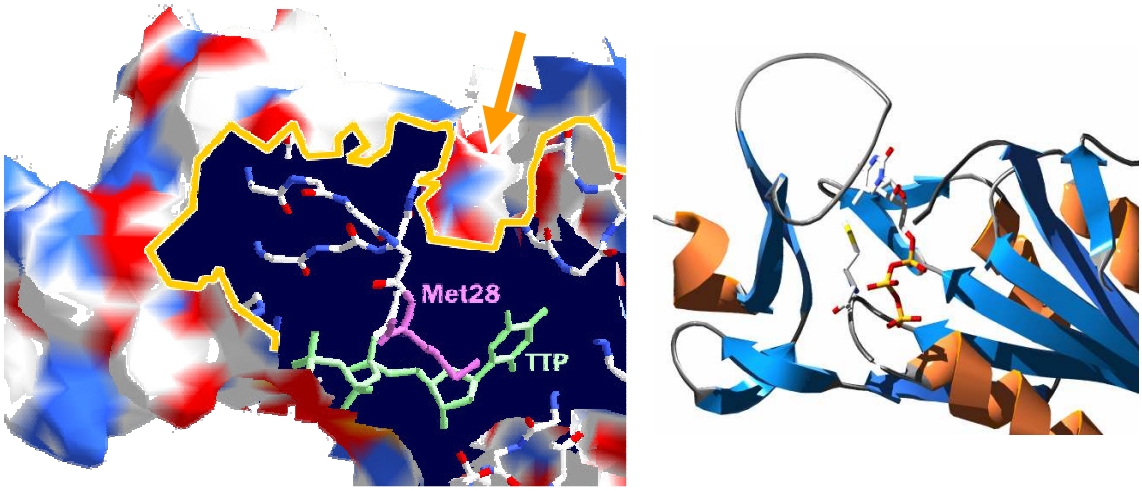


Fig. 2.4.6. – On the left an intresting section of the molecular surface reveals a cleft between the two domains of the enzyme. TTP is shown in green. On the right the context of Met[28] and TTP in the molecule.

With PDB viewer from GSK, a possible vision of the active site after mutagenesis was produced (fig. 2.4.6.) using Human TK1 as model, and choosing between the three isoscored rotamers the one with sidechain in the position similar to Methionine 28.

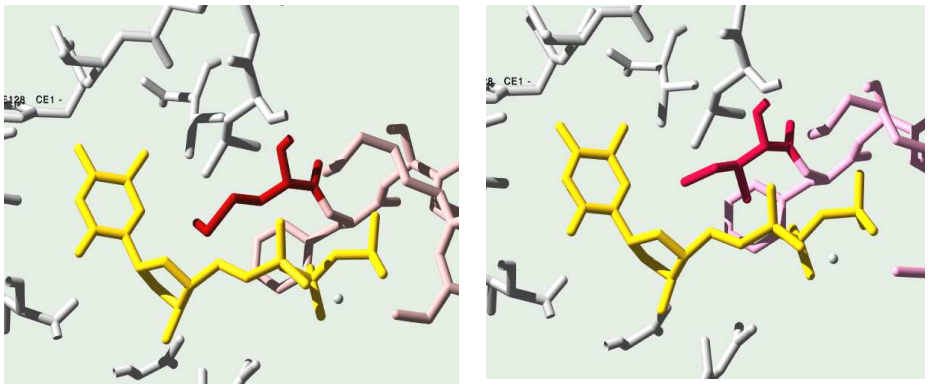


Fig. 2.4.7. – Steric comparison between Met[28] as found in the PDB structure (*left*) and a possible rotamer of Ile[28]. Rotamer was choosen between three that were equally scored by PDB viewer[vi].

Methionine and isoleucine are aliphatic aminoacid of the "VILM group"[25], but isoleucine has a branch in β-carbon (shown in molecular model), that could act as destabilizing element of the structure.

Considering that methionine is highly conserved, and the difference of with isoleucine both from a chemical (lack of S atom) and sterical (β-branch) it's easy to suppose a disruption of wild-type activity, but nontheless a change in specificity range could be achieved with this mutation.

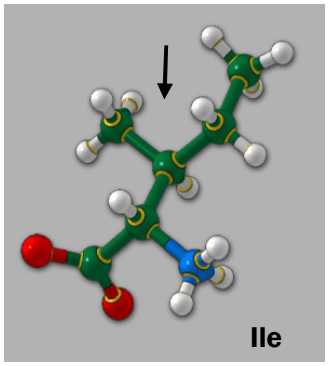


Fig. 2.4.8. – Molecular model for amino acid Isoleucine.

2.5. Project overview

Main steps of the project are summarized in the scheme below:

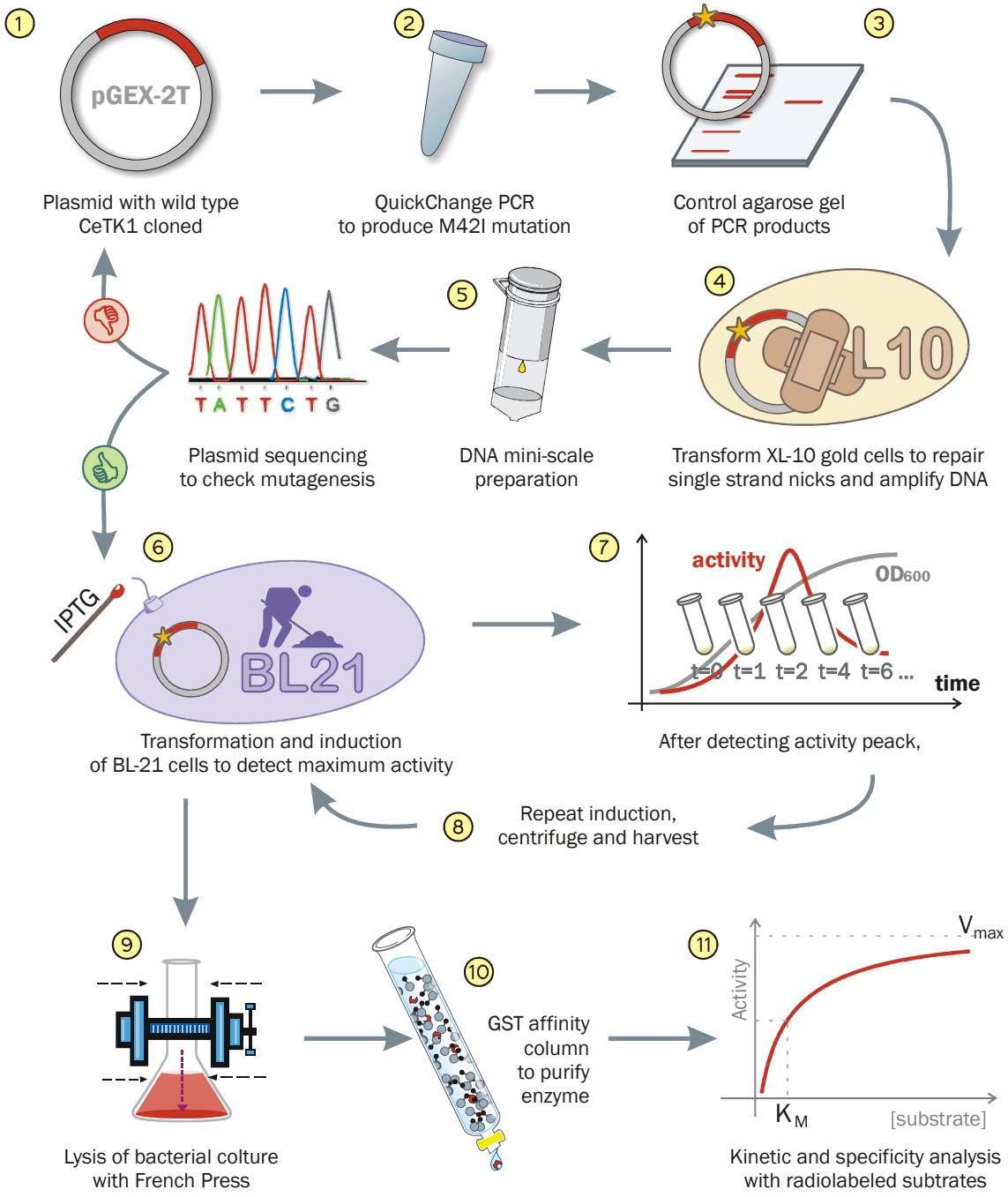
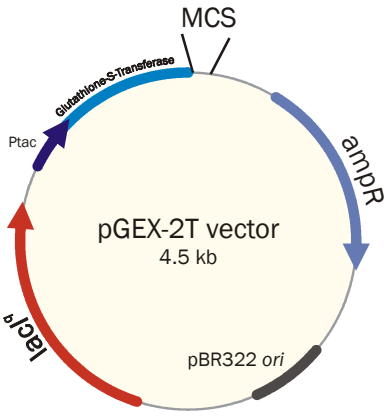


Fig. 2.5.1 – The QuickChange™ protocol is based on a modification of a PCR reaction (2), that is checked with an agarose gel electrophoresis (3). Mutagenesis require also a treatment of PCR products with *Dpn*I endonuclease and transformation of *E. coli* XL-10 gold to repair single strand nicks (4). Plasmid is extracted with a MiniPrep Kit (5), that is sequenced for control. BL-21 strain of *E. coli* is then transfomed with the plasmid extracted, and a first culture is inducted (5) taking several time samples to monitor bacterial growth (via OD lecture) and activity of thymidine kinase (via radiolabeled enzyme essay) through time (7). Once the best time to harvest cells is detected, a new induction is performed. The culture is then passed through the French Press to perform bacterial lysis (9) and TK1 is purified from crude bacterial homogenate via GSH affinity column, thanks to GST domain fused with TK1 in the pGEX plasmid. Kinetics and specifity are then analyzed.

3. Experimental background

3.1. Site directed mutagenesis



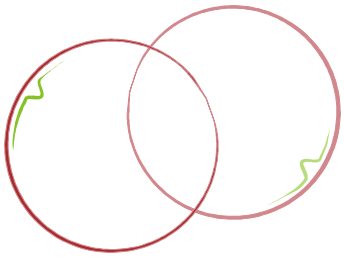
The vector pGEX-2T with inserted *Ce*TK1[vii] was mutagenized with the QuickChange™ protocol by Stratagene®, that allows site-specific mutation in virtually any double-stranded plasmid, thus eliminating the need for subcloning into M13-based bacteriophage vectors and for ssDNA rescue. In addition, the QuickChange™ site-directed mutagenesis system requires no specialized vectors, unique restriction sites, or multiple transformations. QuickChange™ strategy is a variation of a PCR, where the peculiar primer design permits an efficient mutagenesis.

QuickChange™ strategy overview

Step 1 – Primer design

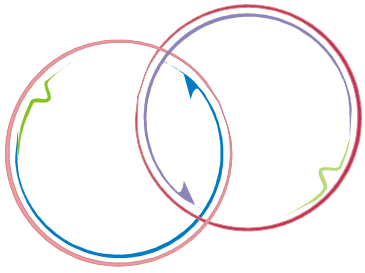
Plasmid pGEX-2T, with cloned CeTK1 insert, is shown with a green fragment in the site to be mutagenisized.

PCR primers are designed to contain the mismatch in the site to be changed, surrounded by two “complementary tails” long enough to bind the target sequence in the vector. In addition they are reverse complementary each other.



Step 2 – Denaturation in the PCR thermal cycler.

Plasmid DNA (pink) melting is followed by mutagenic primers (in green, with mismatch shown as a ^) annealing. Thermal cycler is programmed to permit elongation of the whole plasmid.



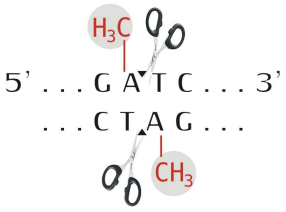
Step 3 – Mutagenesis

The *Pfu Ultra High Fidelity* DNA polymerase[viii] starts replication of both strands (only in the first cycle, see next paragraph), each with its primer, while the primer is encountered. A high fidelity DNA polymerase is recommended, a non strand displacing one is mandatory.

Step 4

After the amplification in the thermal cycler, the DNA is digested with *Dpn*I endonuclease, that cuts only methylated or hemi-methylated DNA. Remaining molecules are mutated, but still contains two nicks, that will be repaired transforming *E. coli* XL-10 gold, a strain lacking exonucleases (having a nick).

