Biochemical Characterization of Human Guanylate Kinase and Mitochondrial Thymidine Kinase: Essential Enzymes for the Metabolic Activation of Nucleoside Analog Prodrugs

Dissertation

for the award of the degree

"Doctor of Philosophy" (Ph.D.)

Division of Mathematics and Natural Sciences

of the Georg-August-Universität Göttingen

submitted by

Nazimuddin

from

Charsadda, Pakistan

Göttingen 2014

Member of the Thesis Committee: **Dr. Manfred Konrad** (Reviewer)
Enzyme Biochemistry Research Group, Max Planck Institute for Biophysical Chemistry

Member of the Thesis Committee: **Dr. Kai Tittmann** (Reviewer)
Department of Bioanalytics, Georg-August-Universität Göttingen

Member of the Thesis Committee: **Dr. Matthias Dobbelstein**
Department of Molecular Oncology, Georg-August-Universität Göttingen

Member of the Examination Committee: **Dr. Thomas P. Burg**
Biological Micro- and Nanotechnology, Max Planck Institute for Biophysical Chemistry

Member of the Examination Committee: **Dr. Jochen Hub**
Department of Molecular Structural Biology, Georg-August-Universität Göttingen

Member of the Examination Committee: **Dr. Henning Urlaub**
Bioanalytical Mass Spectrometry Group, Max Planck Institute for Biophysical Chemistry

Date of the oral Examination: February 5, 2015

Affidavit

I hereby declare that the Thesis "Biochemical Characterization of Human Guanylate Kinase and Mitochondrial Thymidine Kinase: Essential Enzymes for the Metabolic Activation of Nucleoside Analog Prodrugs" has been written independently and with no other sources and aids than quoted.

....

Nazimuddin

Göttingen, November 29, 2014

Dedication

I dedicate this work to my mother "**Fazilat Khan**" who inspired me the most in my life because of her great kindness, love, wisdom, and elegance.

Acknowledgments

I would like to express my deepest gratitude to **Dr. Manfred Konrad** for offering me the opportunity to do my Ph.D. research work in the Enzyme Biochemistry Research Group. It was indeed an honor and great experience to work under his supervision, and in the very dynamic scientific environment of MPI-bpc, Goettingen. I am highly grateful for his guidance and great support in starting several collaborative projects. I am highly thankful to him for reviewing my PhD thesis, and for his very constructive critique and useful comments.

My profound gratitude to the Ph.D. thesis committee members, **Prof. Dr. Kai Tittmann** and **Prof. Dr. Matthias Dobbelstein**, for reviewing my research work, constructive critique and for their useful feedback. I am highly thankful to all other members of the Ph.D. Examination Board **Prof. Dr. Henning Urlaub**, **Dr. Thomas P. Burg**, and **Dr. Jochen Hub** for their time and constructive comments.

I am highly grateful to our research collaborators **Prof. Dr. Wolfgang Parak**, **Nadeem Sabir** and **Dr. Susana Carregal** from Philipps University of Marburg, **Prof. Simone Techert** and **Rohit Jain** from Structural Dynamics of (Bio)chemical Systems group at MPI-bpc Goettingen, **Dr. Donghan Lee** and **Dr. Thomas Michael Sabo** from NMR-based Structural Biology at MPI-bpc Goettingen, **Dr. Vladimir Pena** and **Dr. Tales Rocha de Moura** from Macromolecular Crystallography research group at MPI-bpc Goettingen, **Dr. Andre Skirtach**, **Dr. Alexey Yashchenok** from the department of interfaces at Max-Planck Institute of Colloids and Interfaces Golm/Potsdam, and **Dr. Bogdan Parakhonskiy** from University of Trento Italy, for their valuable discussions and contribution to this work.

My special thanks to our colleague **Frau Ursula Welscher-Altschäffel** for helping me in learning biochemical techniques, and very often for her useful discussions and kind help. I am thankful to **Dr. Christos Karamitros**. It was a good time with him, and I found him a very hardworking guy. I am very thankful to **Dr. Theresa McSorley** for useful discussions and help. I enjoyed with them the weekly intragroup meetings and scientific discussions at daily lunchtime. Let me say thanks to **Dr. Claudia Höbartner** from nucleic acid chemistry and **Dr. Thomas Jovin** from cellular dynamics for our weekly seminars and discussions.

Abstract

Deoxyribonucleoside and nucleotide kinases are key enzymes that catalyze the critical phosphorylation steps in the conversion of antiviral and anticancer nucleoside analogs to their corresponding cytotoxic nucleoside triphosphates for incorporation into DNA. In this work, I characterized three enzymes: Human guanylate kinase (hGMPK), human mitochondrial thymidine kinase (hTK2), and *E.coli* guanosine-inosine kinase (ecGSK). They were recombinantly produced and kinetically characterized. A series of mutations were introduced to understand the catalytic roles of specific residues. The enzymes were structurally characterized for substrate-induced conformational changes, and two novel assays were devised to study their kinetics. Our main aim was to provide a basis for their potential use in cancer chemotherapy.

The hGMPK was recombinantly produced in catalytically active form although it was previously reported to be inactive upon production in *E.coli*. In order to explain the role of certain residues in catalysis, a series of point mutations were introduced in hGMPK by rational design applying the structural information of mouse GMPK, which is 88% identical to hGMPK allowing us to build a homology model. Interestingly, it was found that a single hydrogen bond between the active site S37 and the carbonyl oxygen of guanine in GMP substrate is critical for binding of GMP and catalysis. Disturbing this single hydrogen bond in the form of the S37A mutation adversely affected the catalytic activity. Besides its catalytic role, S37 is required for the dynamics of the hinge part that connects two structural regions designated as NMP-binding region (NMP-BR) and the CORE region. Its mutation to proline (S37P), which is the least flexible amino acid in terms of sterically allowed conformations, reduced the catalytic efficiency of hGMPK by about 10^3-fold making the molecule more like the non-enzymatic guanylate kinase domain of MAGUKs (membrane-associated guanylate kinase homologs). Similarly, the bidentate interaction of T83 with the carbonyl oxygen of guanine in GMP is required for catalysis. Y81 interacts with the phosphate of GMP and has a role in binary complex stabilization. We demonstrated in cell culture experiments that hGMPK, which catalyzes the second phosphorylation step in the final conversion of the antileukemic drug 6-thioguanine (6-TG) to 6-thioGTP/6-thiodGTP for incorporation into RNA and DNA, is the bottleneck enzyme in the metabolic activation of 6-TG, enhancing its cytotoxicity by several fold when overexpressed in HEK293 cells.

Analyzing the SAXS structures of hGMPK in different conformational states, in particular in the open (unliganded) and completely closed (with two bound nucleotides) forms, revealed large conformational changes that occur during catalysis. The open-to-closed conformational transition of hGMPK induced by binding of ligands supports the model of the induced fit mechanism. In addition, we optimized the higher yield production of isotope-labeled (^{15}N, ^{15}N/^{13}C) hGMPK for its structural analysis by NMR. GMP-induced ^{15}N-^{1}H HSQC (Heteronuclear Single Quantum Coherence) chemical shift changes for hGMPK mapped onto its open form confirmed our findings by SAXS studies that hGMPK undergoes substrate-induced conformational changes.

In order to develop novel and advanced approaches for studying the catalytic properties of deoxyribonucleoside and nucleotide kinases, we devised two assays. In one assay, we used a CdS/ZnS quantum-dot (QDs)-modified gold electrode for the detection of hGMPK-catalyzed reaction in an enzyme-coupled assay based on the electrochemical sensing of NADH in a GMP concentration-dependent way. We also demonstrated the proof of concept of a light-controlled sensor for hGMPK immobilized on CdS/ZnS QDs-modified gold electrode. Similarly, we established an Amplex Red-based spectrophotometric and fluorometric enzyme-coupled assay for hGMPK as an alternative to the conventional NADH-dependent spectroscopic assay. Our new assay overcomes the overlapping wavelength problem associated with strong absorption of 6-thioguanine nucleotides at 340 nm, and it has the advantage of being usable both in the absorbance and fluorescence modes.

We investigated the hGMPK loading capacity of calcium carbonate microparticles of different shapes. It was found that ellipsoidal microparticles with loaded hGMPK exhibited higher specific activities, after coating with polyelectrolytes, as compared to microparticles of all other shapes including spherical, rhomboidal, star and cube-like particles. Thus, ellipsoidal particles turned out to be more appropriate for drug loading and cellular targeting experiments.

We determined the mitochondrial localization of hTK2 by expressing it in HEK293 cells as a fusion with C-terminal EGFP, and observed its subcellular localization by confocal microscopy. In addition, we solved the aggregation problem associated with hTK2 upon overexpression in *E.coli*. Our optimized protocol is based on the expression of hTK2 as a fusion with N-terminal His$_{14}$-MBP-SUMObr tag (~60 kDa) under optimum conditions. Similarly, we generated a C- and N-terminal truncated form of hTK2 with improved catalytic activity as compared to wild-type hTK2, and optimized experimental conditions for its crystallization. Additionally, through

directed evolution using error-prone PCR and subsequent screening of mutants against antiviral and anticancer nucleoside analogs, we found that two mutants designated M5 and M17 increased the sensitivity of the TK-deficient KY895 *E.coli* strain to gemcitabine by 25-fold and fourteen mutants by a factor of 10. Similarly, five mutants enhanced the sensitivity of KY895 to AZT by 3-fold.

A unique property was explored for the recombinantly produced ecGSK that phosphorylates the nucleoside form of the clinically used antileukemic drug 6-thioguanine. To determine its structure-function relationship, experimental conditions were optimized for the crystallization of ecGSK.

Contents

List of Figures

List of Tables

Abbreviations

aa	Amino acid
AMP-PNP	5´-Adenylylimidodiphosphate, a non-hydrolyzable analog of ATP
Ap5G	P^1-(5´-adenosyl) P^5-(5´-guanosyl) pentaphosphate, a bi-substrate analog
AZT	3´-azidothymidine
Dm-dNK	*Drosophila melanogaster* deoxyribonucleoside kinase
dNK	Deoxyribonucleoside kinase
dNTP	Deoxyribonucleoside triphosphate
dGK	Deoxyguanosine kinase
dCK	Deoxycytidine kinase
ecGMPK	*E.coli* guanylate kinase
ecGSK	*E.coli* guanosine-inosine kinase
EGFP	Enhanced green fluorescent protein
epPCR	Error-prone polymerase chain reaction
FAD	Flavin adenine dinucleotide
GMP	Guanosine monophosphate
GMPK	Guanylate kinase
hGMPK	Human guanylate kinase
HPRT	Hypoxanthine–guanine phosphoribosyltransferase
HRP	Horseradish peroxidase
hTK2	Human mitochondrial thymidine kinase
IPTG	Isopropyl β-D-thiogalactopyranoside
LbL	Layer-by-Layer
LDH	Lactate dehydrogenase
lpPOX	Pyruvate oxidase from *Lactobacillus plantarum*
MAGUK	Membrane-associated guanylate kinase

MDS	Mitochondrial DNA depletion syndrome
mGMPK	Mouse guanylate kinase
mtGMPK	*Mycobacterium tuberculosis* guanylate kinase
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NADH	Nicotinamide adenine dinucleotide (reduced form)
NDPK	Nucleoside diphosphate kinase
NMPK	Nucleoside monophosphate kinase
PAH	Poly(allylamine hydrochloride)
PEP	Phosphoenolpyruvate
PK	Pyruvate kinase
PSS	Poly(sodium 4-styrenesulfonate)
QDs	Quantum dots
saGMPK	*Staphylococcus aureus* guanylate kinase
SAXS	Small angle X-ray scattering
SD	Standard deviation
TK1	Cytosolic thymidine kinase
TK2	Mitochondrial thymidine kinase
TMPK	Thymidylate kinase
TPMT	Thiopurine S-methyltransferase
TPP	Thiamine pyrophosphate
yGMPK	Yeast guanylate kinase

1. Introduction

1.1 General characteristics of nucleoside monophosphate kinases

Nucleoside monophosphate kinases (NMPKs) catalyze the reversible phosphoryl transfer reaction by which monophosphates (NMPs) are converted to their corresponding diphosphate forms (NDPs). Ribonucleotides are synthesized by two pathways, the *de novo* pathway and the salvage pathway. In the *de novo* pathway, the nucleosides are synthesized from small molecules to ribonucleoside monophosphates, and subsequently phosphorylated by NMPKs to ribonucleoside diphosphates and by nucleoside diphosphate kinases (NDPKs) to nucleoside triphosphates. The ribonucleoside diphosphate can be reduced to the corresponding deoxyribonucleoside diphosphate catalyzed by ribonucleotide reductase [44]. Nevertheless, in the salvage pathway, preformed (deoxy)ribonucleosides are imported into cells by nucleoside transporters. The (deoxy)ribonucleosides are then phosphorylated to their corresponding triphosphates in three consecutive phosphorylation steps, catalyzed by (deoxy)ribonucleoside kinases, NMPKs, and NDPKs (Fig. 1.1) [42, 45]. Nucleoside analog prodrugs are phosphorylated to their corresponding active triphosphate forms via the salvage pathway [43]. In humans, the NMPK family includes four types of enzymes: guanylate kinase (hGMPK), adenylate kinase (hAMPK), uridylate/cytidylate kinase (hUMP-CMPK) and thymidylate kinase (hTMPK). The names are given according to their preferred substrates, (d)GMP, (d)AMP, (d)UMP/(d)CMP and dTMP [46]. Humans have seven isoforms of guanylate kinase (hGMPK), six AMPKs (hAMPK1–6), two UMP-CMPKs (hUMP-CMPK), and one thymidylate kinase (hTMPK). In addition, a putative mitochondrial thymidylate kinase has also been reported [47, 75]. The hGMPK, hAMPK1, hUMP-CMPK, and hTMPK are located in the cytosol while hAMPK2 is found in mitochondria. Similarly, hGMPK, hAMPK1, hAMPK2 and hUMP-CMPK are monomers whereas hTMPK is a homo-dimeric molecule [46]. Amino acid sequence identities of human NMPKs are very low: hUMP-CMPK is 40, 21 and 20% identical to hAMPK1, hTMPK and hGMPK, respectively. The hGMPK and hTMPK are less sequence-related to other NMPKs. However, all have the same highly conserved three-dimensional (3D) fold [48].

a.