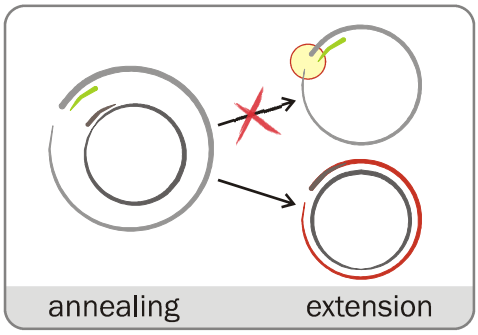
Notes on the QuickChange™ protocol



The basic principle of the protocol is the incorporation of mutated primers in new molecules. The removal of parental molecules, containing a mismatch, is performed cutting the reaction mix with the *Dpn*I endonuclease, that cuts only methylated or hemi-methylated DNA at the specific site shown in picture. This requires that the original plasmid was not duplicated in *E. coli* strains lacking methylases. As only the parental plasmid was methylated[ix] in *E. coli*, the two hybrid molecules descending from the plasmids are degradated.

The protocol suggests using *Pfu* Ultra High Fidelity DNA polymerase. Fidelity level apart, it's important that the DNA polymerase used hasn't a strand-displacing behaviour, so that at the end of the cycle, when the 5' of the primer is encountered.

The yield of DNA is lower than traditional PCR, as the amplification is linear instead of being exponential. As the picture on the right shows, for each molecule, only one copy per cycle can be obtained, as one of the primers is placed near a nick produced on previous cycle, when the DNA polymerase encounters the primer at the end of the extension.



Primer design

Primers have to contain the desired mutation and guarantee an adequate melting temperature for a PCR reaction. I wrote a simple perl program (prim.pl, available §1.3) to autochoose the smallest primer that has the mutation in quite a central position and melting temperature, calculated with the empirical formula reporter below, above or equal to 77 °C.

$$T_m = 81.5 + 0.41 \cdot (\%GC) - (675/Sequence\_length) - \%(mismatch)$$

The program suggested a sequence with T_m of 77.4 °C that seemed suitable for the protocol, and accepted as it was.

Primer sequence suggested:

```
fwd: 5'-GGG GCC AAT TTT CAG TGG CAA AAC CAC CG-3'
rev: 5'-CGG TGG TTT TGC CAC TGA AAA TTG GCC CC-3'

Parameters:
T_M=77.4°C          n=29          %GC=55%
```

Fig. 3.1.1 - Output of the program. The two primers with the changed position highlighted.

Mutagenesis control

The pWhitescript™ 4.5-kb control plasmid is used to test the efficiency of mutant plasmid generation using the QuickChange site-directed mutagenesis kit. The pWhitescript 4.5-kb control plasmid contains a stop codon (TAA) at the position where a glutamine codon (CAA) would normally appear in the β-galactosidase. Competent cells transformed with this control plasmid appear white on LB–ampicillin agar plates, containing IPTG and X-gal, because β-galactosidase activity has been obliterated. But, if the oligonucleotide control primers create a point mutation on the pWhitescript control plasmid that reverts the T residue of the stop codon, a fully functional β-galactosidase is synthesized, thus, colonies transformed with mutated plasmid appears blue.

Final steps

An exonuclease deficient *E. coli* strain was transformed with the PCR products, after the incubation with the *Dpn*I endonuclease. This allows:

- Nick repair by host's DNA ligase;
- DNA amplification in a high-fidelity strain;

Cells were made competent and transformed with the $CaCl_2$ chemical transformation protocol (see Appendix).

A miniscale preparation of DNA was then performed to:

- DNA sequencing to check mutagenesis;
- Obtain pure plasmid to transform *E. coli* BL21 cells (protein expression);

DNA sequencing

Plasmid sequencing was performed by an external company (MWG Biotech), with the standard Sanger method, and both forward and reverse sequence were requested, using universal primers designed for pGEX-2T. The company provides both electropherograms and text sequences. These were aligned with a perl implementation of the Needleman and Wunch algorithm, that provided a complete sequence of the plasmid insert.

This was then aligned with the wildtype sequence from GenBank using the same algorithm, but also the ClustalX program for Windows, that has an embedded multialignment algorithm. The graphical output readily confirms if the mutagenesis protocol properly worked.

3.2. Radio-labelled enzyme assay

To investigate enzyme kinetics it's necessary to have the possibility of measure the progress of reaction (i.e. formation of product or lowering of substrate). The reaction catalysed by deoxynucleoside kinase is:

Main product of reaction (phosphate acceptor) is negatively charged, while corresponding substrate is not. Spotting an aliquot of the reaction to a filter where only charged molecules can bind, and labelling substrate^{x} (again phosphate acceptor) so that only the product will be detectable, allow us to measure the reaction rate.

Radioactive decay of unstable isotopes

Substrates (like deoxythymidine and AZT) used in activity assays are labelled with Tritium, the artificial and unstable isotope of Hydrogen. This atom shows a decay of this type:

$$\text{Neutron} \rightharpoonup \text{Proton}^{+} + \beta^{-} + \text{Antineutrin}$$

The proton produced remains in the nucleus, that has its atomic number increased by one unit (i.e. ^{3}H becomes ^{3}He). The total energy emitted by this process is equally shared by the β particle and the antineutrin, so the β particle (that is the detected product) can be emitted with a range of different energies.

For any population of unstable atoms the **nuclear decay rate** is proportional to the number N of decomposing nuclei present:

$$dN/dt = \lambda N, \qquad \text{or} \qquad N = N_0 e^{-\lambda t}$$

The latter form is useful to introduce the **half-life** concept ($t_{½}$), which is the time after which half of the population of unstable nuclei deceased.

Tritium half life is 12.3 years, and the average energy of the emitted β particle is 0.0055 MeV. The unit of radioactive particle emission is the Curie, Ci, defined as the number of disintegrations per second of a gram of impure Radium ($3.7 \times 10^{10}\ s^{-1}$).

Scintillation counting

Liquid scintillation counting is the most efficient method to detect β^{-} particle emission, that, being of relatively low-energy, needs to be in intimate contact with the molecules used for detection. The sample is placed in a scintillation mixture containing an excitable solvent, and one or more fluorescent compounds, called "**fluors**".

Emitted β^{-} particles collide with the solvent (S), and transfer part of its energy to this molecule, originating an excitated molecule (S*), and a β^{-} particle with less energy. The average

energy of a β^- particle is usually greater than the energy required to excite a solvent molecule, so the same emitted particle can often excite more than one solvent molecule before its energy is depleted. Then, the population of excited solvent molecules transfers its energy to fluor molecules, which then decay to ground state with emission of a photon.

Atomic decay is a statistic process, and thus a large amount of counts has to be recorded to have a measure of the radiolabelled molecule present in the test tube. Detection itself is a difficult task, as, especially for the low-energy isotope Tritium, it's difficult to register all disintegration.

Machine used in this issue, the **liquid scintillation counter** (picture on the right), has to detect photons emitted by the tube, converting them in a difference of potential (voltage). New machines have very sensitive photo-rilevators, thus with a serious problem of back-ground noise. This can be partially solved using two photo-rilevators, and recording only disintegration detected by both of them, this unfortunately reduces sensitivity. CPMs (counts per minute) recorded, in fact, are a fraction of the total emissions. It's common to **use a standard** with known concentration of tritiated substrate to solve this problem.

Enzyme kinetics assay

To detect the rate of enzyme catalysis a tube with enzyme and labelled substrate is incubated at optimal temperature (37°C for TK1). After at least three time gaps an aliquot of the reaction mixture is spotted on a **DEAE filter** (see fig. 3.2.1), that tightly binds negatively charged molecules. This stops the reaction, as one of the reactants (ATP) is subtracted from the mix (see fig. 3.2.2).

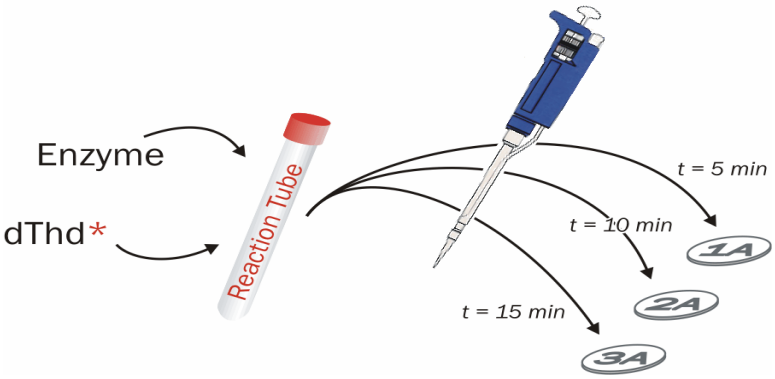


Figure 3.2.1 – A series of tubes with different substrate concentrations are prepared. Then enzyme is added, and the reaction starts. After regular intervals an aliquot is taken from the reaction tube and spotted on a DEAE filter, where the reaction is stopped as ATP (and newly synthesized TMP) binds to filter.

Figure 3.2.2 – DEAE cellulose allows keeping only a fraction of the radiolabelled molecules: the phosphorylated ones (products).

Each filter is washed to remove every uncharged molecule, then is placed into a scintillator tube, where **elution buffer**, that allows bound molecule to leave the filter, is added. Then the scintillation liquid (containing fluor molecules), is added as well, so that samples are ready to be loaded in the liquid scintillation counter.

Experimental design

This general design has been used in our project for two tasks:

- Activity assay on bacterial crude lysate (*Activity assay*)
- Michaelis-Menten kinetic assay (*Kinetic assay*)

In the first case the aim is to detect activity on a sample, usually from time samples taken from protein induction, and in the tube crude lysate is placed first, then reaction starts adding substrate (in a definite concentration for each sample).

When performing a Michaelis-Menten assay there is the need of varying substrate concentration, so substrate itself is placed first, then, at time zero, the reaction mix with enzyme and ATP is added[xi].

Results analysis: concepts

As said before, each kinetic assay is performed with at least four standards (i. e. unwashed filters containing both reactants and products). Using standards is possible to convert CPMs to picomole of substrate as follows.

CPMs recorded for each standard are divided for concentration of substrate and for volume of reaction mix spotted on standard filters (5 µL), resulting in *standard CPMs per picomole of substrate* (mean of value for each filter should be equal unless experimental error).

Now this value establishes a link between CPMs and picomole of substrate. It's easy to convert each CPM value recorded to picomoles of substrate with a conversion factor, shown in red in the formula below:

$$\text{CPM} \cdot \left(\frac{\text{Assay Volume}}{\text{Spotted Volume}} \cdot \frac{\text{[Substrate] x Standard Volume}}{\text{Standard CPMs}} \right) = \text{Substrate pmol}$$

Where "Substrate, pmol" is the amount of substrtae (in pmol) that was spotted in the filter[xii]. The unit "*nmol/min*" is often referred to simply as "units" and is abbreviated u.

Result analysis: tools

Data from scintillometer are transferred into an Excel™ worksheet, where a first analysis is performed to detect slopes as described in the previous paragraph. Obtained data of the Michaelis-Menten plot (velocity against substrate concentration), they should be analyzed with a statistical software capable of non-linear curve fitting to estimate parameters (K_M and V_{max}) without spreading experimental error. See also §3.7.

3.3. Gel electrophoresis

The "*killer technique*" in molecular biology is probably gel electrophoresis, allowing investigators to separate macromolecules, as nucleic acids and proteins. In our project agarose gel electrophoresis was used to separate PCR products, and a polyacrylamide gel for proteins instead.

Common features

In both cases molecules are forced to run in a gel with an electric field acting on macromolecular charge. Small molecules find easily a path within the grid of the gel, while for bigger ones migration takes a longer while. This results in a separation of molecules based on mass/charge ratio[xiii]. Nucleic acids and most proteins are transparent to visible light, so the gel is stained with a coloured molecule binding to them.

Agarose gel electrophoresis

Gel is made with agarose, a polysaccharide extracted from the alga *Agar agar*. DNA is negatively charged at neutral pH, so migration occurs to the anode. Gel is stained with Ethidium Bromide, a planar hydrophobic molecule of condensed aromatic rings, that fits between the base pairs, and that, receiving UV lights, emits a visible pink light.

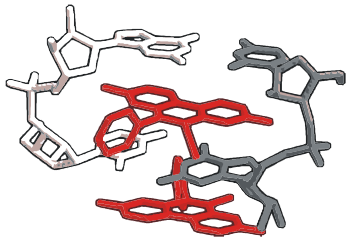


Fig. 3.3.2. – Tho molecules of ethidium (red) intercaled in a dsDNA dinucleotide (gray).

SDS-PAGE

Proteins are usually separated after denaturation with the SDS detergent, that is negatively charged and interacts with protein backbone about one molecule every two residues. Gel is made polymerizing bis-acrylamide. This results in linear proteins, whose charge is mainly determined by the number of molecules of detergent (proportional to peptide length). After running the gel is stained with Coomassie Brilliant Blue, that binds to proteins conferring to resulting band a blue colour.

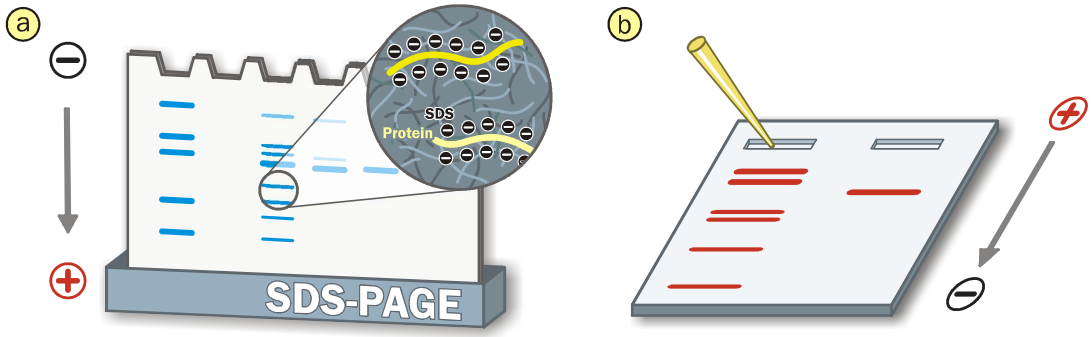


Fig. 3.3.2. – (a) SDS-PAGE is performed with scaffolding that keeps the gel vertical. (b) Agarose gel electrophoresis is performed horizontally, adding Ethidium Bromide to the gel and surrounding buffer to stain DNA.

3.4. Protein expression

The open reading frame of the *C. elegans* TK1 was cloned into a pGEX-2T vector, whose polylinker is in fusion with a domain of the Glutathione-S-Transferase (GST) enzyme, and with a short loop spacing the two domains and containing a Thrombin cleavage site. The resulting fusion-protein thus contains a domain that tightly binds to GSH matrices, simplifying purification.

Transformation of BL-21 cells

E. coli strains used for protein expression lacks some of their proteases. Used strain to produce the mutant TK1 was *E. coli* BL21, that lacks both the *lon* and *OmpT* protease-encoding genes. BL21 competent cells are to be transformed with the miniscale preparation of DNA from XL-10 gold cells and plated on a Petri dish.

From a single colony then a liquid medium culture is prepared, and only after allowing bacteria to freely grow in the medium inductor (IPTG) is added. This ensures quite a high yield of protein, as the expression of a foreign gene is usually grow-inhibiting (when not toxic). In particular, overxpression of a nucleoside kinase will introduce a misbalance of host's nucleotide pool, thus impairing bacterial health. This makes it difficult to obtain large amounts of this enzyme, and explain the importance of and inducible expression system.

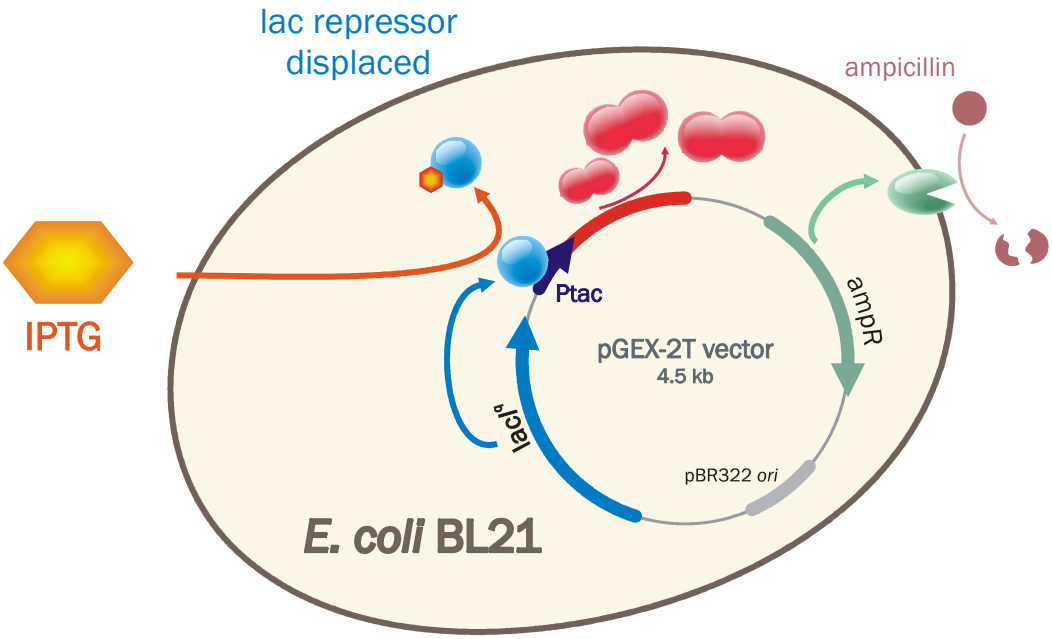
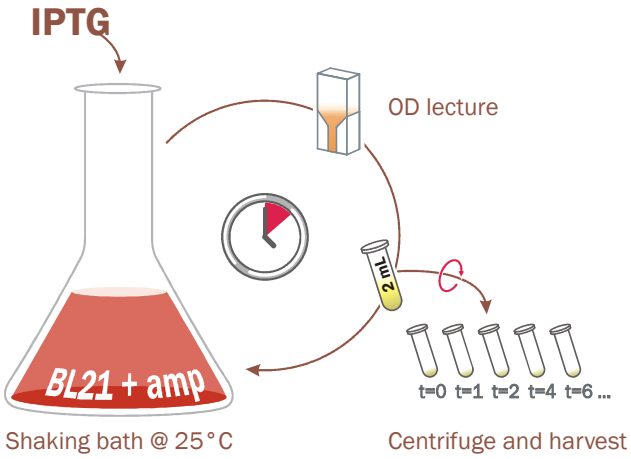


Fig. 3.3.1 – Schematic overview of the expression system based on *E. coli* BL21. After adding IPTG in the medium the strong repressor **lacI^q** is displaced from the Ptac promoter, and the transcription of the fusion protein starts. Selection of transformed bacteria is based on the enzyme β-lactamase, encoded by *amp^R* gene, that confers to the host cell resistance to ampicillin degradating it.

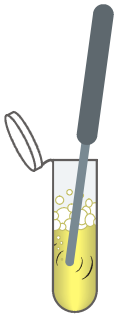
Time samples

A first induction is performed to monitor the TK1 activity, and consequently choose the best interval to lysate cells when there is the maximum of activity. After regular intervals from the addition of IPTG to the flask (that is “time 0”), 2 mL of culture are centrifuged and the supernatant is discarded, and an OD_{600} lecture is performed, in order to follow bacterial growth and see if and how is affected by the production of recombinant protein.

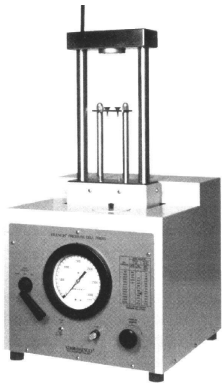
Activity of time samples is measured resuspending pellets with a lysis buffer, and sonicating (see next paragraph about lysis) them to release cytoplasmatic content. Then, the lysate is analyzed with a radio-labelled enzyme assay, resulting in a plot of activity against time.



Lysis



The small time samples are lysated by **sonication**. Resuspended in 400 µL of lysis buffer, that contains detergents to facilitate solubilizing molecules and lysozime that starts and enzymatic lysis of cell wall. An ultrasound emitter is then immerged in the tube – kept on ice to avoid enzyme denaturation – where ultrasound waves produces small bubbles that imploding generate high-energy waves propagated in the medium, resulting in disruption of cells, that release their cytoplasm in the solution (scheme on the right).



The bigger amount of culture is lysated with a **French press** (picture on the left), that is mechanically pressing the resuspention. This is forced to pass through a narrow hole within the machine whose picture is on the left. In this case the lysis buffer is the same as sonication buffer, lacking of lysozime.

In both cases a centrifugation, to remove larger residues as ribosomes and wall pieces, is needed.

3.5. Protein purification

After centrifuging the pressed lysate, this is filtered to remove bigger contaminants, and now we refer to it as CBH (Crude Bacterial Homogenate). This contains the recombinant protein of interest: fusion of GST domain and *Ce*TK1-M42I.

The GST tag is used in the affinity chromatography column, that consists of at Sepharose matrix with glutathione (GSH) molecules conjugated to it. When the homogenate passes through the column matrix the GSH molecules binds the GST part of the GST-*Ce*TK1 fusion protein. Thoroughly washing the column will remove residual proteins and cell residues leaving only GSH- *Ce*TK1 in the column. Addition of a thrombin containing buffer, and subsequent incubation with it, causes *Ce*TK1 to be cleaved from the GST protein and it can thus be washed out of the column and collected.

The purity of protein in the different column fractions collected during the purification process is examined by SDS gel electrophoresis, Bradford protein determination and for thymidine activity.

The purest fraction with the highest protein content (which is often the first cleavage fraction) is then used for further analysis of the enzyme.

Fraction are labelled as CBH (Crude Bacterial Homogenate), SET (elute non-tagged molecules), W1 to W5 (washes of the column), EQ (equilibration buffer for cleavage), CL1 to CL3 (cleavage fractions eluted after Thrombin incubation) and finally GSH for the final GSH wash. Column, after use, has to be regenerated.

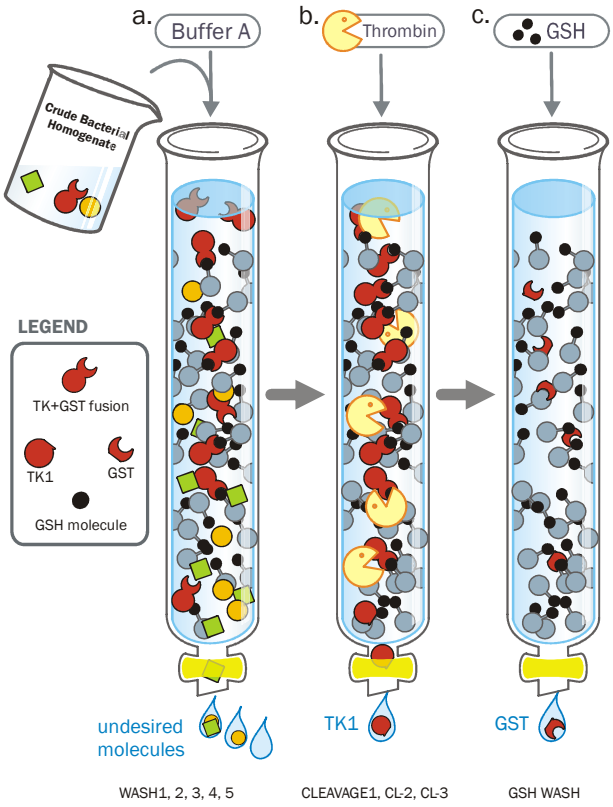
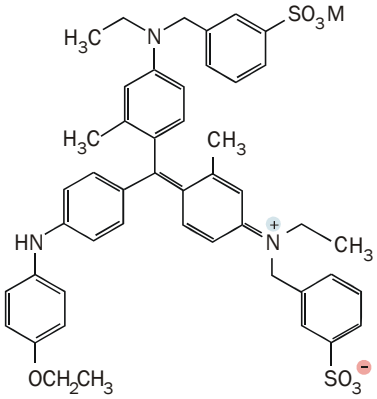


Fig. 3.5. – This scheme shows column affinity purification main steps. First fractions (a) contains molecules of the crude homogenate but the target of purification. Then (b) protease (Thrombin) is added and incubated in the column, resulting in the TK1 fraction. A wash (c) with reduced GSH allows collecting GST domain, to check the performance of cleavage by thrombin.

3.6. Bradford quantification



The Bradford analysis[26] is a very popular protein assay method to determine protein concentration in a solution because it is simple, rapid, inexpensive and sensitive; it's based on the interaction between the **Coomassie brilliant blue G-250** (CBBG, structure on the right) stain and proteins, that is completed in about one minute, and remains stable for over an hour.

The mechanism of dye binding to protein was investigated by measuring CBBG absorbance spectra during titration of the dye reagent in the absence of protein and the response of the CBBG assay to various polyamino acids (keeping the protein /dye ratio high, in order to ensure that only high-affinity sites were occupied). As result of these studies, dye was shown specifically binding to proteins at phenylalanine, tryptophan, tyrosine, histidine, lysine and arginine residues; to the latter eight times as much as the other listed residues.