Free bases   Ribonucleosides   Deoxyribonucleosides

Salvage of free
nucleoside bases   Salvage of
ribonucleosides   Salvage of
deoxyribonucleosides

Small molecules → Ribonucleotides → Deoxyribonucleotides

De novo synthesis of
ribonucleotides   De novo synthesis of
deoxyribonucleotides

DNA

b.

Ado
Cyd
Urd

dAdo dGuo
dCyd dThd
dUrd

dGuo
dThd   CP
dUrd ← dCDA
NP   dCyd
dAdo

cell membrane   CP

Ado
Cyd ┐
Urd ← CDA
Guo

5′NT ↕ dNK
dAMP
dGMP
dCMP ┐ dCMP deaminase
dUMP
dTMP ← TS

5′NT | rNK

NP

NMPK

dUTPase

Small molecules → AMP
GMP
CMP ┐ dCMP
UMP ← deaminase

G
T
U

ADP
GDP   RR
CDP
UDP

dADP
dGDP
dCDP
dUDP
dTDP

NMPK

U
G
A
C
PRT

NDPK   NDPK

dUTP

DNA ← dATP
dCTP
dGTP
dTTP

RNA ← ATP
GTP
CTP ← CTP-S
UTP

Fig. 1.1. *De novo* and salvage synthesis of ribo- and deoxyribonucleotides. (a) Key steps in the ribonucleoside and deoxyribonucleoside metabolism. (b) Detailed steps in the ribonucleoside and deoxyribonucleoside metabolism. A, adenine; C, cytosine; CDA, cytidine/deoxycytidine deaminase; CP, carrier protein; CTP-S, CTP synthase; dNK, deoxyribonucleoside kinase; dUTPase, deoxyuridinetriphosphatase; G, guanine; NP, purine/pyrimidine nucleoside phosphorylase; 5′NT, 5′-nucleotidase; PRT, phosphoribosyl transferase; rNK, ribonucleoside kinase; T, thymine; U, uracil [45].

NMPKs have generally three structural parts, a CORE region, a LID region, and an NMP-binding region. The CORE includes the ATP binding P-loop, the NMP-binding region binds the (d)NMP substrate, and the LID region provides catalytic residues for the reaction (Fig. 1.2) [46, 49]. The NMP substrate is mainly bound to the NMP-binding region, however it also interacts with some residues of the CORE [11, 50]. These three regions are interconnected by hinges. NMPKs are known to undergo large conformational changes upon binding of their substrates [51]. The enzyme attains the open conformation in the absence of any substrate. When the substrates bind, the NMP-binding and the LID parts undergo large hinge-bending motions that make the subunit switch from an open to a closed conformation [9]. These conformational changes were first described for adenylate kinase, and a two-state model of induced-fit mechanism was suggested in which the substrates, in particular AMP, can easily trigger the change between two stable conformations. The B conformation related to the structure of free enzyme, and the A conformation corresponded to the enzyme conformation after an induced-fit [49, 52]. NMP kinases can be inhibited by bisubstrate analogs which occupy both the phosphate donor site and the phosphate acceptor site: for instance, Ap5G inhibits *E.coli*, yeast, mouse and human GMPKs, and Ap5A is an inhibitor for hAMPK1 (Fig. 1.2). These bisubstrate analogs occupy the binding positions of both substrates, NMP and ATP, with a linker of four or five phosphates in between [46].
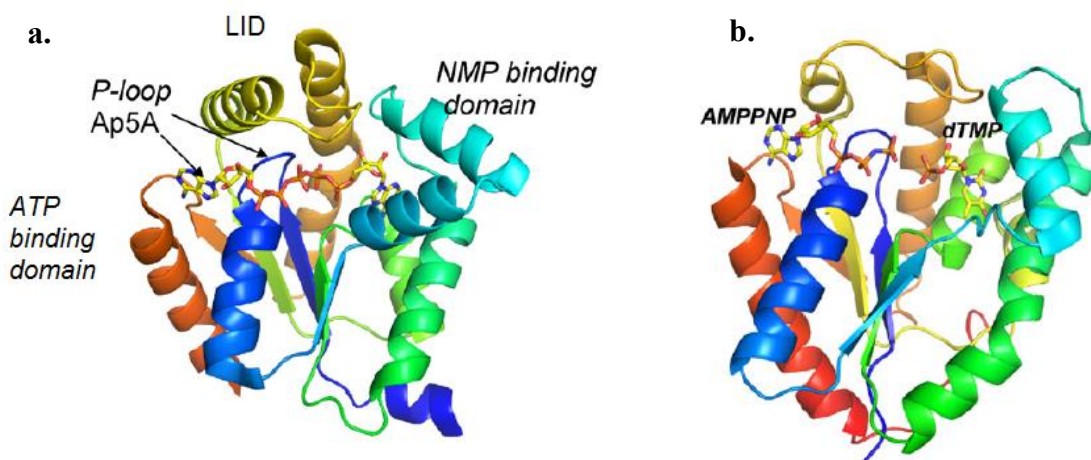
Fig. 1.2. The NMPK monomer fold. (a) Human AMPK1 complexed with Ap5A (P^1-(5´-adenosyl) P^5-(5´-adenosyl) pentaphosphate), an inhibitor that occupies both the donor nucleotide site on the left and the acceptor site on the right. (b) Human TMPK (PDB 1E9E) in complex with TMP and an ATP analog (AMPPNP, Adenylyl-imidodiphosphate). UMP-CMPK and GMPK adopt the same fold as AMPK and TMPK [46].

1.1.1 Guanylate kinases

Guanylate kinase (GMPK, ATP:GMP phosphotransferase) is a member of the family of ATP:NMP phosphoryltransferases (nucleoside monophosphate kinases; NMP kinases; NMPKs; EC 2.7.4.8) which catalyzes the reaction (d)GMP + ATP ↔ (d)GDP + ADP [25]. The reaction and biological function of GMPKs closely resemble that of adenylate kinases which use (d)AMP as a phosphoryl group acceptor [53]. There is a glycine-rich P-loop in the N-terminus of GMPKs (11-18 amino acids in mouse GMPK, and 9-16 residues in yeast GMPK) and adenylate kinases that binds α and β phosphates of ATP. Nevertheless, besides similar ATP binding sites, the chain fold motifs of the GMP binding domain of GMPKs and AMP binding domains of adenylate kinases are totally different [53]. As mentioned before, the 3D fold of NMP kinases is highly conserved. But there are also structural differences: for instance, the NMP-binding region of GMPKs consists of a four-stranded β-sheet and only a short helix [11, 53], whereas the NMP-binding domains of other NMP kinases are α-helical [10]. GMPKs have been studied from several organisms, and many functional and structural details are known, for example, the X-ray structures of *E.coli* GMPK (ecGMPK) [54], *Mycobacterium tuberculosis* GMPK (mtGMPK) [59], *Staphylococcus aureus* GMPK (saGMPK) [60], yeast GMPK (yGMPK) [10, 53, 55] and mouse GMPK (mGMPK) [11]. Analysis of these structures indicates that GMPKs have common structural features including the CORE domain that carries the ATP β-phosphate binding glycine-rich P-loop, NMP-binding domain for binding GMP, and LID domain which interacts with the adenine base of ATP and provides catalytic residues for the phosphoryl transfer reaction [54]. In the absence of any substrate, GMPKs attain an extended open conformation in which the NMP-binding and LID regions are at higher distance from each other and more mobile allowing full access to the ATP and GMP binding sites [10]. GMP binding induces partial closure of the GMPK molecule, and complete closure is achieved by the binding of both substrates, ADP (or non-hydrolysable ATP-analog) and GMP [10, 11]. Up to now, there is no structural data available for human GMPK (hGMPK). Therefore, the three-dimensional structure elucidation of hGMPK will be of great relevance for understanding its structural and functional characteristics. The hGMPK enzyme is 197 amino acids in length and is 88% identical to mGMPK, 51% identical to yGMPK, 41% identical to ecGMPK, 35% identical to saGMPK and 31% to the guanylate kinase-like domain of hDlg (MAGUK from *Homo sapiens*, UniProt identifier Q12959)

(Fig. 1.3). The ecGMPK and saGMPK molecules are multimeric unlike *Mycobacterium tuberculosis*, yeast and mammalian GMPKs which are monomers [35, 58, 60].

```
              P-loop                                    NMP-BR
      β1              α1            β2    β3             β4
hGMPK  --MSGPRPVVLSGPSGAGKSTLLKRLLQ-EHSGIFGFSVSHTTRNPRPGEENGKDYYFVT 57
mGMPK  --MAGPRPVVLSGPSGAGKSTLLKKLFQ-EHSSIFGFSVSHTTRNPRPGEEDGKDYYFVT 57
ecGMPK --MAQGTLYIVSAPSGAGKSSLIQALLKTQPLYDTQVSVSHTTRQPRPGEVHGEHYFFVN 58
saGMPK MDNEKGLLIVLSGPSGVGKGTVRKRIFE-DPSTSYKYSISMTTRQMREGEVDGVDYFFKT 59
yGMPK  ----MSRPIVISGPSGTGKSTLLKKLFA-EYPDSFGFSVSSTTRTPRAGEVNGKDYNFVS 55
hDlg   -----TRPVIILGP---MKDRINDDLIS-EFPDKFGSCVPHTTRPKRDYEVDGRDYHFVT 51
                                          * *         *

                 NMP-BR
        α2          β5      β6    α3        β7        α4
hGMPK  -REVMQRDIAAGDFIEHAEFSGNLYGTSKVAVQAVQAMNRICVLDVDLQGVRNIKAT-DL 115
mGMPK  -REMMQRDIAAGDFIEHAEFSGNLYGTSKEAVRAVQAMNRICVLDVDLQGVRSIKKT-DL 115
ecGMPK -HDEFKEMISRDAFLEHAEVFGNYYGTSREAIEQVLATGVDVFLDIDWQGAQQIRQK-MP 116
saGMPK -RDAFEALIKDDQFIEYAEYVGNYYGTPVQYVKDTMDEGHDVFLEIEVEGAKQVRKK-FP 117
yGMPK  -VDEFKSMIKNNEFIEWAQFSGNYYGSTVASVKQVSKSGKTCILDIDMQGVKSVKAIPEL 114
hDlg   SREQMEKDIQEHKFIEAGQYNNHLYGTSVQSVREVAEKGKHCILDVSGNAIKRLQIA-QL 110
                              * *

                        LID region
      β8        α5                α6                    β9
hGMPK  RPIYISVQPPSLHVLEQRLRQRNTETEESLVKRLAAAQADMESSKEPGLFDVVIINDSLD 175
mGMPK  CPIYIFVQPPSLDVLEQRLRLRNTETEESLAKRLAAARTDMESSKEPGLFDLVIINDDLD 175
ecGMPK HARSIFILPPSKIELDRRLRGRGQDSEEVIAKRMAQAVAEMSHYA---EYDYLIVNDDFD 173
saGMPK DALFIFLAPPSLEHLRERLVGRGTESDEKIQSRINEARKEVEMMN---LYDYVVVNDEVE 174
yGMPK  NARFLFIAPPSVEDLKKRLEGRGTETEESINKRLSAAQAELAYAET-GAHDKVIVNDDLD 173
hDlg   YPISIFIKPKSMENIMEMNKRLTEEQARKTFERAMKLEQEFTEHFT-----AIVQGDTLE 165

         α7      α8
hGMPK  QAYAELKEALSEEIKKAQRTGA------------ 197
mGMPK  KAYATLKQALSEEIKKAQGTGHA----------- 198
ecGMPK TALTDLKTIIRAERLRMSRQKQRHDALISKLLAD 207
saGMPK LAKNRIQCIVEAEHLKRERVEAKYRKMILEAKK- 207
yGMPK  KAYKELKDFIFAEK------------------- 187
hDlg   DIYNQVKQIIE---------------------- 176
```

Fig. 1.3. Sequence alignment of bacterial and eukaryotic GMPKs. From top to bottom, GMPK amino acid sequences are from human, mouse (88% identity), *E.coli* (41% identity), *Staphylococcus aureus* (35% identity), yeast (51% identity) and human GK domain (31% identity). Conserved residues are shown with orange background. On the top of the sequences are indicated the mGMPK secondary structures. The CORE region and hinges include all regions outside NMP-binding region (NMP-BR) and LID region. Below the sequences are indicated the residues (∗) which were mutated for kinetic characterization in hGMPK.