Samples for Bradford quantification

The CBBG dye reagent can exist in three forms: a cation (red form with maximum absorbance at 470 nm), a neutral form (green with maximum at 650 nm), and an anionic form (blue with maximum at 595 nm), shown to form a stable complex with basic amino acid residues of proteins, mainly with arginine and aromatic amino acids.

This blue anionic form of the CBBG dye exists in negligible concentration at the pH of the assay reaction mixture but binding to proteins shifts the equilibrium toward its formation, thus the solution gets as more blue as more proteins are present. In the photo on the left different cuvettes shows a gradient of colour from blue (highest concentration of protein). to brown (no protein is present).

The estimation of protein concentration measuring OD_{595} is based on a standard curve performed with serial dilutions of a known protein (BSA). The curve is automatically computed by the spectrophotometer that gives for subsequent measures the protein concentration (expressed as mg/mL). At higher protein concentration a saturating effect is detectable as a bend of the curve; for this reason protein samples have to be diluted if their adsorbance is over 0.8 A.

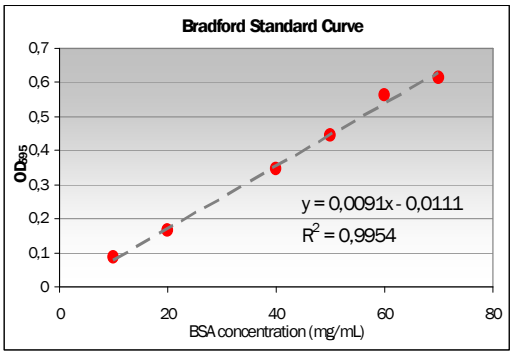


Fig. 3.6.1 – Standard curve obtained staining serial dilutions of BSA (10, 20, 40, 50 and 60 mg/mL) with Bradford reagent, and measuring adsorbance at 595 nm with the spectrophotometer [CSD, see §1.2].

3.7. Kinetic analysis

The kinetic behaviour of the mutant was analyzed with the Michaelis-Menten model (see §2.1.) by a radio-labelled assay (see §3.2.). The experiment is designed to measure the initial velocities at an adequate range of substrate concentrations to see the typical saturation curve of enzymes. Estimation of V_{max} and K_M from the Michaelis-Menten plot is quite difficult if hand-made, as guess of the asymptote is arbitrary. Michaelis and Menten originally proposed a linearization of their equation. Broadly used is the Lineweaver-Burk plot, a linearization obtained plotting the reciprocal of velocity against reciprocal of substrate concentration. This treatment, nevertheless, suffers of bad error propagation, also resulting in good-looking plots with inaccurate data. On the other hand the Eadie-Hofstee plot has better error propagation, and even small inaccuracy is made evident with this plot.

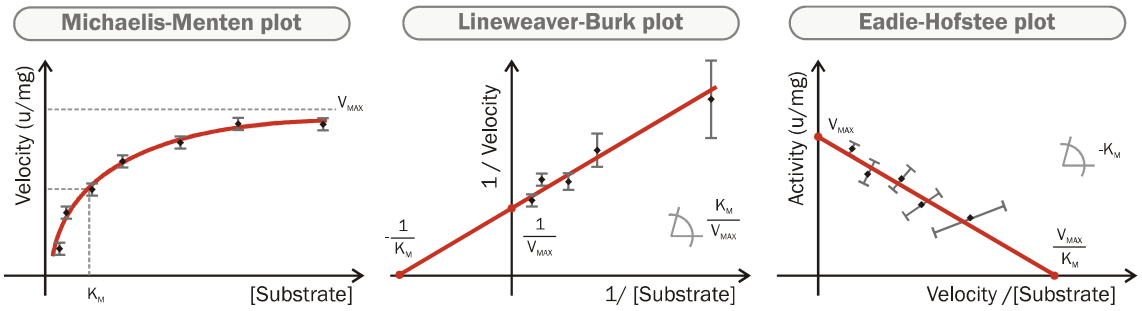


Fig. 3.7.1. – Popular kinetics plots. (a) Michaelis-Menten hyperbole, (b) Lineweaver-Burk linearization (also known as double reciprocal plot) and (c) Eadie-Hofstee linearization. Error bar shown for comparison.

With the advent of personal computing linearization should avoided, unless for a quick first look at data[xiv]. Non linear curve fitting allows estimating parameters from a dataset, given a model. The power of this tools raises the risk of under valuating the importance of choosing the right model, as a result is always given.

> *“A mathematical model is neither a hypothesis nor a theory. Unlike scientific hypotheses, a model is not verifiable directly by an experiment. For all models are both true and false.... The validation of a model is not that it is "true" but that it generates good testable hypotheses relevant to important problems.*
>
> **R. Levins, Am. Scientist, 1966**

For the simple Michaelis-Menten equation two parameters has to be estimated, so a planar space represents all possible Michaelis-Menten equations.

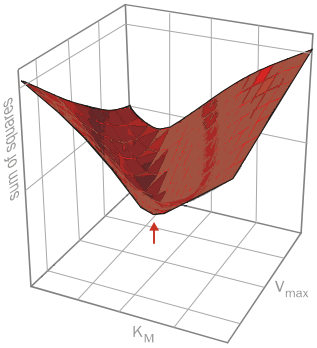


Fig. 3.7.2. – Solution of the non linear fit in a multidimensional space.

A “**fit criterion**” has to be used to choose the best curve of the space: most analytical software uses “sum-of-squares”: this is the third dimension of the space fig. 3.7.2., the one to be minimized. The user has to initialize parameters, then the software look for the direction downhill to the minimum.

4. Results

4.1. Mutagenesis

Mutagenic PCR

After running the PCR for mutagenesis (QuickChange™) an agarose gel electrophoresis was performed, loading on lanes a marker, the PCR products, and the negative control (PCR mix without plasmid) to check the absence of contaminating DNA.

A band is clearly visible where expected, so remaining PCR products are treated with *Dpn*I endonuclease to remove unchangend parental plasmid.

Competent cells (*E.coli* XL10 gold) were then transformed for nick repair. Three plates were made: one with the plasmid (different aliquots plated to increase the probability of having single colonies) and both a negative and a positive control. Analysis of plated bacteria growth is consistent with expectation. Plasmid plate has well separated colonies, in a more dense fashion in the left, where 100µL were plated. The control plate without ampicillin contains a homogeneous and unresolved growth, while on the other hand the control with antibiotic (but without plasmid conferring resistance to it) has no sign of bacterial growth on it, as depicted in the scheme below:

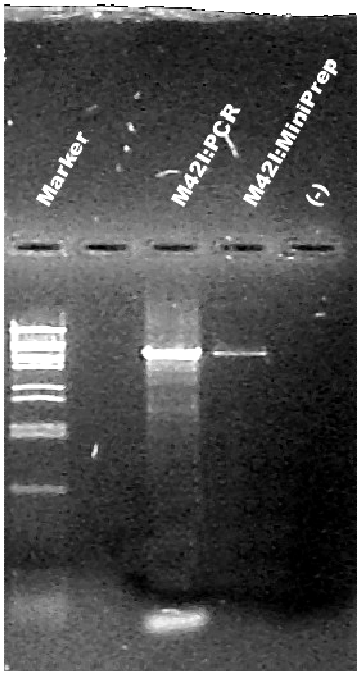
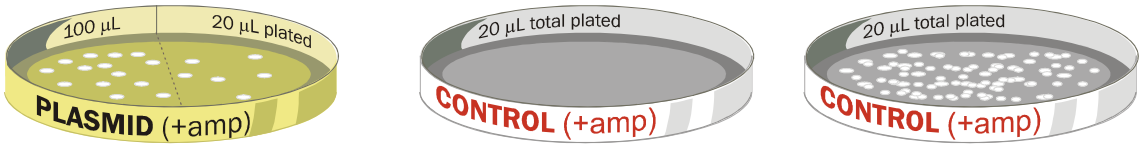


Fig. 4.1.1. – Agarose gel electrophoresis of PCR products (second lane) and miniprep (third lane).



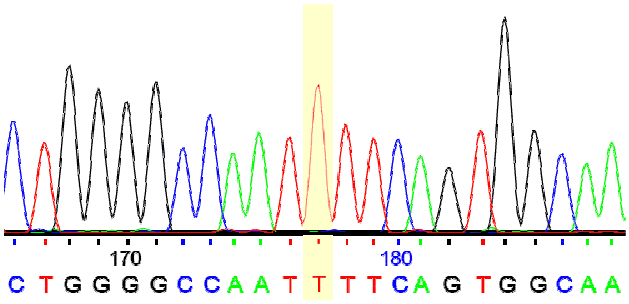
To ensure that mutagenesis had the expected effect on the plasmid, this was extracted to be sequenced. The mini-scale preparation was performed with a MiniPrep kit by Sigma-Aldrich. Agarose electrophoresis shows a neat band, cleaned from the smear present in PCR products.

Quantification was performed comparing intensity of bands with markers ones (λBst EII marker). Intensity of MiniPrep band is comparable to 2323 bp band, that results in an estimation of 12 ng/µL.

Although *Webrepository* (§1.3) contains all results presented in this section, most relevant results are also present at the end of this report.

DNA sequencing

Extracted DNA was evaporated to send the e-tube to MWG Biotech for sequencing from both extremities (using universal primers fro pGEX-2T). Sequence was assembled by alignment with Needleman and Wunch algorithm, and the final sequence was aligned with Wild Type *Ce*TK1 sequence.



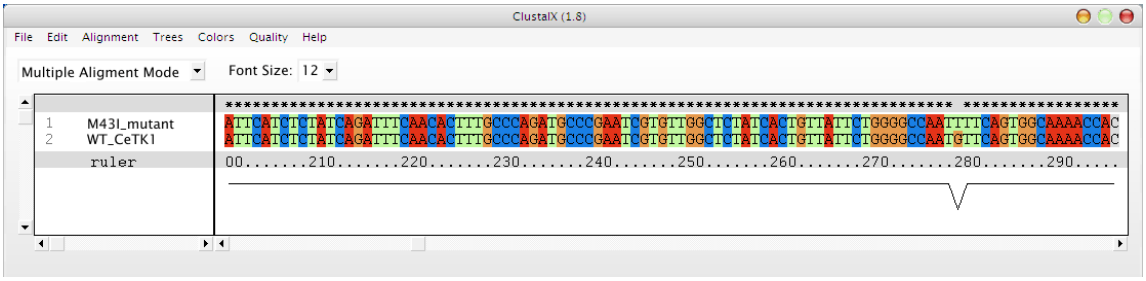


Fig. 4.1.2. – (*left*) Portion of the complete electropherogram provided by MWG Biotech with the sequence of the plasmid pGEX-2T-Ce-TK1-M42I. The mutated base is highlighted in yellow. (*bottom*) ClustalX was used to compare mutated sequence with the wild type one. Alignment confirmed that only the desired base was changed.

Transforming *E. coli* BL21

Three controls of transformation and recombinant bacteria plates were prepared and gave the expected result:

1.	No plasmid, no Ampicillin (-AMP)	Visible layer of bacteria
2.	No plasmid, with ampicillin (+AMP)	No growth (not shown)
3.	Plasmid pGEX-2T, Ampicillin (+PLASMID)	No growth
4.	Mutant plasmid (CeTK1-M42I)	Single colonies.

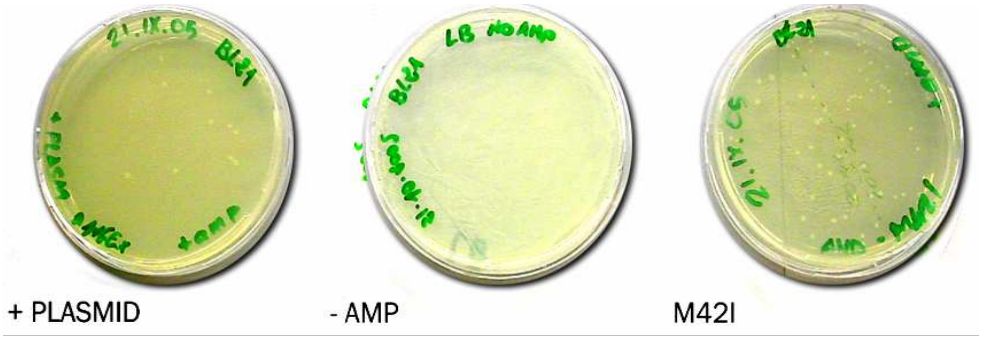


Fig. 4.1.3. – Plates with transformed *E. coli* BL21, the expression host of this project. Control with plasmid without insert plated on ampicillin medium and without are first two plates. Then plate with cells transformed with pGEX-2T-Ce-TK1-M42I is visible with well separated single colonies.