1.1.1.1 *Escherichia coli* **guanylate kinase**

Gunaylate kinase (GMPK, ATP:GMP phosphotransferase, EC 2.7.4.8) from *Escherichia coli* (ecGMPK) was preliminarily characterized by Oeschger and Bessman in 1966 [39]. It is 207 amino acids in length (UniProt ID P60546) having 23,462 Da molecular mass of the monomer as determined by electrospray ionization-mass spectrometry (ESI-MS) [60]. Like other GMP kinases, ecGMPK uses both GMP and dGMP as substrates. The guanylate kinase reaction using dGMP as a substrate is stimulated by K^+ and NH_4^+ ions [58]. The ecGMPK was found to exist in equilibrium between various oligomeric species (dimeric, tetrameric and hexameric) in solution depending on the ionic strength and protein concentration [60]. As its catalytic activity is similar to that of monomeric GMPKs and shows no cooperativity for the GMP substrate, it is likely that its active sites are essentially independent in the oligomers. The crystal structure of ecGMPK was reported in nucleotide-free form, in complex with GMP, GDP [60], with the monophosphate of the antiherpes drug ganciclovir (GCV-MP), and with the bi-substrate inhibitor Ap5G [54]. It was crystallized as a hexamer. The ecGMPK subunit shares the same characteristic 3D fold with monomeric yeast GMPK and mouse GMPK consisting of CORE, LID and NMP-binding domains (Fig. 1.4) [10, 11]. The binding of substrates induces domain closure in ecGMPK from a fully open conformation when no nucleotides are bound (apo-ecGMPK) to a partially closed NMP-binding domain in the presence of GDP (ecGMPK$_{GDP}$), and essentially closed LID and NMP-binding domains in the presence of both GMP and sulfate ions (ecGMPK$_{GMP.SO4}^{2-}$) [60]. The ATP-binding site remains partially obstructed by inter-subunit interactions in all the presented crystal structures, indicating that either local conformational changes or dissociation of the hexamer are required. The conserved arginine residues for example R138LID and R45NMP play an important role in transition state stabilization. The R138 (R137 in mGMPK) which interacts with α and β phosphates of ADP may bridge the ADP leaving group and the transferred phosphates at the transition state (TS) (Fig. 1.4b), whereas R45 (R44 in mGMPK) which binds α phosphate of GMP may bridge the GMP to the transferred phosphoryl group in TS. The R149LID (R148 in mGMPK) which binds to the α-β bridging oxygen atom of GDP may rather be involved in stabilizing the GMP substrate in the ground state. It is because charges at this atom are expected to decrease in the TS and hence the contribution of this residue to the stabilization of the TS [60].
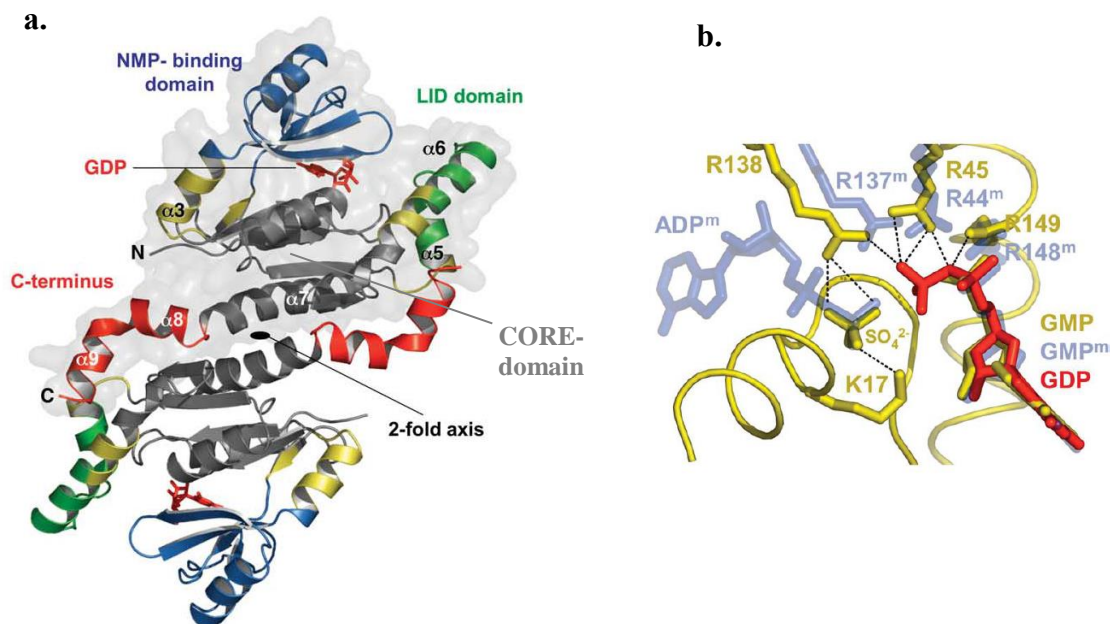
Fig. 1.4. *E.coli* guanylate kinase. (a) Dimeric interface of ecGMPK. The CORE domain is shown in *grey*, the NMP-binding domain in *blue*, the LID domain in *green*, hinges in *yellow* and the C-terminal extension in *red*. GDP is shown in red ball-and-stick. One monomer is contoured with its van der Waals surface. (b) A model of catalytic interactions of conserved arginine residues. Overlay of GDP (red) from the ecGMPK$_{GDP}$ structure onto GMP in the closed conformation of ecGMPK$_{GMP\cdot SO4}^{2-}$ (in yellow), based on the superposition of the NMP-binding domain. Candidate hydrogen bonds of the conserved arginine residues to the phosphate groups are shown as dotted lines. ADP, GMP and the invariant arginine residues from the mGMPK$_{GMP.ADP}$ structure are superposed to show the equivalence of the sulfate ion with the β-phosphate group of ADP (in blue) [60].

Unlike the eukaryotic orthologs including yeast, mouse and human GMPKs, the ecGMPK has a unique C-terminal extension. The C-terminal extension has a role in forming the basic dimeric folding unit of the hexamer, and the tyrosine insertion in the CORE domain is a specific feature of the trimeric interface. Most of the enterobacterial and the *Vibrio cholerae* GMPKs are highly related to ecGMPK with respect to these signature regions. Therefore, the oligomeric nature of ecGMPK in solution and its hexameric arrangement in the crystal serve as a model for the quaternary structure of GMPKs from these bacteria including human pathogens such as *Salmonella typhimurium*, *Yersinia pestis* and *V. cholerae*. Being an essential enzyme of bacteria, GMPK is a potent drug target. As none of the ecGMPK structures was able to bind ATP due to blockage of the adenine-binding site by subunit interactions at the trimeric interface, this auto-inhibited conformation could be specifically stabilized for bacterial cells. By doing so, the

pathogenic bacteria could be targeted with the designed stabilizing inhibitor. Similarly, the non-substrate GDP binding site at the trimeric interface in ecGMPK interferes with both ATP binding and LID domain closure and can be used as a potential drug target [60].

1.1.1.2 Yeast guanylate kinase

Yeast guanylate kinase (yGMPK, GUK1, ATP:GMP phosphotransferase, EC 2.7.4.8, UniProt ID P15454) is 187 amino acids in length and has 20,637 Da molecular weight. Its primary structure contains one cysteine and one tryptophan residue. It is a monomeric protein, and has the N-terminally located phosphate-binding loop (GXXGXGKS, 9-16 aa, GPSGTGKS) [15, 25]. The enzyme was purified to homogeneity by Berger et al in 1989 [25], and one year later the three-dimensional structure (crystal structure) was determined in complex with GMP by Stehle and Schulz [55]. The steady-state kinetic parameters for both forward and reverse reactions were determined by Li et al in 1996 [35]. The k_{cat} (turnover number) was 394 s^{-1} for the forward reaction (formation of ADP and GDP) and 90 s^{-1} for the reverse reaction (formation of ATP and GMP). The K_m values were 0.20, 0.091, 0.017, and 0.097 mM for MgATP, GMP, MgADP, and GDP, respectively. It was observed that at concentrations above 0.22 mM, the initial velocity decreased with increasing GMP concentration and leveled off at ~50% of the apparent maximum velocity. This partial substrate inhibition was not competitive with MgATP and may be due to the formation of an abortive complex, yGMPK.MgADP.GMP [35]. Like other NMP kinases, yGMPK consists of three structural regions; the CORE, LID, and NMP-binding domains. The NMP-binding domain of yGMPK is composed of a four-stranded β-sheet (β3, β4, β5 and β6) and one α-helix (α2). It is unlike other NMP kinases whose NMP-binding domains are all α-helical [10]. The CORE domain of yGMPK consists of a five-stranded parallel β-sheet (β1, β2, β7, β8 and β9) and six α-helices (α1, α3, α4, α5, α6 and α7) as shown in (Fig.1.5). The LID domain is composed of one loop (seven residues in length, 135-141 aa) which connects α5 and α6. The comparison of the crystal structures of apo-yGMPK (no bound nucleotide) and yGMPK$_{GMP}$ (bound GMP) revealed that the binding of GMP induces a major movement of the NMP-binding domain and a smaller movement of the LID domain. Remarkably, the NMP-binding domain moves toward the central CORE domain upon binding of GMP, the LID domain moves away from the CORE domain. The overall effect is the closing of the yGMPK molecule (partially

closed conformation). Helix3 which connects the CORE and NMP-binding domains was supposed to play an important role in the substrate-induced domain movements [10].
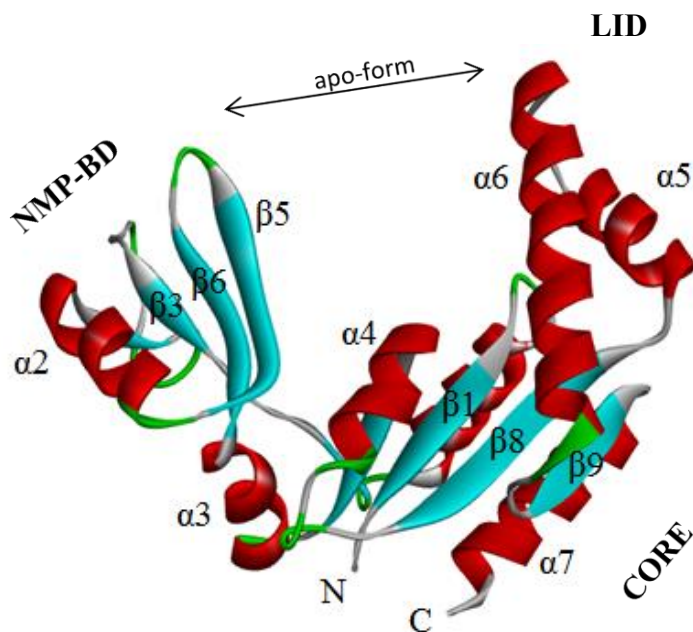
Fig. 1.5. Dynamic domains in apo-yGMPK. Three main structural regions of unliganded yGMPK (PDB IEX6), CORE, LID and NMP-binding domains are indicated. The NMP binding domain contains α2, β3, β4, β5, and β6; the CORE domain contains α1, α3, α4, α5, α6, α7, β1, β2, β7, β8, and β9; and the LID domain contains the loop between α5 and α6 [10].

N-terminal acetylation is one of the most common co-translational covalent modifications of proteins in eukaryotes [62, 63]. It was suggested that one function of N-acetylation of cellular proteins is to prevent their degradation by the ubiquitin system [64, 65]. However, recombinant eukaryotic proteins produced in *E.coli* are not acetylated. The yGMPK is N-terminally acetylated when purified from its natural source, but its recombinant form was not acetylated. When both acetylated and non-acetylated structures were compared for any structural and functional consequences, there was no significant impact on the three-dimensional structure [10, 53]. Nevertheless, a smaller hydrogen bond distance was found between T94 and the N-terminus in native yGMPK (2.78 Å) as compared to the hydrogen bond distance of 3.19 Å in case of recombinant yGMPK. The stronger interaction in case of native yGMPK may have a role in its stability [10].

NMP kinases undergo large conformational changes upon binding of substrates as shown by comparing 17 crystal structures of NMPKs in various states [9]. These conformational changes were mainly attributed to the movements of two domains, NMP-binding domain and LID domain. In case of yGMPK, the unliganded enzyme and its complex with GMP were determined by X-ray crystallography [10]. By aligning the two structures, it was found that yGMPK indeed undergoes substrate-induced conformational changes from an unliganded open conformation to a liganded partially closed conformation. The large domain movement involved the rotation around the hinge axis parallel to helix3 which connects the NMP-binding domain and CORE domain. Hinges are normally flexible regions that connect NMP-binding and LID domains to the CORE domain. The rotation of the amphipathic helix3 alters interactions between helix3 and the CORE domain. GMP binding brings major mobility in the C-terminal part of helix 3 which is unlikely due to the loss of van der Waals interactions (vdw) between the helix and the CORE domain because most of the vdw contacts between the C-terminal part of helix3 and the CORE domain remain upon GMP binding. It is the N-terminal part of the helix that loses its contacts with the CORE domain. It was proposed that helix3 acts like a spring in the movements of the NMP-binding domain which may facilitate the binding of substrates and release of the products. GMP-binding also induces small movements of the LID domain which moves away from the active center to make the molecule slightly more open. It is unlike the NMP-binding domain that moves towards the active center of yGMPK making the molecule more closed. The net effect is the partial closing of the molecule. Such domain movements were also observed in case of mouse GMPK upon binding of GMP and ADP [11]. In case of *E.coli* adenylate kinase, as studied by the method of time-resolved fluorescence resonance energy transfer [66], AMP binding caused the closure of the LID domain. The distances between the excitation energy donors and acceptors attached to residues 73 at the CORE domain and 142 at the LID domain were shortened by 9 Å. This means that the LID domain moved towards the active center which is in contrast to what we see in case of yGMPK where GMP causes the LID domain to move away from the active center causing the ATP binding site to further open, and thus facilitating ATP binding [10].

The glycine-rich P-loop is located in the N-terminus of GMPKs (9-16 aa in yGMPK), adenylate kinases and UMP-CMP kinases, and binds α and β phosphates of ATP. The ATP-binding site is highly conserved among NMP kinases. Nevertheless, besides similar ATP binding sites, the

chain-fold motifs of the NMP-binding domain of GMPKs are totally different from those of adenylate and UMP-CMP kinases [53]. The NMP-binding domain of AMPKs and UMP-CMPKs is completely α-helical whereas in GMP kinases it consists of a four-stranded β-sheet and a short helix. Only two arginine residues are involved in binding the phosphate group of NMP in adenylate kinase or UMP/CMP kinase. The phosphate moiety of bound NMP in GMPKs interacts with a pair of arginine residues as well as a pair of tyrosine residues. When the Y78 in yGMPK was substituted by phenylalanine, the k_{cat} was decreased by a factor of 131, and the K_m value for GMP was increased by a factor of ~20 [61]. The K_m for ATP was increased by only a factor of ~2. The conformational stability of the wild-type and Y78F yGMPK was studied by GdnHCl denaturation experiments [61]. The results showed that the hydroxyl group of the Y78 side-chain contributes to the conformational stability by ~1.0 kcal/mol. A single mutation. Y78F. changed both the kinetic properties and conformational stability of yGMPK. However, these changes were not due to global structural perturbations as investigated by 2D NMR [61]. It is because the Y78F mutant was properly folded and its conformation was highly similar to that of the wild-type yGMGK. Also, the change in the kinetic properties of ATP was insignificant. The mutation, however, significantly changed the chemical shift of Y50 which like Y78 interacts with the phosphate of GMP [53]. The changes in the chemical shift of Y50 could be due to changes in the local microenvironment caused by the Y78F mutation. Similarly, the Y50F mutation also changed the chemical shift of Y78 in the GMP-bound form [67]. Nevertheless, Y78F or Y50F did not change the chemical shift of the other tyrosine residue in the unliganded state of yGMPK because both residues were no longer interacting. The changes in the kinetic properties were unlikely due to a decrease in the conformational stability because the mutant remained stable for days at room temperature. All these observations suggest that the changes in kinetic properties and conformational stability of Y78F mutant are due to the loss of a hydrogen bond between its side-chain hydroxyl group and the phosphate of the GMP substrate. It was estimated from the kinetic data that the hydrogen bond between Y78 and GMP phosphate stabilizes the binary complex by 1.7 kcal/mol, the ternary complex by 1.8 kcal/mol, and the transition state by 4.6 kcal/mol [61]. The Y50 also forms a hydrogen bond with GMP phosphate and contributes more (~0.5 kcal/mol) to the stability of binary and ternary complexes as compared to Y78 [67]. It is because this hydrogen bond is slightly shorter and a bit stronger than that between Y78 and GMP. It was reported that in comparison to Y50, the Y78 contributes

more to the stabilization of the transition state by 1.4 kcal/mol. Therefore, Y78 plays a more important role in the enzymatic catalysis of yGMPK. It also indicates that the hydrogen bond between Y78 and GMP becomes stronger than that between Y50 and GMP as the reaction proceeds to the transition state [67].

1.1.1.3 Mouse guanylate kinase

Mouse guanylate kinase (mGMPK, GUK1, ATP:GMP phosphotransferase, EC 2.7.4.8, UniProt Q64520) is a small polypeptide of 198 amino acids in length [36]. It is a monomeric protein with molecular weight of 21,917 Da as calculated from its amino acid sequence. The mGMPK is 11 residues longer than yGMPK (187 aa), two of these amino acids are located at the N terminus and nine are located at the C-terminal part of the protein. It has two cysteine residues but no tryptophan, and has the characteristic N-terminal P-loop pattern (GXXGXGK, 11-18 aa, GPSGAGKS). Like other NMP kinases, GMP kinases undergo substrate-induced conformational changes as part of their catalysis [3, 10]. GMPKs are bi-substrate enzymes which catalyze the transfer of a phosphoryl group from ATP to GMP which acts as a phosphate acceptor. The binding of either substrate or their analogs induce conformational changes as studied by X-ray crystallography and NMR [10, 11, 13, 53-55, 60, 61]. Comparing the structures of nucleotide-free GMPK to the binary complex in which one substrate is bound (ATP or GMP), and to the ternary complex in which both substrates are bound, indicated conformational changes. Such different conformational states are called open form with no bound nucleotide, partially closed form with one bound substrate and completely closed form with two bound substrates. These conformational changes are due to rigid body movements of the three structural regions CORE, LID, and NMP-binding regions (NMP-BR). The crystal structure of mGMPK (PDB 1LVG) in the closed conformation with bound GMP and GDP was reported by Sekulic et al in 2002 [11]. The fold of mGMPK is very similar to that of yGMPK consisting of three structurally and functionally distinct parts (Fig. 1.6) [10, 11]. These are CORE region (residues 5–31, 97–123, and 165–194; helices α1, α4, α7, and α8; strands β1, β7, β8, and β9), NMP-BR (residues 37–89; helices α2 and α3; strands β3, β4, β5, and β6) and LID region (residues 126–156; helices α5 and α6). The NMP-binding domain of yGMPK is composed of a four stranded β-sheet and one α-helix, and in case of adenylate kinases it is all helical [10, 68]. The CORE region of mGMPK is connected to NMP-BR and LID region via four hinges. Hinge 1 (residues 32–36) and hinge 2

(residues 90–96; part of helix 3) connects CORE and NMP-BR, whereas hinge 3 (residues 124–125) and hinge 4 (residues 157–164; part of helix 6) join CORE and LID region [11].
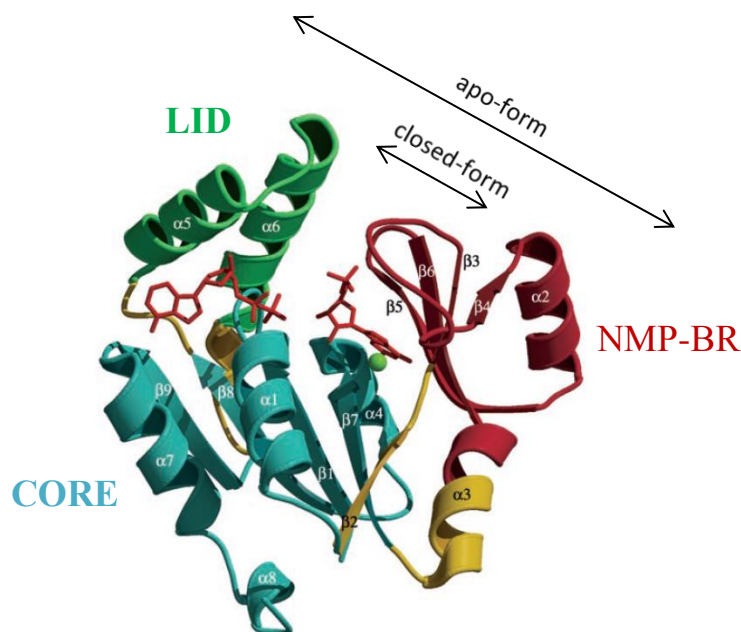
Fig. 1.6. Ribbon diagram of mGMPK$_{GMP-ADP}$ in closed conformation. Different regions of the enzyme are color-coded. *cyan*, CORE region; *red*, NMP-binding region (NMP-BR); *green*, LID region; *yellow*, interconnected with four hinges. The nucleotides ADP and GMP (*red*) and the potassium ion (*green sphere*) are also shown [11].

The closed conformation of mGMPK$_{GMP.ADP}$ (with bound GMP and ADP) was overlaid on the apo-yGMPK (with no bound nucleotide) and yGMPK$_{GMP}$ (with bound GMP) to see the effect of substrate binding on the conformation of GMPK. It was found that the NMP-BR and LID region move upon binding of substrates. The farthest distance between NMP-BR and LID region was found in the apo-yGMPK designated as open conformation [10]. GMP binding caused a significant movement of the NMP-BR towards the LID region with relatively small move of the LID in the same direction. The net effect was to bring the two regions closer for binding GMP and resulted in a partially closed conformation. The binding of two nucleotides, GMP and ADP, further pulled the two regions closer to each other and to the CORE region, forming a more compact closed conformation.

The binding interactions of GMP in mGMPK and yGMPK are similar. The specificity of GMPKs for GMP is achieved by discriminating the guanine base of GMP from adenine of AMP. The active site residue S37 interacts with the carbonyl oxygen (an amino in adenine) of guanine at position 6 via a single hydrogen bond as shown in (Fig. 1.7a-d).
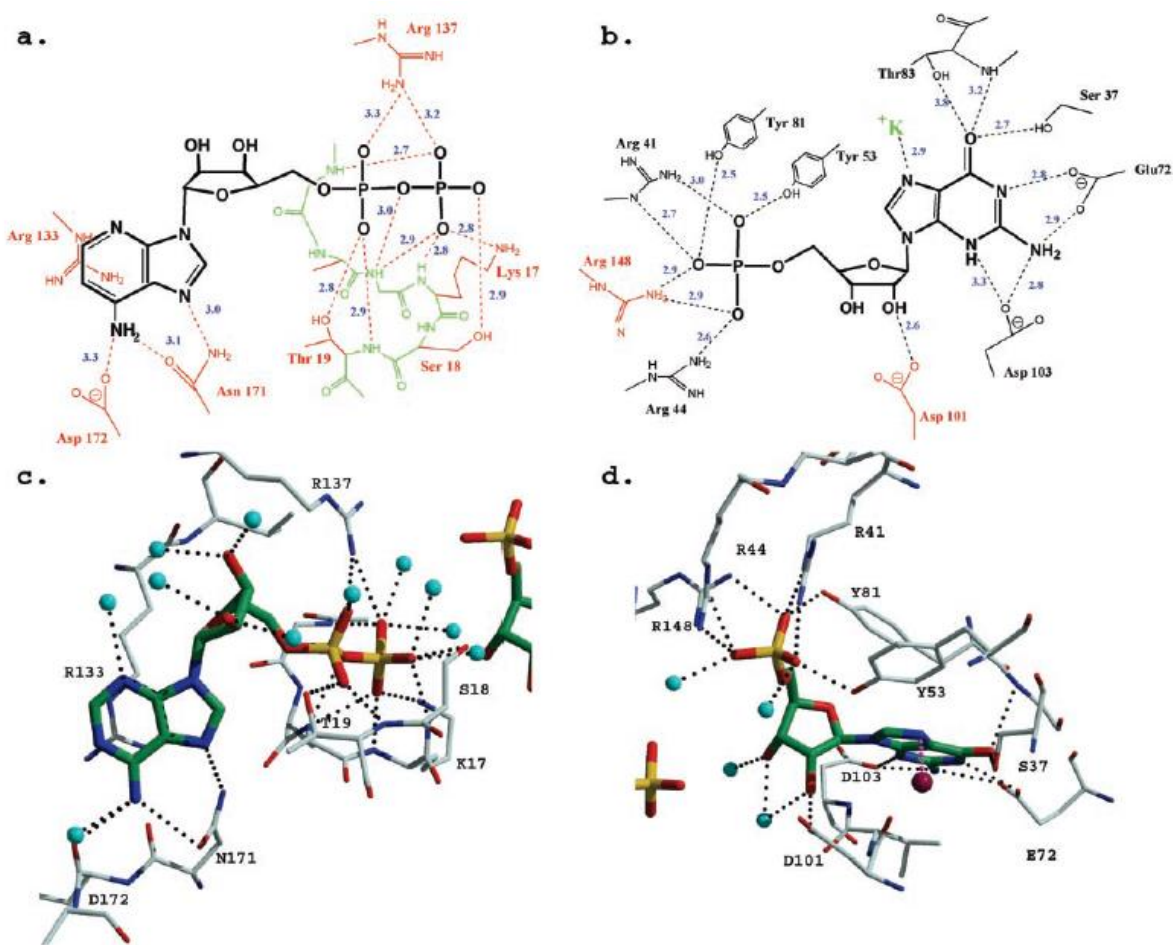
Fig. 1.7. Binding sites of GMP and ADP. *a* and *b*, distance map showing residues involved in binding of ADP (*a*) and GMP (*b*). For clarity, backbone atoms of the P-loop are shown in *green*. Residues that are making interactions not previously observed in the yGMPK$_{GMP}$ structure are shown in *red*. For example, the D101 interaction with the GMP ribose observed in the mGMPK$_{ADP.GMP}$ complex is made possible by the additional closing of the structure as a result of ADP binding and is absent in the yGMPK$_{GMP}$ structure (D98 in yGMPK). The distances are in angstroms. *c* and *d*, ball-and-stick representation of the ADP-binding (*c*) and GMP-binding (*d*) sites [11].