4.2.z Protein expression, purification and quantification

Time samples

Time samples were taken for 8 hours, and activity was continuously increasing, suggesting that this mutant seems not to heavily destabilize bacterial growth. The 14 hours of the experiment were not enough to see a peak, so estimation was done to harvest cells in the next induction. The plot on the right shows growth (OD) and activity normalized to the maximum value of each series.

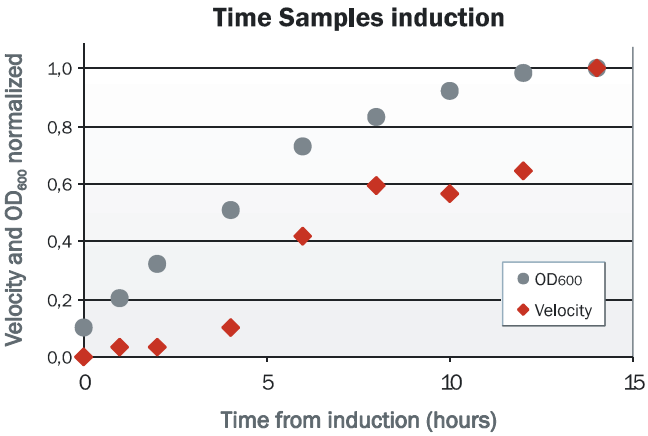


Fig. 4.2.1. – Plot of activity and OD_{600} from time samples. [CSD. see §1.2]

Induction was made in two flask: one to harvest cells from which extract and purify protein (after 16 hours), and one to take again time samples for a longer time (up to 24 hours). The choice to harvest cells after 16 hours reflect the caution needed not to overcome the peak, but rather harvesting cells before.

Time samples from induction

Up to 16 hours after induction OD_{600} were recorded for both harvest and time samples flasks, as a control. In the plot on the right results from the "time samples" flask are presented as in the same form as above, normalizing results to allow a handy comparison in the same scale. A peak is now clearly visible at 18 hours, thus the choice to harvest at 16 appeared adequate.

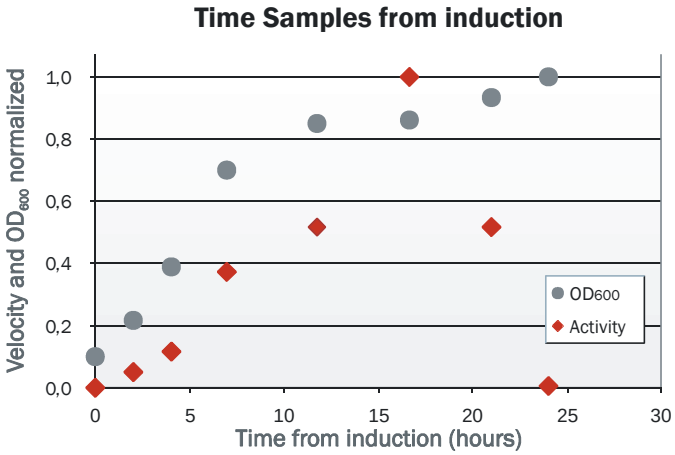


Fig. 4.2.2. – Plot of activity and OD_{600} from induction. [CSD, see §1.2]

Activity per litre of culture

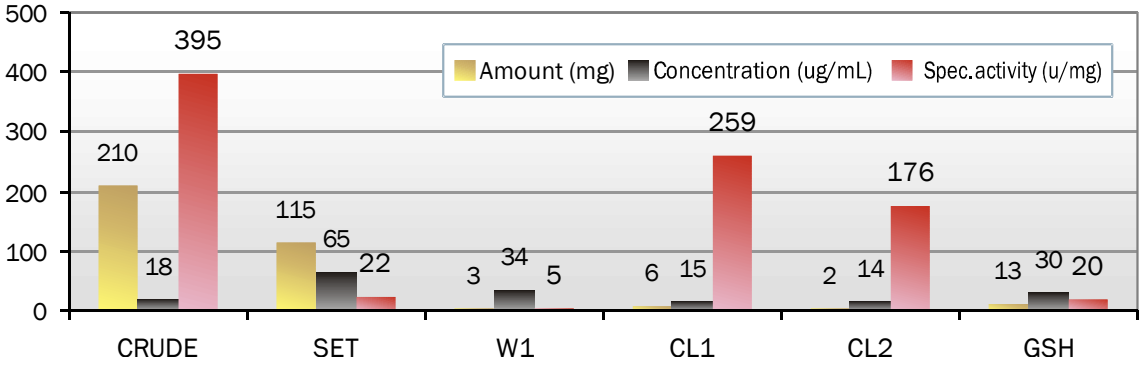
Time samples are made taking 2 mL of liquid culture, and resuspending the pellet in 0,4 mL of lysis buffer. The activity of sample time samples from the second induction was:

Sample	u/mL	u/mL of culture
Induction Time sample (16h)	153.034	382.584

Purification

All fractions[xv] were quantified with Bradford, analyzed with SDS-PAGE and finally activity was measured. Result of **Bradford quantification** is reported in the plot below, where for each fraction both total amount of proteins of the whole fraction (in mg) and concentration (µg/mL) are shown.

Purification plot for significant fractions



Purification Table

Purific. STEP	Volume (ml)	Activity (u/ml)	Total activity (U)	Total Protein (mg)	Spec. activity (u/mg)	Yield (%)	Purification factor
CBH	20,0	590,26	11805,17	209,6	56,3	100%	1,00
CL1	3,0	386,90	1160,69	5,8	199,3	10%	3,54
CL2	2,0	262,58	525,17	2,0	258,7	45%	4,59

Fig. 4.2.3. – Graph (*top*) showing amount and sconcentration of proteins and specific activity in different purification fractions. CRUDE is bacterial homogenate, SET is first elution, W1 the first wash, CL1-2 cleavage fractions. Purification table (*bottom*) showing yield and purification factor.

Purification table shows that yield of purification was poor, but enough to go on with characterization. An aliquot of each fraction was loaded on a **SDS-gel** to detect the degree of purification. Cleavage fraction shows a clear band situated between markers band of 30 and 20 kDa, that is where *Ce*-TK1 should be.

On both cleavage fractions a light band over the *Ce*-TK1 one is visible, meaning that purification was not absolute. The extra band could be some protein interacting with *Ce*-TK1, while it's difficult could be generically a molecular chaperone for the small weight of the same.

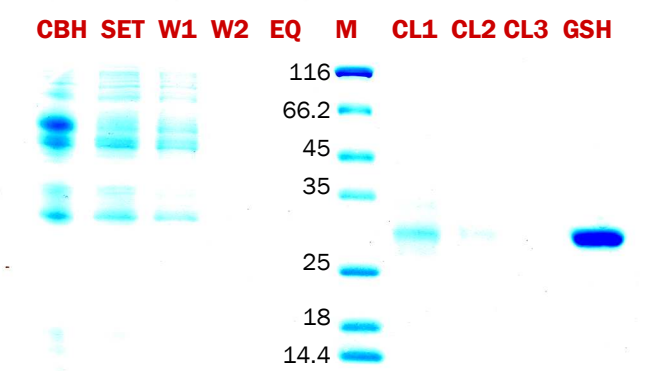


Fig. 4.2.4. – Digital acquisition of SDS-PAGE for M42I fractions. M is protein marker (MW in kDa is written near corresponding band).

4.3. Kinetic analysis with Thymidine

Purified CeTK1-M42I mutant was then subject of kinetics analysis as previously described. After choosing ideal assay conditions (i.e. enzyme dilution and amount of labelled substrate) an assay with different substrate concentration was performed as described.
Except for thymidine assay, only Michaelis-Menten plot is provided, for reasons anticipated in § 3.7., as Eadie-Hofstee plot were made (and it's available §1.3.) for first look only.

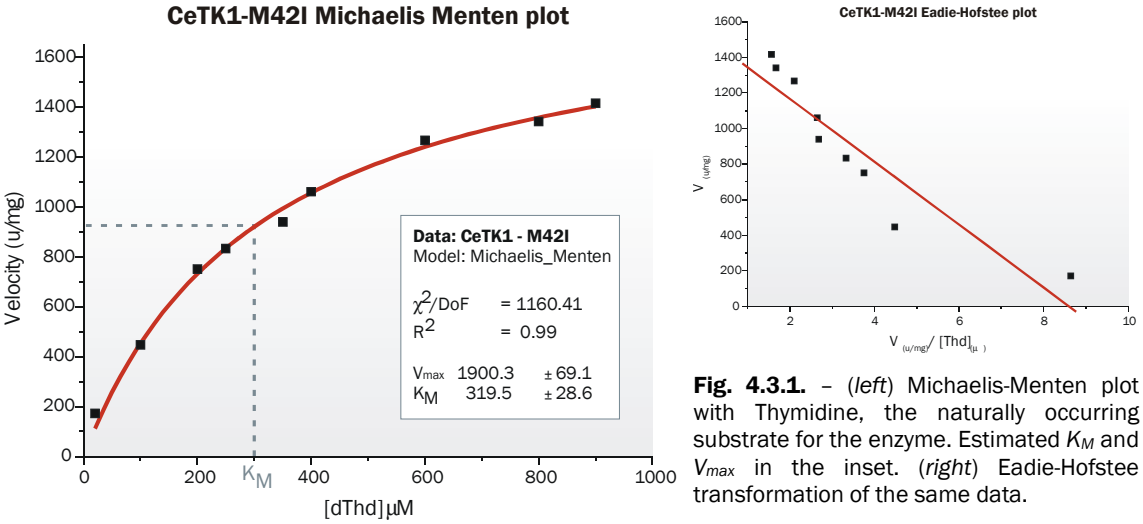


Fig. 4.3.1. – (*left*) Michaelis-Menten plot with Thymidine, the naturally occurring substrate for the enzyme. Estimated K_M and V_{max} in the inset. (*right*) Eadie-Hofstee transformation of the same data.

It's clear that experimental data well fit the Michaelis-Menten model (R^2=0.992). Estimated parameters for the Michaelis-Menten equation were:

V_{max}	1900.3± 69.1 u/mg	k_{cat}	0.94 ± 0.00 sec^{-1}
K_M	319.5 ± 28.6 µM	k_{cat}/K_M	2900 M^{-1} sec^{-1}

4.4. Substrate specificity

TK1-M42I activity was tested also with the other three deoxynuclesides, and two Thymidine analogs (AZT and AraC) having for each test a substrate concentration of 100 µM.

Substrate	Activity (u/mg)	Relative (dThd=100)
dAdo	0.089	0.03%
dCyd	0.157	0.04%
dThd	**350.000**	-
dGuo	0.206	0.06%
AZT	17.718	5.06%
AraC	0.032	0.01%

Only deoxycytidine showed a sufficient activity to further investigate its kinetics, while phosphorylation of purines was too little to detect activity with lower substrate concentrations. Activity with AZT was also relatively high, and thus further probed.

Kinetic analysis with dCyd

From curve fitting, kinetics parameters with dCyd as substrate were estimated as:

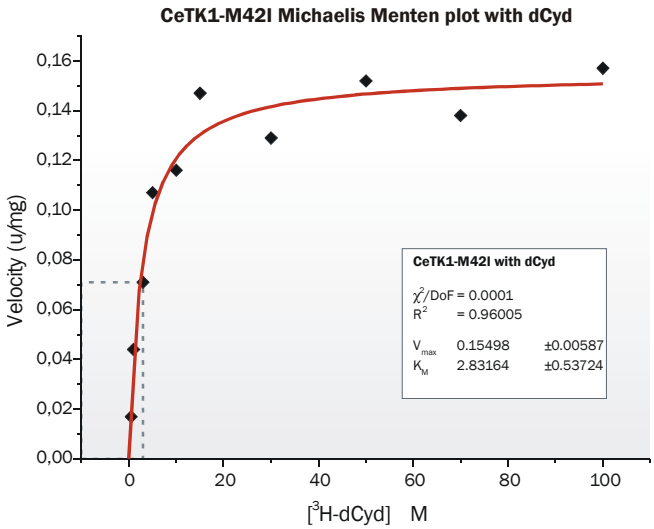
V_{max} 0.15 ± 0.005 u/mg

K_M 2.83 ± 0.53 µM

k_{cat} 7.4 x 10^{-5} ± 0.00 sec^{-1}

k_{cat}/K_M 26 M^{-1} sec^{-1}

The whole plote is shown on the right, with fitting results in the inset.