Similarly, the E72 carboxyl group forms two hydrogen bonds with guanine, one with protonated N1 (unprotonated in adenine) and the other with the amino group at position 2 (hydrogen in adenine). In contrast, the specificity for adenine in adenylate kinases is accomplished by glutamine (E72 in mGMPK) via a bidentate interaction with the amino group at position 6 and the N1 (as hydrogen bond acceptor) [69, 70]. The D101 interacts with the 2′ hydroxyl group of the ribose sugar in GMP; however, this interaction cannot be formed in the dGMP complex, and this explains why it was reported as poor substrate for yGMPK [35]. Like yGMPK, in mGMPK$_{GMP.ADP}$ the Y53 and Y81, and R41 and R44 interact with the phosphate of GMP. However, the hydrogen bond between R44 and phosphate in mGMPK is stronger (2.6 Å) than that in yGMPK (3.4 Å). Besides that, an additional R148 from the LID region also interacts with the phosphate of GMP, and this interaction does not exist in yGMPK. In analogy to the uridylate kinase structure where in a complex that mimics the transition state, it was found that the arginine that would correspond to R148 in mGMPK interacts with the transferred phosphoryl group. It means that in case of mGMPK the R148 would interact with the γ-phosphate of ATP on one side and with the α-phosphate of GMP on other side to facilitate the phosphoryl group transfer. It was reported that the mGMPK double mutant, E72Q/D103N, can phosphorylate AMP although this adenylate kinase activity was much less than that of wild-type adenylate kinases i.e., the specific activity of E72Q/D103N mutant was ~310 times less than the wild-type adenylate kinase [72]. In adenylate kinases, the glutamine analogous to E72 interacts with N1 and the amino group at position 6 of adenine, however in the mouse double mutant E72Q/D103N, the interaction of glutamine with N1 is favorable, but binding with the amino group at position 6 is intervened by residues S37 and T83 [11]. As a result, the double mutant could not mimic its counterpart in adenylate kinase. Therefore, to achieve this goal, S37 and T83 may also need to be mutated to residues whose side chains are not bulky like glycine for instance [11].

In case of yGMPK, it was found that GMP causes partial substrate inhibition, and at concentrations above 0.22 mM, the initial velocity of the reaction was decreased by increasing the GMP concentration [35]. This is due to the formation of an abortive complex yGMPK$_{GMP.ADP}$ which however does not arrest the release of ADP. The structure of this complex was determined in case of mGMPK as explained above [11]. ADP binds at the ATP binding site which is located between CORE and LID region. The glycine-rich P-loop which resides in between β1 and α1

(11-18 aa) in the CORE region binds the phosphates of ADP/ATP [71]. Moreover, K17 and S18 from the P-loop interact with the β-phosphate, and R137 from the LID region interacts with both β and α-phosphates of ADP/ATP. T19 binds the α-phosphate. The latter residues could potentially bind the γ-phosphate of ATP to play their role in the phosphoryl group transfer reaction. As observed in other NMP kinases [53], the ADP/ATP-ribose is stabilized by water molecules and does not interact with any surrounding residues. The adenine base is bound by R133 by stacking interaction and through hydrogen bonds by N171 and D172 [11].

1.1.1.4 Human guanylate kinase

1.1.1.4.1 Biological significance

Guanylate kinase (GMPK, ATP:GMP phosphotransferase, EC 2.7.4.8) is an essential enzyme involved in guanine nucleotide metabolism of unicellular and multicellular organisms, which acts by catalyzing the reversible phosphoryl group transfer from ATP to (d)GMP resulting in (d)GDP and ADP [11-13]. These nucleotides are required for a variety of cellular metabolic processes, as well as for RNA and DNA synthesis [42]. Guanylate kinase activity was first reported by Klenow and Lichtler in 1957 [24], and was initially characterized from different sources including human erythrocytes, hog brain, mouse, yeast, *Arabidopsis thaliana*, and *Escherichia coli* [25-29, 37, 39]. It plays an important role in the recycling of second messenger (cGMP) via the cyclic GMP cycle (cGMP→GMP→GDP→GTP→cGMP), and thereby regulates the supply of guanine nucleotides to signal transduction pathways [14, 15, 36, 38]. Besides its physiological roles, the enzyme is required for intracellular activation of numerous antiviral and anticancer purine nucleoside analog prodrugs [11, 312]. Prominent examples are the FDA-approved drugs azathioprine, 6-mercaptopurine, 6-thioguanine, ganciclovir and acyclovir. Azathioprine is commonly used as an immunosuppressive agent to prevent graft rejection in organ transplant patients and for treating autoimmune diseases [31]. 6-mercaptopurine and 6-thioguanine are commonly prescribed for the treatment of acute lymphoblastic leukemia [17-19, 31]. These two drugs are also effective against other diseases including colitis, psoriasis and rheumatoid arthritis [17]. In addition, the 2´-deoxy-guanosine analog prodrugs ganciclovir and acyclovir, which are used as efficient agents for the treatment of herpes infections, are first phosphorylated by viral thymidine kinase, and then converted to the diphosphate forms by cellular GMPK [11, 34]. Similarly, the anticancer drug 9-β-D-arabinofuranosylguanine (araG)

after initial phosphorylation by deoxynucleoside kinases is diphosphorylated by cellular GMPK before being converted into the pharmacologically active triphosphate form [41]. Guanylate kinases from microorganisms like *Staphylococcus aureus* and *Mycobacterium tuberculosis* offer new chemotherapeutic targets [32, 33]. It has been suggested that the GMPK enzyme evolved into a non-enzymatic GMPK domain of MAGUKs (membrane associated guanylate kinase homologs) which perform completely different functions (neo-functionalization) such as cell junction formation and mitotic spindle orientation [40]. Despite all these important roles, no structural data is available for the medically most relevant hGMPK enzyme. Therefore, it is of great interest to study the structure and function of hGMPK.

1.1.1.4.2 General characteristics

In human tissues, seven isozymes of guanylate kinase were found [75]. Three are primary isozymes called a, c and e. Four are secondary isozymes; b is secondary to a, d is secondary to c, and f & g are secondary to e. All three groups, a-b, c-d, e-f-g (GUKs 1-3) are differentiated by their tissue distribution, thermostability, and molecular masses [42, 75]. There are three separate gene loci coding for these isozymes. GUK1 codes for the e, f and g isozymes [38]. Up until now, only GUK1 (isoform 1 of human and mouse guanylate kinases) has been cloned [36], and most of the data concerning guanylate kinases have been derived from studies of isoform1. Human guanylate kinase (hGMPK, GUK1, GMK, ATP:GMP phosphotransferase, EC 2.7.4.8, UniProt Q16774) is an essential enzyme involved in guanine nucleotide metabolism [73]. The hGMPK was identified and partially purified from erythrocytes in 1971 [73]. Later, the cDNA sequence was determined, and the chromosomal localization was refined, 1q31-1q42 [38]. Its primary structure consists of 197 amino acids with a molecular mass of 21,725 Da as computed from the sequence [36]. It has one cysteine residue but no tryptophan and is a monomeric protein. Like other GMPKs, it has the N-terminal nucleotide binding pattern of GXXGXGK (11-18 aa, GPSGAGKS). The apparent K_m values for GMP, dGMP and acyclo-GMP measured were 22 μM, 72 μM and 330 μM, respectively [74].

The guanylate kinase domain (GK^{dom}) of guanylate kinase enzymes (GK^{enz}) is also found in the membrane-associated guanylate kinase (MAGUK) family of proteins. MAGUKs are scaffolding proteins which organize protein complexes at cell or synaptic junctions [76] and play an important role in cell signaling, regulation of synaptic structure, and function in mediating

specific interactions, and mitotic spindle orientation [40, 77, 79]. All MAGUKs have a multidomain structural architecture which consists of one or several PDZ domains (except CACNB, calcium channel β subunit), one Src homology 3 (SH3) domain (except for MAGI, membrane-associated guanylate kinase inverted) and a guanylate kinase domain (GK^{dom}) [76]. The GK^{enz} is widely distributed throughout evolution from bacteria to animals and was evolved into the non-enzymatic GK^{dom} of MAGUKs which is restricted to only choanoflagellates and animals [76, 78]. The GK^{dom} of MAGUKs and GK^{enz} share high sequence and structure similarities, for instance hDlg (human discs large homolog, Uniprot Q12959) has sequence similarities with human GMPK (53%), mouse GMPK (54%), yeast GMPK (54%) and *E.coli* GMPK (47%). Nevertheless, their function is different as GK^{enz} catalyzes the phosphorylation of (d)GMP into (d)GDP using ATP as a phosphate donor while the MAGUK GK^{dom} has no phosphotransferase activity, but functions as a protein interaction domain. The MAGUK GK^{dom} potentially binds GMP and ATP with varying affinities and may have a regulatory role by inducing different conformations between the nucleotide-free and the nucleotide-bound forms [77, 80]. Besides many similarities, there are also critical sequence differences between GK^{enz} and GK^{dom} which may have given rise to their functional divergence. However, the mechanism by which this transition from a nucleotide kinase enzyme to a protein-binding domain occurred is not clear yet. One explanation was provided by introducing the point mutation S35P in yeast GMPK that converted the yeast GK^{enz} into a phosphoprotein binding GK^{dom} with spindle orientation activity of MAGUK as tested in an *in vitro* pull-down assay [40]. Although the P35 mutant had lost kinase activity, it could still bind ATP and GMP. It was suggested that protein dynamics may play a role in switching between catalysis and spindle orientation function. The S35P mutant in yeast GMPK had a similar apo-form like the wild-type GMPK with a large cleft in between the ATP-binding LID and NMP-binding regions (GBD) as shown in (Fig. 1.8). Nevertheless, fluorescence quenching and NMR experiments confirmed that the S35P mutant although may bind GMP but failed to induce conformational closing upon binding of GMP [40]. As proline is least flexible in terms of sterically allowed conformations, therefore, its presence in the hinge part between CORE and GBD regions may hinder the closing movement of the two regions. This suggests that inhibition of substrate-induced guanylate kinase closing was sufficient for functional conversion of GK^{enz} into GK^{dom}.
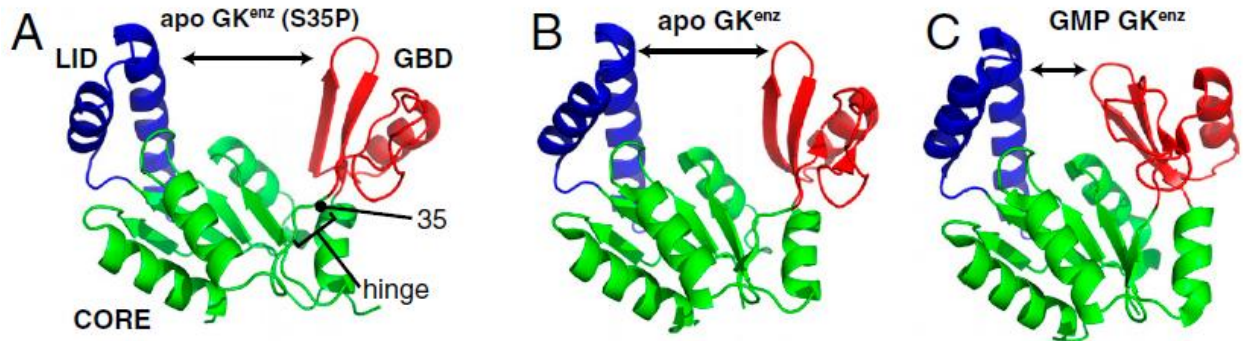
Fig. 1.8. Structure of the yeast guanylate kinase serine to proline mutant (S35P). (A) Structure of yeast guanylate kinase (GK^{enz}) with the serine-to-proline mutation that confers spindle-orienting function. LID (*blue*), CORE (*green*), and GBD (*red*, NMP-binding region) regions are shown. The mutated residue is marked by its sequence number "35". The hinge denotes residues that undergo large dihedral angle changes upon GMP-induced closing [9]. The arrow shows the extent of cleft opening between LID and GBD regions. (B) Structure of wild-type apo GK^{enz} (PDB 1EX6) [10]. (C) Structure of GMP-bound GK^{enz} (PDB 1EX7) [10]. Note the proximity of the LID (*blue*) and GBD (*red*) domains compared with B, demonstrating the large GMP-induced conformational changes in the enzyme GK fold [40].

As stated in section 1.1.1.4.1, hGMPK is physiologically and medically the most relevant enzyme as compared to GMPKs of non-human origin, therefore studying its structural properties and conformational behavior in the presence of different nucleotides is crucial to explain its mechanism of action. It will help us to rationally design mutants and study their kinetics to understand its structure-function relationship.

1.1.1.4.3 Substrate-induced conformational changes in hGMPK

Substrate binding to an enzyme may cause an appreciable amount of changes in the protein structure bringing the catalytic groups into proper orientation for catalysis [1]. Substrate-induced conformational changes suggested by the induced-fit theory have been supported by the findings from numerous studies making use of different techniques such as X-ray crystallography, NMR, and small angle X-ray scattering (SAXS) [2, 3, 23]. Kinases, in particular, have been reported to undergo large movements during catalysis [4], to shield their active center from water in order to avoid ATP hydrolysis.