[CSD numbers in Y axis, see §1.2]

Kinetic analysis with AZT

AZT, analogue of Thymidine used in AIDS cure, is incorporated by the mutant with a surprisingly low efficiency (see discussion for comparison with wild type):

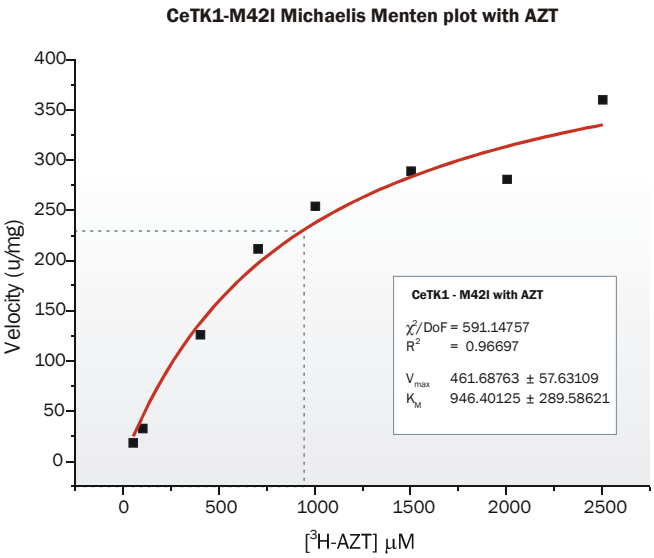
V_{max} 461.7 ± 57.63 u/mg

K_M 945.4 ± 289.5 µM

k_{cat} 0.23 ± 0.02 sec^{-1}

k_{cat}/K_M 200 M^{-1} sec^{-1}

The whole plote is shown on the right, with fitting results in the inset.



5. Discussion

5.1. Effect of substitution on activity

Data overview for wild type were presented in §2.4. Michaelis-Menten plot and non-linear curve fitting is reported below for *Ce*TK1-WT and for HuTK1-Δ40-M28I, that is the analogous mutation in Human enzyme.

Humant mutagenesis was made on the truncated form of TK1, because this resulted in increased stability for the Human, but not for the *C. elegans* Thimine Kinase 1. Data for Human mutant and *C. elegans* was kindly provided by Tine Skovgaard, and for this reason are reported here rather than in results.

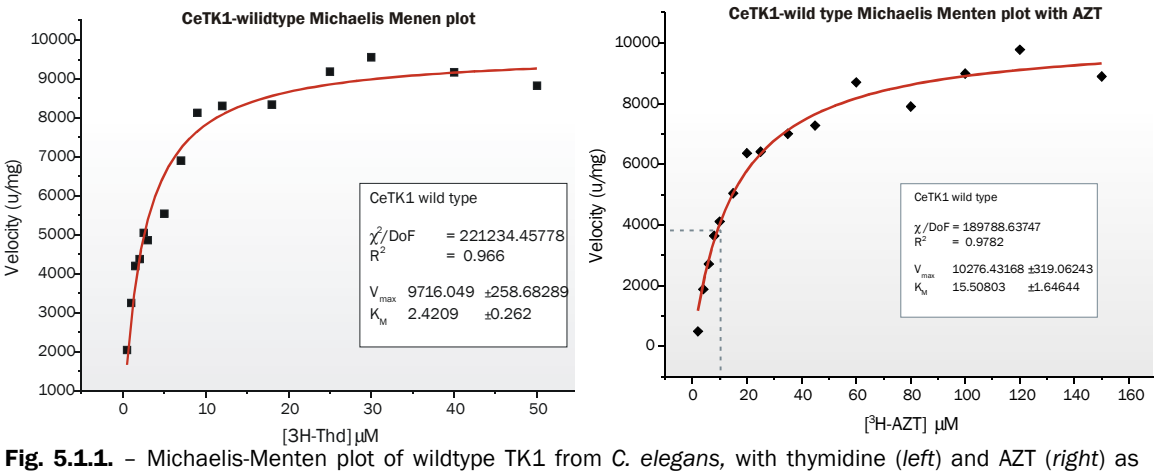


Fig. 5.1.1. – Michaelis-Menten plot of wildtype TK1 from *C. elegans,* with thymidine (*left*) and AZT (*right*) as substrate.

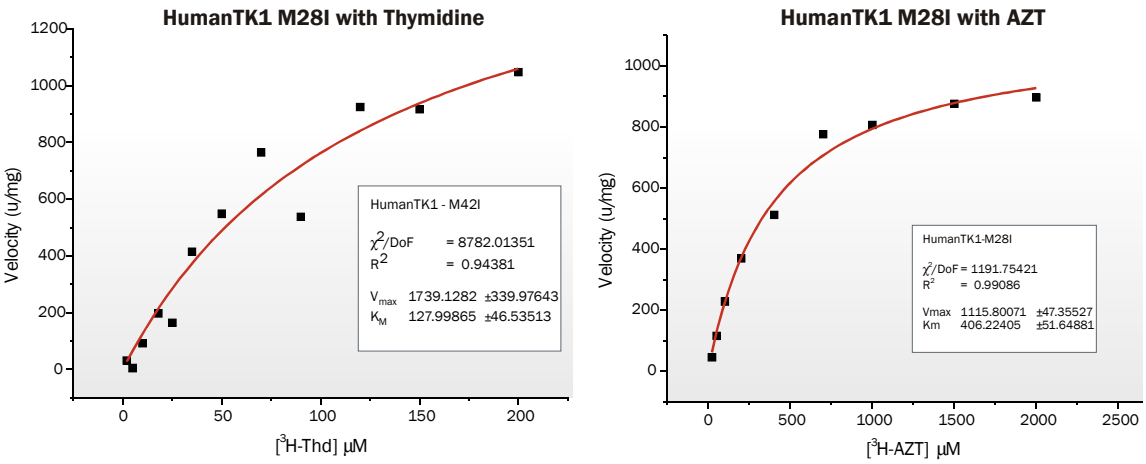


Fig. 5.1.2. – Michaelis-Menten plot of Human TK1-Δ40-M28I, with thymidine (*left*) and AZT (*right*) as substrate, shown for comparison.

From this plots a first comparison of the presented mutant with wild type is summarized in the following table. Table 5.1.1 shows kinetic parameters of the Michaelis-Menten equation, while Table 5.2.2. shows also k_{cat} and specificity constant (k_{cat}/K_M).

	K_M	V_{max}
HuTK1-WT [26]	Thd: 0.5 AZT: 2.2 (increased 4.4 X)[a]	Thd: 9455 AZT: 4957 (52%) [a]
HuTK1-Δ40-M28I	Thd: 128 ± 46 AZT: 408 ± 52 (3.2 X)	Thd: 1729 ± 339 AZT: 1115 ± 47 (64%)
CeTK1-WT	Thd: 2.42 ± 0.2 AZT: 15.51 ± 1.6 (6.2 X)	Thd: 9716 ± 258 AZT: 10276 ± 319 (105%)
CeTK1-M42I	Thd: 319.5 ± 28 AZT: 945.4 ± 289 (3 X)	Thd: 1900.3 ± 69.1 AZT: 461.7 ± 57.6 (24%)

Table 5.1.1. – Comparison of kinetic parameters of Michaelis-Menten equation for *C. elegans* and Human kinase and mutants in the same position[21, 27]. [a] Comparison of AZT vs Thd usage is made.

HuTK1	**WT** [27]	**CΔ40-M28I** [a]
k_{cat} (sec^{-1})	6.3 ± 0.9	0.61 ± 0.11
K_m (μM)	16 ± 3.4	128 ± 46
k_{cat}/K_M	0.4 x 10^6	0.005 x 10^6
CeTK1	**WT**[21]	**M42I**
k_{cat} (sec^{-1})	5.0 ± 0.1	0.94 ± 0.01
K_m (μM)	2.3 ± 0.3	319.6 ± 28.6
k_{cat}/K_M ($M^{-1}sec^{-1}$)	2.2 x 10^6	0.003 x 10^6

Table 5.1.2. – Comparison of kinetic parameters using Thd as substrate. [a] Data for Human truncated from Tine Skovgaard.

It's remarkable the high V_{max} of *C. elegans* wild type with AZT analogue, almost equal than with Thd. K_M, on the other hand, is definitely higher (six times). The mutant has a decreased activity, but even more evident with AZT (with a twice as high a K_M and only 25% V_{max} compared with thymidine catalysis).

The disruption of catalytic efficiency was predictable substituting a so highly conserved amino acid. But the mutant seems also less tolerant with the substitution of nitrogen in 3' position than wild type.

Comparing both mutants with their corresponding wild type, show a really similar alteration of catalytic efficiency. Values of V_{max} themselves are similar, and both mutants have a 20% decrease of maximum velocity. Affinity for substrate decrease is more dramatic, as K_M increases 130 times for *C. elegans* mutant, and 250 times for human mutant, suggesting a heavy involvement of this amino acid in ligand binding, rather than in catalysis.

5.2. Effect of substitution on substrate specificity

Preliminary data[xvi], because obtained from a single point assay, was obtained screening mutant's ability of phosphorylating different substrates.

For each mutant a pie plot were made with relative activity, to compare the relative[xvii] preference with different substrates. Two different type of pie charts are presented: first a simple chart with percentage referring to relative activity-modified pie plot using logarithm of such values is also presented, to make a visual comparison easier. Both mutants with this visualization appear to be more tolerant with other substrates than wild type. CeTK1-M42I, in particular, seems to be the only one capable of AraC phosphorylation; in relative terms, its activity with dGuo is also relevant.

With dCyd the activity is really low (see Michaelis-Menten plot, §4.4), but with a low K_M. This is another clue for Met[42] involvement in ligand binding.

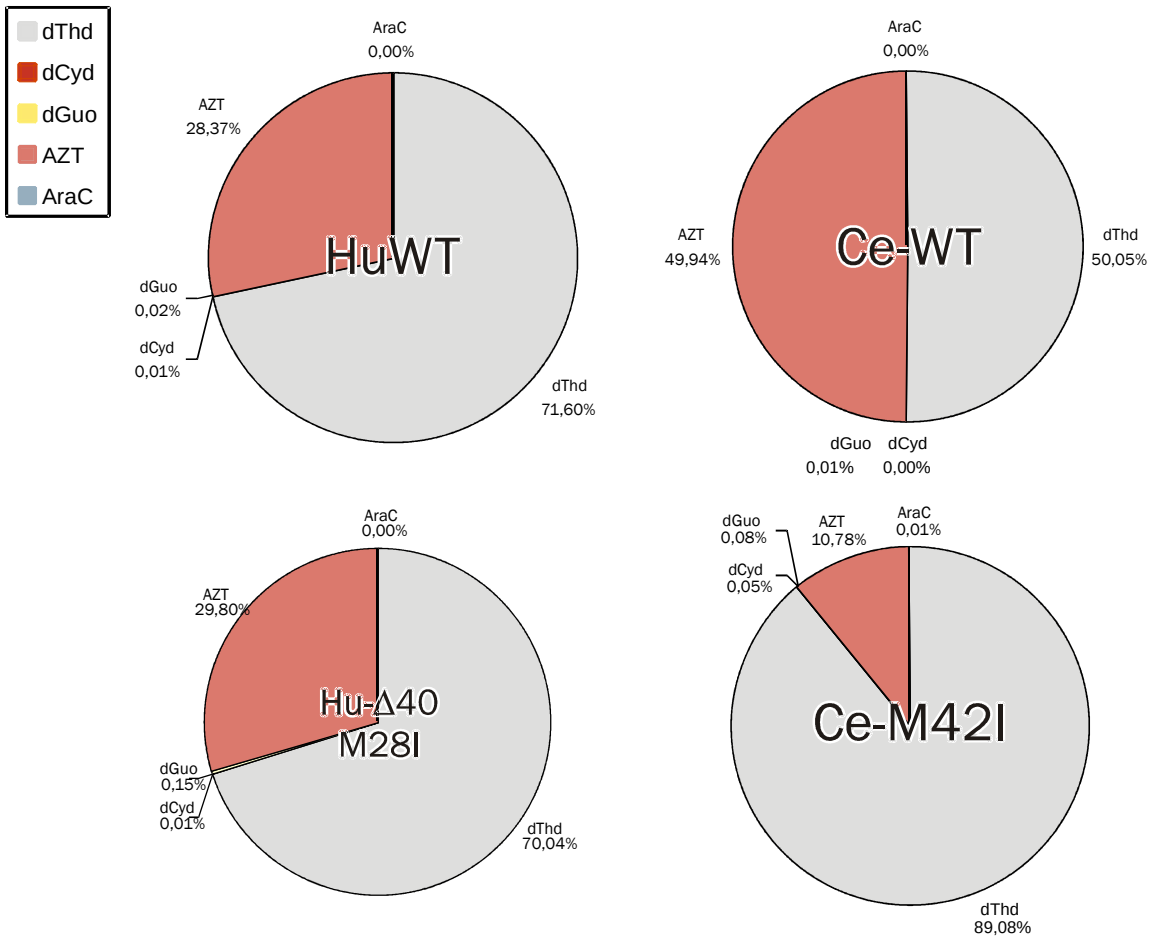


Fig. 5.2.1. – Pie plots relative usage of different substrates for each considered mutant (unmodified scale to compare real values). [CSD numbers, see §1.2]

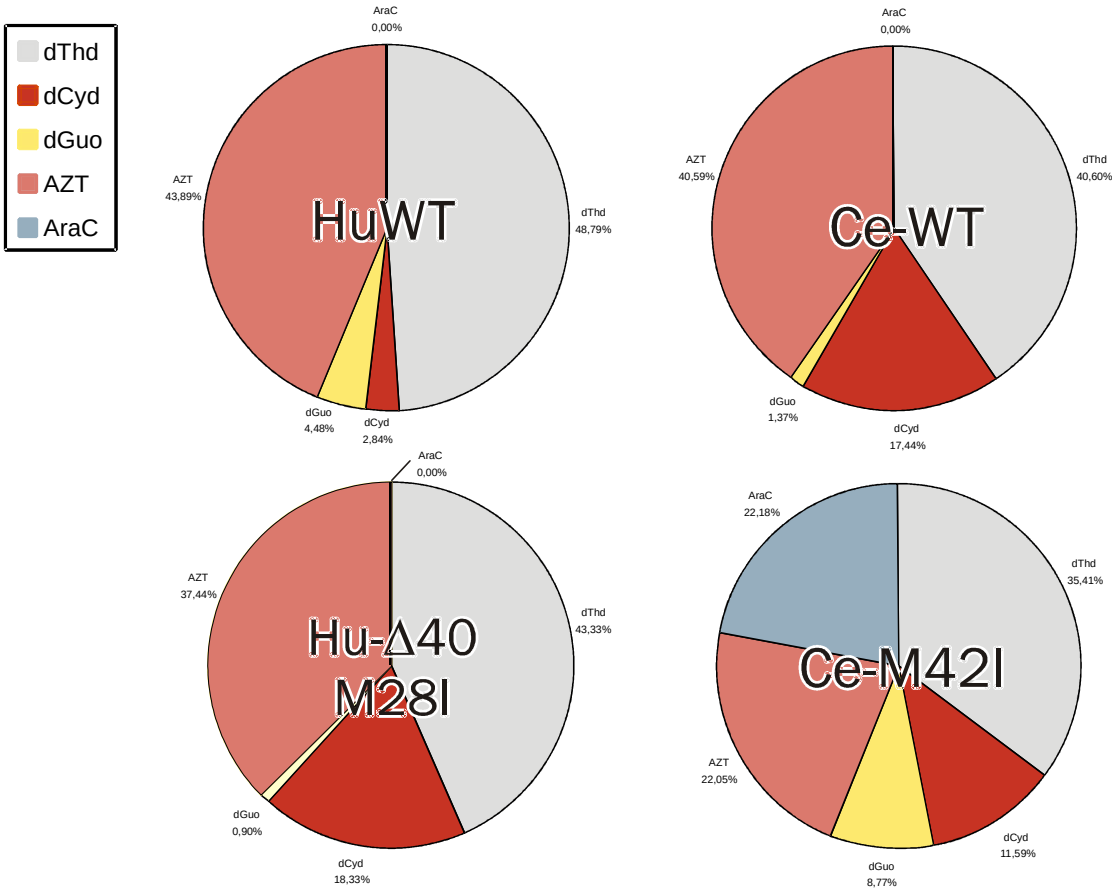


Fig. 5.2.2. – Pie plots relative usage of different substrates for each considered mutant. In this plot a logarithmic scale is used to enlarge slice referred to poorly used substrates. [CSD numbers, see §1.2]

The most remarkable aspect raising from these data is the relative activity of CeTK1-M42I with AraC analog. Not only in relarive terms (see plots), but also in absolute terms (data not shown), this mutant seems to have the highest activity.

6. Appendices

6.1. Material and Methods

Media

Bacteria were grown in LB medium, composed as follows: 10 g Bacto™ tryptone; 5 g Bacto™ Yeast Extract; 10 g NaCl; ddH_2O up to 1 litre. When needed Ampicillin or IPTG were added to reach final concentration of 100 µM.

Agarose gel electrophoresis

1×TBE-buffer: 108 g Tris-base, 55 g Boric acid, 7.5 g EDTA, add H_2O to a volume of 2000 ml and autoclave, then dilute 5 times with H_2O.

1% Agarose gel: 50 ml 1× TBE, 0.5 g agarose,

Agarose is dissolved by heating in a microwave oven for ~2 min. and is then cooled to ~60°C. 5 µl 4 mg/ml Ethidium bromide; pour in a form and leave to set for app. half an hour.

Running buffer: 750 ml 1× TBE, 90 µl 4 mg/ml Ethidium bromide

Loading buffer: 50 mM Tris-HCl (pH 8.0), 25 mM EDTA, 20% sucrose, add H2O to a volume of 100 ml. Add bromo phenole blue until a dark blue colour is reached.

MW Markers: *1 Kb Plus DNA Ladder* from Invitrogen: Provides fragments from 100 up to12.000 bp. *λ-BstE-II-marker*: Fragments from 117 to 14 kbp.

SDS-PAGE

Gels (12% SDS-polyacrylamide) were supplied by Bio-Rad (BIO-RAD) was used. Samples were prepared as 5 µl protein dilution; 1 µl 6× SDS loading buffer, heated at 95°C for 2 min. Gel run 120 V for 1 hour. Staining O/N with Coomassie Brilliant Blue solution.

SDS-PAGE loading buffer 6×: 0,35 M Tris-HCl; pH 6,8 at 22°C; 10,28% w/v SDS; 36% v/v glycerol; 5 % v/v β-mercaptoethanol; 0,012% w/v bromophenol blue; ddH_2O.

Coomassie Brilliant Blue Stain: 1g CBBG-250; 450 ml CH_3OH 96%; reach 500 mL ddH_2O.

Miniscale preparation of DNA

A single isolated colony from the ON plate was put in 5 mL LB medium + 5 µl / Amp 100 mg/ml and grown over night. The day after, a miniprep of plasmid was performed, with the GenElute™ Plasmid Miniprep Kit (supplied by Sigma-Aldrich®). The purified plasmid was stored at -20°C.

Chemical competent cells

Competency buffer I: 100 ml ddH2O; 20 ml 1 M KCl; 1.2 ml 5 M Potassium acetate; 12 ml 1 M $CaCl_2$; 30 ml ultrapure glycerol;

Mix to homogeneity and titrate to pH 5.8 with 0.2 M acetic acid. Bring to a final volume of 200 ml with ddH2O and filter sterilize. Store at 4 °C.

Competency buffer II: 100 ml ddH2O; 4 ml 0.5 M MOPS; 2 ml 1M KCl; 15 ml 1 M $CaCl_2$; 30 ml ultrapure glycerol

Mix to homogeneity and titrate to pH 6.8 with 10 N NaOH. Bring to a final volume of 200 ml with ddH2O and filter sterilize. Store at 4 °C.

An *E. coli* XL10-Gold glycerol culture is inoculated to 5 ml LB-medium and incubated in a 37°C water bath with shaking over night. The over night (ON) culture is diluted 100 times by transferring 1 ml ON culture to 99 ml LB-medium and the new culture is left to grow at 37°C until OD_{600} = 0.3-0.5. The flask is placed in an ice/water bath for 15 min. and the culture is transferred to pre cooled centrifuge glasses and spun down in a Sorvall centrifuge at 3000 rpm and 4°C for 15 min. The supernatant is discarded and the pellet is suspended in 1/3 culture volume (~33 ml) ice cold competency buffer I. After 1 hour of incubation on ice the cells are centrifuged again at 3000 rpm and 4°C for 15 min. and the supernatant is discarded. The cells are resuspended in 1/25 culture volume (~4 ml) competency buffer II and incubated on ice for 15 min. The cells are now competent and can be stored in small portions at -80°C.

Protein engineering

Plasmid pGEX-2T containing the ORF of *Ce*-TK1 was used to obtain the point mutation M42I using Quick-Change™ kit by Stratagene®, with minor changes, as the adoption of *Pfu Ultra High Fidelity DNA polymerase*. Primer used: 5'-**GGG GCC AAT TTT CAG TGG CAA AAC CAC CG**-3' and its reverse complementary, as required by the protocol.

PCR was performed with the following PCR mix: 5 µl *Pfu* buffer (stock solution: 10×); 1 µl template DNA (pGEX-2T-Ce-TK1 plasmid, ~60 ng/µl); 2,5 µl primer forward (stock solution: 50 ng/µl); 2,5 µl primer reverse (stock solution: 50 ng/µl); 1 µl dNTPs MIX (stock solution: 10 mM); 1 µl *Pfu Ultra High Fidelity* (stock solution: 2,5 U/ml); 37,0 µl sterile and deionised water, thermocyclated with this program: 95°C x 30"; 95°C x 30"; 16x[55°C x 1'; 68°C x 5'30"; 68°C x 5'30"]; 4°C x ∞.

The PCR products were treated with restriction enzyme *DpnI:* 1 µl DpnI 10 U/µl added and incubated at 37°C for 1 hour. Sequence was verified by DNA sequencing, performed by MWG-Biotech. Electropherogram analyzed with Chromas2, alignment with ClustalW.

Cell lysis

Cell were lysed with sonication for time samples and French press for induction colture. I

Sonication - Time-samples pellets, stored at -20°C were resuspended in 400 µl of lysis buffer and sonicated 3×5 seconds, with 1 min pause. After the sonication, the samples were centrifugated 10 minutes at 20.000g at 4°C, and the supernatant stored at -80°C.

French press - The bacterial pellet collected after the induction was added with 1/20 volume of the original centrifuged culture volume. The bacteria were passed 3 times in French Press and centrifuged for 15 minutes at 20.000g (GSA rotor), at 4°C. The supernatant filtered through 1 µm and 0,45 µm filters.

Lysis buffer: 1× PBS; 5 mM EDTA; 5 mM DTT (dithiothreitol); 10% glycerol; 0,1% Triton X-100; 0,1 mM PMSF (phenyl-methil-sulfonyl fluoride); 5 mM benzamidine; (only sonication: 150 µg/ml lysozyme); 50 mM ε-aminocapronic acid; ddH_2O.

Protein Expression

The pGEX-2T *Ce*-TK1 M42I plasmid was trasformed into *E. coli* BL21 competent cells, and selected using ampicillin resistance encoded by the plasmid. Induction was performed adding IPTG (100 uM) for 16 hours at 25°C in shaking bath, and was started after reaching OD_{600} 0.6.

Cells were harvested by centrifugation in a Sorvall centrifuge (GSA rotor) at 4000 rpms for 15 minutes. Pellet stored at -80°C, then dissolved in lysis buffer (1x PBS, 5 mM EDTA, 10% glycerol, 0.1% Triton X-100, 0.1 mM PMSF, 5 mM benamidine). Lysys performed with French press.

Protein Purification

Purification was performed in a Sephadex GSH column with following buffers.

Buffer A: 1× PBS; 87% glycerol; 0,1% Triton X-100; 1 mM DTT; 0,1 mM PMSF; 5 mM Benzamidine; 50 mM ε-aminocapronic acid; 5 mM $MgCl_2$; ddH_2O;

Thrombin buffer [+ Thrombin]: Triton X-100 0,1%; PBS 1×; [Thrombin 0,05 U/ml;] ddH_2O;

Glycerol buffer: 30% glycerol; 15 mM DTT; 15 mM $MgCl_2$; 30% Triton X-100; ddH_2O;

Reduced GSH: buffer: 50 mM Tris-HCl pH 8,0; 10 mM reduced GSH; in ddH_2O

Bradford Quantification

Assay: 100 µl of the protein-sample; 900 µl of Bradford reagent. OD_{595} measured after 5 min incubation.

Bradford reagent: 100 µg/ml Coomassie Brilliant Blue G-250; 4,75%ethanol; 8,5% H_3PO_4; ddH_2O;

Kinetic/activity assay

Initial velocity was determined with three time samples (5, 10, 15 minutes) by the DE-81 filter paper assay, using tritiated substrates. Standard assay conditions: 50 mM Tris-HCl pH 7.5, 2.5 mM $MgCl_2$, 10 mM DTT, 0.5 mM CHAPS, 0.5 mg/mL BSA, 2.5 mM ATP. labeled/unlabeled substrate ratio was optimized in each experiment, and is reported in the data available in the Web Repository (§ 1.3.).

In kinetics assay the reaction mix was: 25 µL of Tritium-labeled substrate + 25 µL of reation mixture with the enzyme. 10/13 µL of the mixture, incubated at 37°C, were spotted on filters depending of the diameter. 5 µL were spotted on standards, which filters were not washed.