The active centers of ATP:NMP phosphotransferases (nucleoside monophosphate kinases; NMP kinases; NMPKs) are assembled with large domain movements upon binding of both substrates [5]. The bi-substrate NMP kinases can exist in at least four forms: unliganded form (apo-form), an NMP-bound form, an ATP-bound form, and a ternary complex (closed form). Most of the information on the mechanism of domain movements of an NMP kinase can be extracted by comparison of its open and closed forms. Substrate-induced conformational changes were first identified by comparing the various forms of homologous adenylate kinases (AMP kinases; AKs) [6, 7]. It was observed that both ATP and AMP induce substantial conformational changes upon binding to the enzymes. Binding of AMP results in the closure of the NMP-binding domain, whereas binding of ATP causes the closure of the LID domain [8]. Analysis of 17 crystal structures from the NMP kinase family confirmed the existence of large conformational changes which were mainly attributed to rigid body movements of a LID region and an NMP-binding region with respect to a CORE region [9]. Although comparison of the structures of homologous proteins is informative, precise analysis of the domain movements is complicated by considerable sequence differences, including truncations, substitutions, insertions, and deletions [8]. Therefore, it is more reliable to compare structures of the same enzyme in different ligand-complexed forms. In case of yeast GMPKs, large movements of GMP-binding domain and smaller but significant movements of the LID domain have been reported by comparing $yGMPK_{apo}$ and $yGMPK_{GMP}$. The conformational state of the apo-form of yGMPK is much more open than the substrate-bound closed form [10]. The mouse GMPK structure was previously determined in a fully closed form only, $mGMPK_{GMP.ADP}$ [11]. Like other nucleoside kinases (human deoxycytidine kinase) and NMP kinases (*E.coli*, yeast and human thymidylate kinases) that we structurally and functionally characterized, GMPKs share three distinctive structural parts termed NMP-binding region, LID region and CORE region connected by hinges. The substrate-induced conformational changes can be described by the movement of these regions making the molecule either more compact or extended depending on the presence or absence of substrates, respectively.

Despite several important catalytic roles, no structural data is available for hGMPK. In addition, no mammalian GMPK structures have been determined in all four forms (apo-form, NMP-bound form, ATP-bound form, and ternary complex) that would reflect transitions between various conformational states. Therefore, we embarked on studying hGMPK structures in all four forms

by SAXS to unravel the conformational and dynamic behavior of this medically most relevant enzyme [20-22]. These conformational changes will help to delineate the catalytic reaction trajectory of hGMPK and will provide a basis for determining its high resolution structure. We evaluated our *ab initio* SAXS models by comparison with the available mGMPK crystal structure. As hGMPK is highly identical (88%) to mGMPK at the amino acid sequence level, a reliable homology model was constructed using the mGMPK structure as a template.

1.1.1.4.4 Human GMPK as a critical enzyme for phosphorylation of thiopurines

The FDA-approved thiopurine drugs 6-thioguanine (6-TG), 6-mercaptopurine (6-MP), and azathioprine (Aza) have been used as anti-cancer agents, as immunosuppressants, and in the treatment of inflammatory diseases. The metabolic activation of thiopurine prodrugs involves the action of several cellular enzymes converting thiopurines to 6-thioguanosine monophosphate (6-TGMP). It is phosphorylated in a rate-limiting step catalyzed by hGMPK to form 6-thioguanosine diphosphate (6-TGDP) which is further phosphorylated to thio-GTP and thio-dGTP by NDPK and reductase enzyme. Both nucleotides are substrates for incorporation of 6-TG into RNA and DNA. Once integrated into DNA, a fraction of DNA 6-TG undergoes non-enzymatic methylation (6-meTG) probably by S-adenosyl-L-methionine. During replication, 6-TG and 6-meTG base pair with thymine rather than cytosine. The aberrant 6-TG:T and 6-meTG:T base pairs, which escape proofreading by DNA polymerases, activate the DNA mismatch repair (MMR) system. Incomplete processing due to excessive accumulation of incorrect base pairs ultimately leads to cell death [16-19, 31, 106, 107, 109, 110, 129, 254]. In addition to the MMR pathway, thiopurine treatment causes reduced expression of proteins in the electron transport chain of mitochondrial respiratory complex inducing mitochondrial dysfunction and elevated generation of oxidatively induced DNA lesions that may further contribute to cytotoxicity [19]. Thiopurine metabolic activation is shown below in Fig. 1.9.
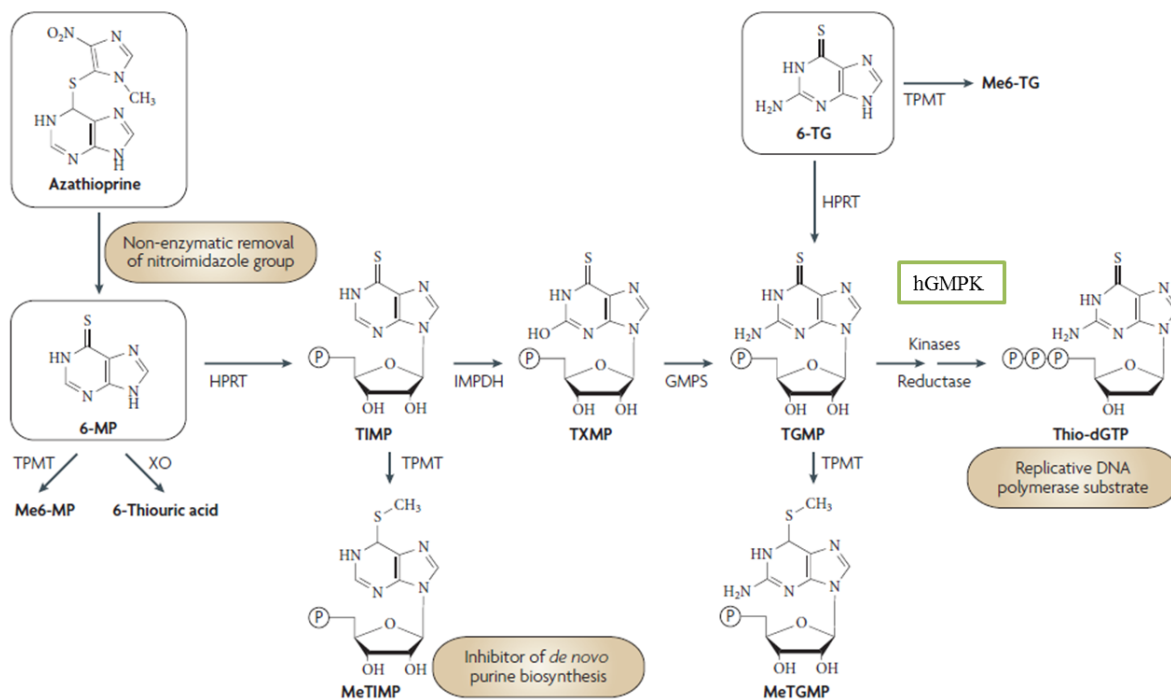
Fig. 1.9. Thiopurines and their metabolism. Azathioprine is converted to 6-mercaptopurine (6-MP) by non-enzymatic activation. Hypoxanthine–guanine phosphoribosyltransferase (HPRT) catalyzes the conversion of 6-MP and 6-TG to their respective nucleoside monophosphates (TIMP and TGMP). Both free thiopurines (6-MP and 6-TG) and their monophosphates (TIMP and TGMP) are substrates for thiopurine S-methyltransferase (TPMT) which methylates them to Me6-MP, Me6-TG, MeTIMP and MeTGMP. In addition, xanthine oxidase (XO) converts 6-MP to 6-thiouric acid. Unlike MeTGMP, MeTIMP is an effective inhibitor of *de novo* purine biosynthesis. TIMP is metabolized to TGMP by successive action of inosine monophosphate dehydrogenase (IMPDH) and guanine monophosphate synthetase (GMPS). TGMP is phosphorylated by guanylate kinase (hGMPK) to TGDP, and further action by reductase and nucleoside diphosphate kinase to form thio-dGTP for incorporation into DNA [16-18].

It was demonstrated by *in vitro* experiments that certain point mutations at S37 in mGMPK (analogous to S37 in hGMPK) can cause resistance to phosphorylation of 6-TGMP to 6-TGDP which is a critical step in the metabolic activation of thiopurine prodrugs [16]. The wild-type mGMPK and its three mutants S37A, S37T and S37Y were recombinantly produced and their activities were determined by using the NADH-dependent enzyme coupled assay [16]. Only the wild-type mGMPK phosphorylated 6-TGMP to 6-TGDP although with catalytic efficiency ~8,000 times lower than that for the natural substrate GMP [16]. All three mutants did not show

any activity for the drug. It was suggested that the steric hindrance between active site T83, bound GMP or 6-TGMP and the presence of certain residues at position 37 such as S37T and S37Y plays an important role in substrate-binding and catalysis. Similarly to the mutant S37A, the substrate may not be properly positioned due to inefficient side chain interactions. Moreover, the growth of wild-type and three mutants expressing conditional GMPK-deficient *E.coli* strain (TS202A(DE3)) was tested on anti-leukemic 6-TG containing plates and broth cultures. The conditional GMPK-deficient *E.coli* strain requires the presence of a functional, plasmid-borne guanylate kinase for growth under selective conditions. It was observed that the wild-type mGMPK expressing cells were non-viable due to utilization of the drug, however all three mutants grew well indicating lack of drug activation by the GMPK mutants [16, 108].

Due to very weak activity of GMPK for 6-TGMP, it will be of great relevance to engineer hGMPK for enhanced activity against the drug, such that it could be used as a suicide gene in cancer cell lines. Nevertheless, preliminary experiments for testing the effects of wild-type hGMPK in combination with 6-TG prodrug in mammalian cells such as HEK293 will provide a useful basis for advanced experiments with engineered hGMPK mutants.

1.1.1.4.5 Electrochemical detection of guanosine monophosphate with a quantum dot-based biosensor modified with human GMPK

Quantum dots (QDs) are semiconductor nanocrystals usually composed of atoms from groups II–VI, III–V, or IV–VI of the periodic table [290-292]. QDs have small sizes of a few nanometers and have very high surface-to-volume ratios enabling better interactions between these surface atoms and the surrounding molecules. They have many advantages over organic fluorophores such as increased photostability, narrow emission bands and high brightness [293, 294]. Due to their unique optical and chemical properties they have been used in biological macromolecule sensors, organic small molecule sensors, inorganic ion sensors, biological labeling, cell labeling, animal imaging and therapy [295].

QDs have been investigated in the construction of light-addressable electrochemical sensors [284-289]. When light is illuminated on QDs surface, electron hole pairs are generated. The excited conduction band electrons can be transferred to an electrode or to an electron acceptor in solution. Electrons can also be transferred from an electrode or solubilized electron donor to valence band holes in QDs. An anodic or cathodic current can be generated depending on the

applied potential. The former is caused by the electron transfer from the conduction band of QDs to the electrode, and the latter one is produced by the electron transfer from the electrode to the valence band of QDs. In this way, a QDs layer on electrode surface facilitates the direct electron transfer between the enzyme redox center and the electrode. The redox reaction of QDs surface can be switched on and off by light [87, 88, 112, 113, 284, 296-300]. Different enzymes have been combined with QD-modified electrodes for making biosensors. Riedel et al [89] reported a photobioelectrochemical sensor for the detection of sarcosine with sarcosine oxidase (SOD) immobilized on QDs. Moreover, Khalid et al [87] presented an electrochemical sensor for *p*-aminophenyl phosphate (*p*APP) based on the electrochemical conversion of 4-aminophenol (4AP) to 4-quinoneimine (4QI) at QDs surface under light illumination. *p*APP was converted to 4AP catalyzed by alkaline phosphatase. As NADH is involved in many enzymatic reactions as co-substrate or co-product, it can be used for the indirect detection of the respective enzyme substrate. The electrochemical detection of NADH was reported by Schubert et al [88] using CdSe/ZnS QDs. They found that the change in photocurrent was proportional to the concentration of glucose utilized and to the production of NADH catalyzed by glucose dehydrogenase (GDH). NADH was detected in the concentration range of 20 μM to 2 mM.

Here, we report on the use of QDs for the enzymatic detection of guanosine monophosphate (GMP) with hGMPK coupled to the utilization of NADH in an enzyme-coupled assay [99]. The products of the hGMPK-catalyzed reaction were coupled to the reactions catalyzed by pyruvate kinase (PK) and lactate dehydrogenase (LDH) as helper enzymes. During this process, LDH oxidized NADH (nicotinamide adenine dinucleotide, reduced form) to NAD^+ in proportion to the hGMPK activity. Therefore, we borrowed the redox reaction of NADH for the indirect detection of GMP by hGMPK. The enzymatic reaction was detected with the QD-modified electrode [87, 88]. In most cases, enzyme-catalyzed reactions are monitored using the enzyme protein in free form on the surface of electrode [88, 284]. However, we immobilized hGMPK on a QD-modified electrode with polyelectrolyte multilayers using layer-by-layer assembly [87, 89]. The approach which we adapted for hGMPK can equally be applied to other therapeutically important nucleoside and nucleotide kinases.