The reaction mix contained: 5 µl MIX ATP/$MgCl_2$ (stock solution: 10×); 2,5 µL BSA (stock solution: 60 mg/ml); 1 µL CHAPS (stock solution: 25 mM); 1 µl NaF (stock solution: 0,15 M); 1 µL ^{3}H-dThd (added only in activity assay); 29,5 µl ddH_2O; 10 µL bacterial crude extract.

ATP/$MgCl_2$ 10X MIX: 0,5 M Tris-HCl; 0,1 M DTT; 0,025 M; 0,025 M ATP $MgCl_2$; ddH_2O.

Enzyme buffer: 50mM Tris pH 7.5; 1mM CHAPS; 3 mg/mL BSA; 5mM DTT; ddH_2O.

Wash of the DE81 filters was made in AMF (10 mL calammonuim formiate 2,5 M in 5 L ddH_2O); the eluition was performed with dTMP Eluition Buffer-TES (KCl 0,2 M in HCl 0,1 M), then filters were left on agitator for 30'.

Fluor: 2,5 mL Eco-Scint scintillator liquid added, mixed, and put in the scintillator counter.

6.2. References

A global review to this broad field of biochemistry, cell biology and cancer medicine were found in Dvora Berenstein PhD thesis (9) and a book on the topic by Prof. Birgitte Munch-Petersen (10). More specific details on *C. elegans* TK1 were obtained in (20). General review of topics related with this project were found in Encyclopedia of Life Science (www.els.net) by Nature Publishing Group (6, 7, 8).

General References

1. Nelson and Cox; *Lehninger Principles of Biochemistry*, 4nd edition; 2004, W. H. Freeman
2. Sambrook, J., Fritsch, E.F., and Maniatis, T.*; Molecular Cloning: A Laboratory Manual,* 2nd edition; 1989, CSHL Press
3. Watson *et al.*; *Molecular Biology of the gene*, 5th edition; Pearson & Benjamin/Cummings, 2004
4. Athel Cornish-Bowden P.; *Foundamentals of Enzyme kinetics* 2nd edition; 1995, Portland Press
5. Lodish H. *et al.*; *Molecular Cell Biology*; 2004, W. H. Freeman Ed.
6. Dunn M. B.; Engineered Enzymes; 2001, Encyclopedia of Life Sciences, Nature Publishing Group (NPG)
7. El-Gewely F., Fenton C., Salomonsen J., Xu H.; *Mutagenesis: Site-specific*; 2001, Encyclopedia of Life Sciences, NPG
8. Ioannou P. A. , *Human Gene Therapy*; 2001, Encyclopedia of Life Sciences, NPG
9. Munch-Petersen B.; *DNA Precursor Pool Balance and Thymidine Kinase Isoenzymes…*; 1996, Roskilde University Press
10. Berenstein D.; *Structure-function studies of human cytosolic thymidine kinase*; 2001, PhD Thesis [Roskilde University]

Articles & Reviews

11. Evans D. R., Guy H. I.; *Mammalian Pyrimidine Biosynthesis: Fresh Insights into an Ancient Pathway*; 2004, JBC
12. Blumenthal, T. *et al.*; *A global analysis of Caenorhabditis elegans operons*; 2002, Nature 417, 851–854
13. WormBase web site; *http://www.wormbase.org; release WS149, date 10-10-2005.*
14. Timmons L., Fire A.; *Specific interference by ingested dsRNA*; 1998, Nature 397, 851
15. Welin *et. al.*; *Structures of thymidine kinase 1 of human and mycoplasmic origin*; 2004, PNAS n. 52, vol. 101, 17971
16. Wild K. *et al.*; *The structure of thymidine kinase from Hepes Simplex Virus type 1 …;* 1997, Protein sci. 6, 39125-9
17. Reichard. P.; *Interactions between deoxyribonucleotide and DNA synthesis*; 1988, Ann. Rev. Biochem, 57:349-7
18. Reichard. P.; *The Anaerobic Ribonucleotide Reductase from Escherichia coli*; 1993, JBC n. 12, 8383-8386
19. Song S., Wheeler J., Mathews C. K.; *Deoxyribonucleotide Pool Imbalance Stimulates…*; 2003, JBC n. 7, 43893-6
20. Elion G.B.; *The Purine Pathway to Chemoterapy*; 1988, Nobel Lecture
21. Skovgaard T.; *Purification and Characterization of Wild-Type and Mutant TK1 Type Kinases from C. elegans*; in press
22. Bianchi V. and Spychala J.; *Mammalian 5'-nucleotidases*; 2003, JBC n. 47, 46195-8
23. Michaelis L. and Menten, M. L.; *Die Kinetik der Invertinwerkung*; 1913, *Biochemische Zeitschrift*, 49:333.
24. Zhu C. *et al.*; *Effect of C-terminal of human cytosolic thymidine kinase (TK1) on in vitro stability…*; . In Press, this issue.
25. Fasman G. D. (Ed.); *Prediction of Protein Structure and the Principles of Protein Conformation*; 1989, Springer
26. Bradford M.*; A Rapid and Sensitive Method for the Quantitation of Microgram…*"; 1976, Anal. Biochem. 72:248-254.
27. Munch-Petersen. B *et al.*; *Diverging Substrate Specificity of Pure Human Thymidine Kinases…*; 1991, JBC n.14 9032-8

Images

Images were made in CorelDraw. Photo of the liquid scintillation counter and Bradford samples were taken in Roskilde Universitetscenter's laboratories. Photo of *C. elegans* and French press were found on the Internet using "Google Images".

7. Acknowledges

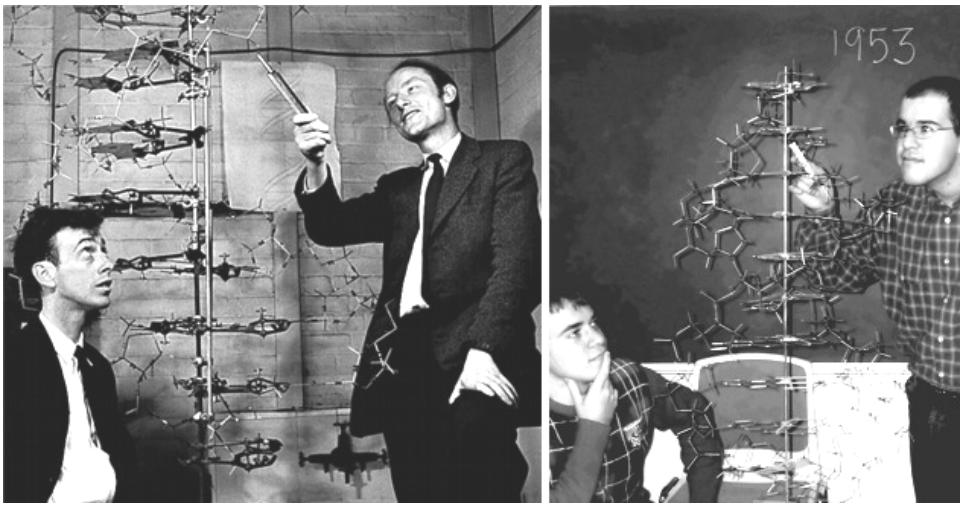
Denmark, Roskilde and RUC itself were very welcome and comfortable places to stay within. My very first approach to the lab life could't be in a better place, and the best picture of this feeling were the pleasant yet useful lab meetings (and lab dinners). *Cooperativity* is probably one of the main ingredients of this laboratory, probably this lesson came from enzymes showing us new properties emerging from cooperation between polypeptides ☺

Tine was my tutor in the white lab, with her clever and nice care of my working, always available to help me with my problems, while the experienced Dvora helped me in assimilataing what I where doing; they were an invaluable resource. I also have to thank Tine to kindly giving to me her data for the Human mutant (TK1-Δ40-M28I) and for *C. elegans* wild type TK1.

I'm also deeply indebted with Marianne, whose "Hello" in the morning I was everyday waiting for, and whose words were helpful in each subsequent moment from that greeting! Marianne is a really friendly and well organized "lab-keeper".

Last but not least, my thanks is for Prof. Munch-Petersen that, with her radiant character, was always ready to answer our doubts, and listen to our requests. Learning from you was a real joy.

Lab mates were a nice team, with Gabriella solar presence, little Maria and Nicolai; they all contributed in creating that good atmosphere. I was also lucky in coming to this experience abroad with Stefano, whose previous lab experience has been so useful for me at the beginning. We share the same passion and enjoyment for scientific research, good food and wine.



"It has not escaped our notice" that a good team achieve better results with less effort, expecially when passion is merged with the ability of having fun with our own work.

Footnotes

[i] For instance, Michaelis and Menten introduced the practice of controlling pH of reaction; pH was a new concept, introduced by Sørensen in 1909.

[ii] It's easy to understand that V_{max} is directly proportional to the amount of enzyme in the reaction mix. Get rid of the quantity "[E]" is necessary to compare different experiments, made – for example – with enzyme from different purifications and so with a different concentration. The proportionality constant, called k_{cat}, is also referred as the "turnover number". Because the maximum rate is obtained athigh substrate concentrations, when all the active sites are thus occupied, the "turnover number" measures the velocity at wich the enzyme operates once the active site is occupied.

[iii] The significance of specificity constant becomes clear rearranging Michaelis-Menten equation as follows: $v \approx (V_{max} / K_M) \cdot [S]$. This approssimation is done assuming that $[S] << K_M$, that, as said, is true under physiological conditions. The specificity constant is so a measure of the efficiency at low substrate concentrations.

[iv] HuTK1, after incubation in ATP containing buffers, forms homotetramers, while the *C. elegans* doesn't.

[v] Homology modeling of the target (CeTK1), performed via secondary structure alignment (real vs Jpred prediction) and consensus based methods with fully automated servers, Human TK1 was confirmed as a good template, even if some difference in secondary structure seems possible. Some server doesn't suggest Human TK1 as possible model, but it's possible that being this structure not yet catalogued in SCOP and CATH, fold recognition based on these servers could be limited.
It's important to notice that the Human enzyme was recognized as related to the RecA family (ATP binding domain present in the α/β domain), and the DALI server (Sawaya, 1999) suggests as most related structure CobA, involved in transfers the adenosyl moiety from ATP molecule.

[vi] Formula: Score = 4 x (number of clashes with backbone N, CA, and C atoms) + 3 x (number of clashes with backbone O atoms) + 2 x (number of clashes with side chains atoms) – number of H bonds – 4 x (number of SS bonds). Best rotamers are those with lowes score.

[vii] Zoran Gojkovic provided the gene, that was subcloned in pGEX-2T and engeneered to restore original worm sequence by Tine Skovgaard in Birgitte Munch-Petersen Laboratory, resulting in the pGEX-2T-Ce-TK1 plasmid used in this project.

[viii] This is a variation on original protocol, as this polymerase was found (T. Skovgaard, personal communication) to be more efficient.

[ix] For this reason the parental plasmid has to be propagated in a strain non methylase deficient.

[x] The assumption that an heavier isotope doesn't change kinetic behavior is essential, and it's true providing that the labeled atom is not involved in bond to be broken [4].

[xi] For one Michaelis-Menten analysis (performed with ^{3}H-AZT analog) a slightly different experiment design was adopted, because of the high concentration needed to assay and the elevate cost of the analog itself. This consisted in using different concentrations of unlabelled AZT into which adding a constant amount of labeled AZT. Even if this changed required a new Excel® template for data analysis (availability §1.3), the substrantial theory remains obviously the same.

[xii] This paragraph briefly explains the most important calculation in result analysis, that is how to convert CPMs recorded by the scintillation counter to amount of substrate spotted in the filter. The template I made for this is however selfexplicative and available at the WebRepository (see §1.3.).

[xiii] Migration is proportional to the mass, as effect of electric field, and then there is the effect of the shape, but dealing within denaturing condition for SDS gels, each molecule is linear shaped, so the main difference is the length. Approximately there is a logarithmic relationship between lenghts of molecule and migration.

[xiv] In this project only Eadie-Hofstee linearization has been used, but estimation of kinetics parameters relied on more statistically robust non linear curve fitting.

[xv] Fractions are labelled as: CBH or CRUDE (crude bacterial homogenate), SET (elution), W1-W5 (eluted washing column with Buffer A), EQ (equilibration), CL1-CL3 (cleavage fractions) and GSH (final wash with glutathione).

[xvi] One point assay, with substrate concentration of 100μM for each substrate. This sort of "multi-substrate assay" is useful to plan further experiments, but data so obtained are not reliable.

[xvii] Each value was calculated referring to the sum of activities with all substrates, rather than referring to the naturally occurring one (Thymidine).