1.1.1.4.6 A novel spectrophotometric and fluorometric enzyme-coupled assay for human GMPK

Thiopurine prodrugs including 6-thioguanine (6-TG), 6-mercaptopurine (6-MP), and azathioprine have been widely used as antileukemic and immunosuppressive agents that are activated by cellular enzymes most importantly by the rate limiting hGMPK [17, 19, 310, 311]. 6-thioguanosine monophosphate (6-TGMP) can be considered as the active metabolite of all thiopurines which is successively phosphorylated to 6-TGDP and TGTP by hGMPK and nucleoside diphosphate kinase, respectively, for incorporation into RNA and DNA [16, 18]. Unlike the canonical bases of RNA and DNA which all absorb light maximally in the UVC spectral region (~260 nm), thiopurines in the form of 6-TG and 6-TGMP are UVA chromophores having an absorbance maximum at approximately 340 nm (Fig. 1.10) [18]. The most commonly used assay for studying the steady-state kinetics of hGMPK and other nucleoside and nucleotide kinases is the NADH-dependent spectroscopic assay [84, 99, 100]. This assay is based on the maximum absorbance of NADH at 340 nm which unfortunately overlaps with that of the 6-TGMP. This spectrophotometric artifact was addressed by Miller et al [301], and they minimized the error associated with the assay by using a blue filter in the light path and monitored the absorbance at 373 nm rather than 340 nm, though it drastically reduced the sensitivity of the assay due to poor absorbance of NADH at 373 nm. Similarly, a number of radioisotopic assays have been reported for nucleoside kinases which are based on using radioactive substrates (nucleosides) and are mostly time consuming [302-304]. Such assays are not appropriate for *in vitro* evolution of hGMPK that aims to improve its catalytic efficiency against nucleoside analogs such as thiopurines. In addition, the limited sensitivity of absorption-based spectroscopic assays [99, 100] could become a major problem when dealing with reduced volumes. For instance, high-throughput screening for directed evolution of proteins by droplet-based microfluidics relies on reaction volumes of nanoliters and femtoliters [116, 317, 318].

In order to solve these problems, we developed a novel assay that can be used both in absorbance mode as well as in fluorescence mode at higher wavelengths than 340 nm with no background absorbance from the nucleotide analogs. In this four step assay (see section 2.2.8), the activity of hGMPK is determined by coupling the formation of nucleoside diphosphate (NDPs) products to the reactions catalyzed by pyruvate kinase (PK) and pyruvate oxidase (lpPOX) in the presence of phosphoenolpyruvate (PEP) and pyruvate, respectively. Hydrogen peroxide, which is produced

by the action of lpPOX [305-309], reacts with Amplex Red in the presence of horseradish peroxidase (HRP) to produce the red-fluorescent oxidation product, resorufin [135]. Resorufin has excitation and emission maxima of about 568 nm and 584 nm, respectively, and because of its high extinction coefficient (5.4 x 10^4 $M^{-1}cm^{-1}$), the assay can be performed spectrophotometrically and fluorometrically [116].
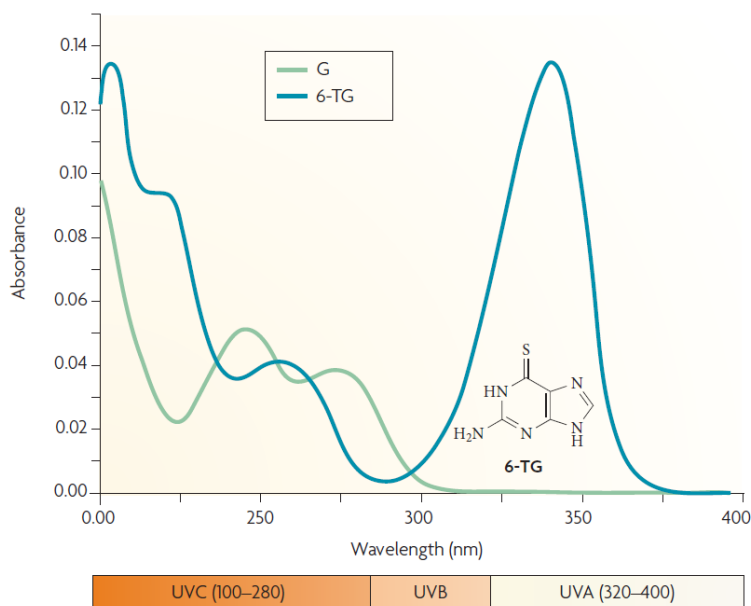
Fig. 1.10. Absorbance spectra of 6-TG and G. Canonical purines such as guanine (G) have no significant absorption at wavelengths longer than 300 nm. Substitution of 6-O atom by sulphur (e.g. 6-TG and 6TGMP) shifts the absorbance maximum into the UVA region (~340 nm) [18].

1.1.1.4.7 Human GMPK-catalyzed reactions in polyelectrolyte containers of various shapes and sizes

A promising strategy to enhance the targeting efficacy in drug delivery systems is to mimic and manipulate the way molecules interact with each other that may enable us to design carrier structures of desired properties like with different shapes, sizes and surface modifications [255, 256]. To date, spherical microparticles have been used extensively in research, and only few methods are available for controlled synthesis of non-spherical microparticles. Nevertheless, there is growing interest in developing non-spherical carriers with elongated or filamentous

morphologies [260] or anisotropic delivery systems [261] because conventional isotropic spherical drug carriers cannot satisfy the current demands [260]. In addition, non-spherical carriers are preferred for their distinctive properties such as anisotropic responses to external fields, large surface areas, and unique structure formation. Examples of anisotropic carriers include particles of various shapes such as ellipsoidal, rhomboidal, stars, cubes, rods, discs and triangles. They could be used to mimic anisotropic cells, to study self-assembly and packing of anisotropic colloids, and encapsulating macromolecules such as therapeutic enzymes to study their reactions and other properties [86, 257-259]. Moreover, the number of studies in the literature continually increases which report about the alteration of biological responses if particles with the same composition but different geometry enter the bio-distribution pathways [262-265]. Particle shape plays therefore a major role during cellular uptake of drug carriers [266].

Among the other popular drug carriers, e.g. nano- and microparticles, liposomes and red blood cells, polymeric capsules fabricated through the layer-by-layer (LbL) technique have attracted particular attention in the last decades because they offer a high degree of multifunctionality in drug delivery applications [267, 268]. The process of LbL assembly is based on the consecutive adsorption of oppositely charged polyelectrolytes around a preformed charged template core, which is chemically dissolved after the coating process. This technology enables the formation of shell-in-shell capsules, where multicompartments can be loaded with different functional biomolecules such as therapeutic enzymes and drugs [86, 269, 283]. Medically and biologically relevant macromolecules can be introduced into LbL films via non-covalent interactions under physiological conditions. The biological properties of the encapsulated macromolecules do not change significantly during loading. In case of loaded enzymes, their catalytic properties and structural stability may be maintained both *in vitro* and *in vivo* experiments. Microcapsules provide the possibility to study enzymatic reactions in confined volumes. Moreover, these capsules are not cytotoxic and can be made biodegradable, which makes them suitable candidates for cellular targeting with the loaded cargo proteins. As the LbL technology combines the versatility in manipulating the composition, surface chemistry, and dimensions of nanostructured LbL thin films with the easy functionalization of diverse therapeutics and/or biomolecules, it provides a powerful tool for the nanoscale synthesis of novel drug-delivery systems [86, 267-270].

Due to their highly porous structure, bio-friendliness and ease of fabrication, $CaCO_3$ particles are commonly used templates for LbL-engineered capsules in drug delivery applications [267]. The $CaCO_3$ cores can easily be dissolved with EDTA after deposition of multilayer polyelectrolyte shells to produce hollow microcapsules with entrapped macromolecules. The shape of commonly used $CaCO_3$ templates is spherical and their crystal structure usually is vaterite [271]. Among three polymorphs (calcite, aragonite, or vaterite) of $CaCO_3$, vaterite is thermodynamically the least stable one [272]. However, for drug delivery applications, it is very attractive since vaterite molecules have high porosity and therefore can take up many molecules into their interior structure [270, 273, 274]. In other application areas such as biomineralization or nanomaterial processing, $CaCO_3$ particles are utilized in a diverse range of shapes because their crystallization mechanism can be easily altered using organic additives [275-279]. If additives such as dipeptides [275], polypeptides [276], peptide type block copolymers [277] or inorganic binding peptides [278, 279] are added to the precipitation process of $CaCO_3$, the shape and size of $CaCO_3$ particles can be varied from nanorods to mesocrystals. On the other hand, organic additives may cause unpredictable cytotoxic effects when they are used in drug delivery applications.

In this study, we present anisotropic LbL containers templated on non-spherical vaterite $CaCO_3$ particles, which do not contain any organic additive. To obtain stable vaterite templates, the working pH range was chosen as 9.0-9.5 [280]. The size and shape of vaterite particles was controlled as a function of reaction time, the concentration of initial salt solutions, solubility of salts and mixing speed [281, 282]. Calcium carbonate microparticles in ellipsoidal, rhomboidal and spherical geometries were produced, by changing the concentration of initial salt solutions and the solubility of salts at constant pH, mixing speed and reaction time. Implications of the influence of initial salt solution concentration and the solubility of salts on the $CaCO_3$ precipitation kinetics are discussed in detail. These calcium carbonate templates were further used for capsules fabrication via LbL-assembly by using conventional polyelectrolytes. Capsules were loaded with hGMPK, and its catalytic activity was determined using the NADH-dependent spectroscopic assay. Our results have revealed that polyelectrolyte containers with different geometries exhibit promising applicability in drug delivery applications due to their high loading capacity, biocompatibility, and easy fabrication and handling.

1.2 General characteristics of deoxyribonucleoside kinases

Deoxyribonucleoside kinases (dNKs) are key enzymes that catalyze the first step in the nucleoside salvage pathway by the transfer of a phosphoryl group from ATP (or other phosphate donor) to deoxyribonucleosides to form deoxyribonucleoside monophosphates (dNMPs) [193]. The dNMPs are finally converted to deoxyribonucleoside triphosphates (dNTPs) by the action of two other kinases called nucleoside monophosphate kinases (NMPKs) and nucleoside diphosphate kinases (NDPKs) [45]. As there are no carrier proteins in the cell membrane for the transport of (deoxy)ribonucleotides, they have to be synthesized within the cells either *de novo* or via the nucleoside salvage pathway [44, 45, 194]. In the *de novo* pathway, small molecules like amino acids, ribose-5´-phosphate, and CO_2 are utilized for the synthesis of ribonucleotides. Deoxyribonucleotides are synthesized *de novo* at the diphosphate level by the reduction of NDP to the corresponding dNDP catalyzed by ribonucleotide reductase [193, 195-198]. However, the nucleosides or deoxyribonucleosides originating from food or degraded cells are transferred across the cell membrane through nucleoside transport proteins and are phosphorylated via the nucleoside salvage pathway [194]. In the salvage pathway, dNKs are usually rate-determining enzymes which add the first phosphate group to deoxyribonucleosides. Once phosphorylated, the deoxyribonucleotides are trapped intracellularly due to their negative charges [193]. In humans, there are four dNKs: the two cytoplasmic dNKs are thymidine kinase 1 (TK1, ATP:thymidine 5´-phosphotransferase, EC 2.7.1.21) and deoxycytidine kinase (dCK, NTP:deoxycytidine 5´-phosphotransferase, EC 2.7.1.74), and the two mitochondrial dNKs are thymidine kinase 2 (TK2, ATP:thymidine 5´-phosphotransferase, EC 2.7.1.21) and deoxyguanosine kinase (dGK, nucleoside triphosphate: deoxyguanosine 5´-phosphotransferase, EC 2.7.1.113) [151]. The four enzymes have distinct but overlapping specificities [46]. The nomenclature of dNKs is based on their preferred substrates. TK1 phosphorylates deoxythymidine (dThd) and deoxyuridine (dUrd) into their monophosphate forms using ATP as phosphate donor, however, it also accepts other phosphate donors [199, 200]. dCK phosphorylates the pyrimidine deoxycytidine (dCyd), and the purines deoxyadenosine (dAdo) and deoxyguanosine (dGuo) [84]. It uses several nucleotides as phosphate donors, but UTP has been reported to be a better phosphoryl group donor than ATP [201-203]. TK2 catalyzes the transfer of a phosphate group to the 5´-OH of dThd, dUrd, and dCyd using ATP as a phosphate donor [139]. dGK phosphorylates dGuo, dAdo, and dIno (deoxyinosine) [188, 204]. It uses several phosphate donors, but UTP is the preferred phosphoryl

group donor [205]. The four dNKs can be divided into two families based on their amino acid sequence similarities. The first family contains dCK, and the mitochondrial enzymes dGK and TK2 that share more than 40% sequence identities [46]. The *Drosophila melanogaster* deoxyribonucleoside kinase (Dm-dNK, nucleoside triphosphate: deoxyribonucleoside 5´-phosphotransferease, EC 2.7.1.145) which uses all four nucleosides as substrates, and the viral HSV1-TK can be grouped in this family of dNKs [46, 206]. The sequence alignment of dCK, dGK, TK2, and Dm-dNK are shown below in Fig. 1.11. These enzymes are homodimers and have similar three-dimensional folds. The other family contains the cytosolic TK1 which is a homotetramer [207]. TK1 shows little sequence similarity to other dNKs, and is related to thymidine kinases from *E.coli* and poxviruses [214, 215].
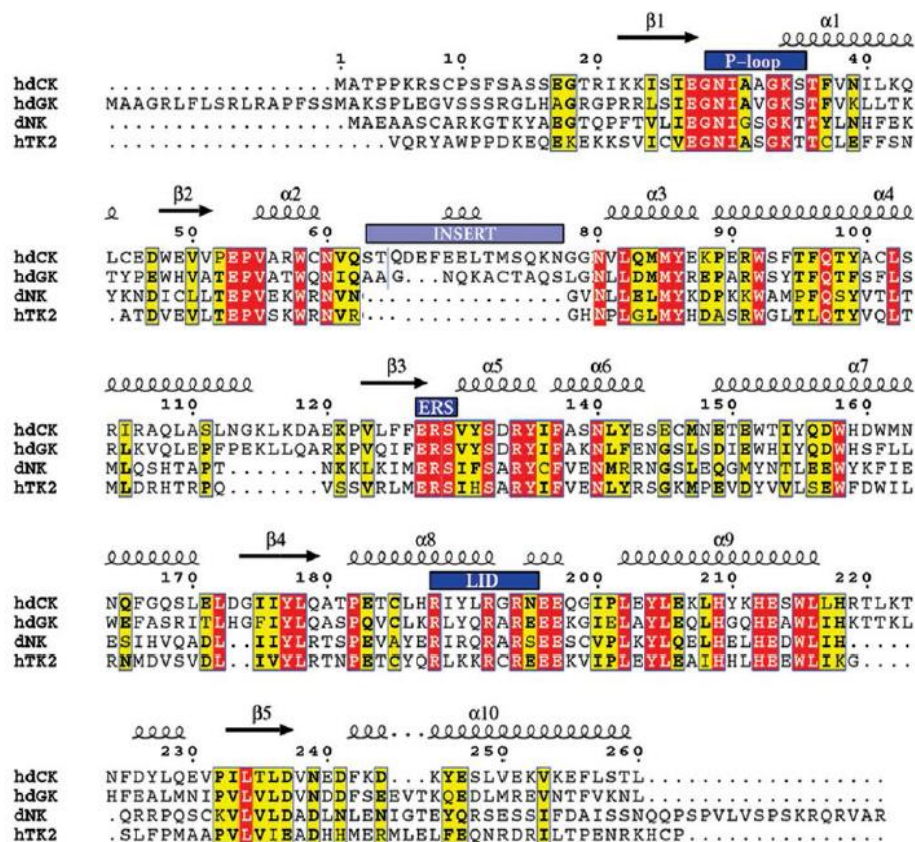
Fig. 1.11. Sequence alignment of deoxyribonucleoside kinases. All four deoxyribonucleoside kinases have a conserved P-loop, the ERS motif and a LID region (boxed in blue). The conserved residues are shown red; and similar residues are yellow colored [208]. The abbreviations used are: hdCK, human deoxycytidine kinase; hdGK, human deoxyguanosine kinase; dNK, *Drosophila melanogaster* deoxyribonucleoside kinase; hTK2, human mitochondrial thymidine kinase.

Crystal structures of all human dNKs are known except TK2. The three-dimensional structure of dCK (PDB 2NOA) is shown below in Fig. 1.12 [46]. Each monomer has an α/β architecture which consists of five-stranded parallel β-sheet surrounded by ten α-helices [208]. There are three conserved motifs in dNKs: the P-loop, ERS motif, and LID region. The P-loop binds the phosphate of ATP, glutamate (E) of ERS contributes to Mg^{2+} binding, and arginine (R) plays a role in catalysis, whereas the LID region closes on the bound phosphoryl group donor and provides residues for ATP binding and catalysis [46]. Besides natural substrates, dNKs bind and phosphorylate nucleoside analogs. The common feature is the interaction of a conserved glutamate-arginine pair (E53 and R128 in dCK) with the 5´-OH of a nucleoside. In addition, the conserved tyrosine-glutamate pair interacts with the 3´-OH (Y86 and E197 in dCK) [208].
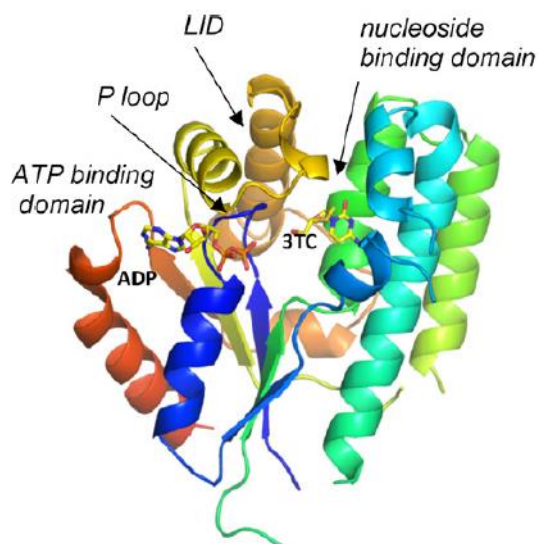
Fig. 1.12. Human deoxycytidine kinase. The ribbon diagram of human dCK (PDB 2NOA) monomer is shown with bound 3TC (2´-deoxy-3´-thiacytidine, lamivudine) at the nucleoside acceptor site on the right and ADP at the nucleotide donor site on the left of the central cleft at the active site [46].

The mechanism of action of dNKs involves the activation of 5´-OH of the deoxyribonucleoside which facilitates the nucleophilic attack on the γ-phosphate of the phosphate donor as shown below in Fig. 1.13. The glutamate carboxyl group close to the 5´-OH at the active site of HSV1-TK was suggested to act as a base in the reaction that deprotonates the 5´-OH [151, 216]. The arginine residues in the close proximity help in the deprotonation. The equivalent glutamate and

arginine residues in other dNKs support this mechanism. dNKs require Mg^{2+} for their activity that interacts with the β- and γ-phosphates of the phosphate donor [69, 151]. dNKs are relatively poor catalysts. For instance, human dCK which has broad substrate specificity for purine and pyrimidine deoxyribonucleosides has a catalytic efficiency of only 4.8×10^3 M^{-1} s^{-1} for dCyd with a low K_m of 6.2 μM, and a turnover rate of 0.03 s^{-1} [208]. dCK also phosphorylates dAdo and dGuo with a lower catalytic efficiency than dCyd. The catalytic efficiency of UMP-CMPK (uridylate/cytidylate kinase), which belongs to the family of NMP kinases and catalyzes the second phosphorylation of dCMP, is about 10-fold higher, i.e. 7×10^4 M^{-1} s^{-1} [46]. dNKs are inhibited by their end products (dNTPs) which act as competitive inhibitors for binding at the enzyme active site. Crystal structures of several dNKs have been solved as complexes with their respective dNTPs [46, 210].
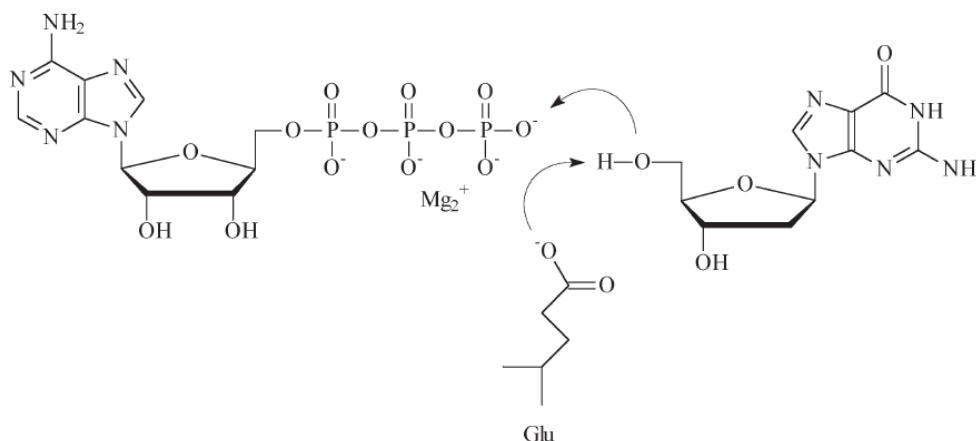
Fig. 1.13. Mechanism of action of deoxyribonucleoside kinases. ATP as a phosphate donor is shown to the left and the phosphate acceptor dGuo to the right. It is suggested that glutamate (Glu) acts as a general base and Mg^{2+} as a counter ion [151].

dNKs are peculiar enzymes for their low enantioselectivity [211]. For example, dCK phosphorylates β-D-2'-deoxycytidine (D-dCyd) which is its natural substrate, and β-L-2'-deoxycytidine (L-dCyd), its enantiomer, with the same efficiency [212]. Similarly, the human mitochondrial TK2 accepts both D and L enantiomers of dThd with almost equal efficiency [46]. It accepts L-dCyd, L-BVDU ((E)-5-(2-bromovinyl)-2'-deoxyuridine), and L-dUrd as substrates [199]. The dGK phosphorylates the L-dAdo more efficiently than D-dAdo and D-dGuo [213]. In

comparison to mitochondrial dNKs, the cytosolic human TK1 displays strict enantioselectivity and is limited to the D-enantiomers of thymidine and its analogs [46].
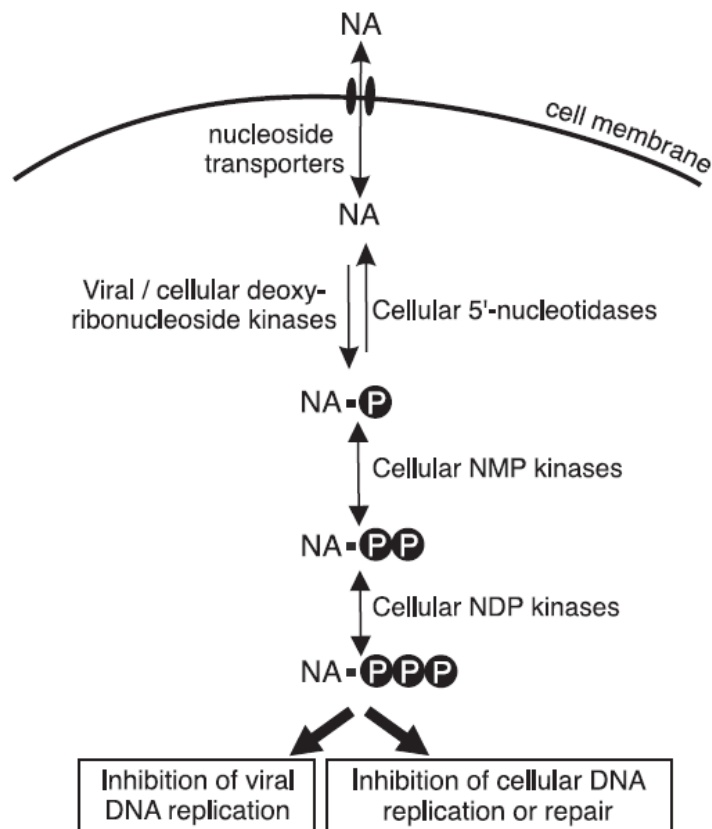
Fig. 1.14. Metabolic activation of nucleoside analogs (NA). The anticancer and antiviral nucleoside analogs require intracellular phosphorylation for gaining pharmacological activities. NAs are transported across the cell membrane, and are phosphorylated by cellular or viral kinases to their triphosphate forms for incorporation into DNA [45].

The chemotherapeutic treatment of cancer and viral diseases depends largely on the action of dNKs that catalyze the first step in the conversion of innocuous nucleoside analogs (NAs) to their corresponding cytotoxic nucleoside triphosphates which are incorporated into cellular or viral DNA by DNA polymerase or a viral reverse transcriptase, as shown schematically above in Fig. 1.14. The presence of NAs in DNA causes stalling or premature termination of replication forks, and also offers resistance to proofreading exonucleases. Some of the NAs act by inhibiting the key enzymes (e.g., ribonucleotide reductase, thymidylate synthase, or dCMP deaminase)

involved in the generation of purine and pyrimidine nucleotides, RNA synthesis, and directly activate the caspase cascade. The overall effect is the inhibition of the DNA synthetic machinery and initiation of apoptosis [45, 217, 218]. Some of the clinically used anticancer NAs are 2-chloro-2´-deoxyadenosine (Cladribine, CdA), 2-fluoro-9-β-D-arabinofuranosyladenine (Fludarabine, F-AraA), 1-β-D-arabinofuranosylcytosine (Cytarabine, Ara-C), and 2´,2´-difluorodeoxycytidine (Gemcitabine, dFdC). The antiviral NAs comprise acyclovir, ganciclovir, 2´,3´-dideoxyinosine (Didanosine, ddI), 2´,3´-dideoxycytidine (Zalcitabine, ddC), 2´-deoxy-3´-thiacytidine (Lamivudine, 3TC), 3´-azido-2´,3´-dideoxythymidine (Zidovudine, AZT), BVDU (bromovinyldeoxyuridine or brivudin), 2´,3´-didehydro-3´-deoxythymidine (Stavudine, d4T), and abacavir (ABC) [45, 217]. Some of the structures of NAs are shown below in Fig. 1.15.
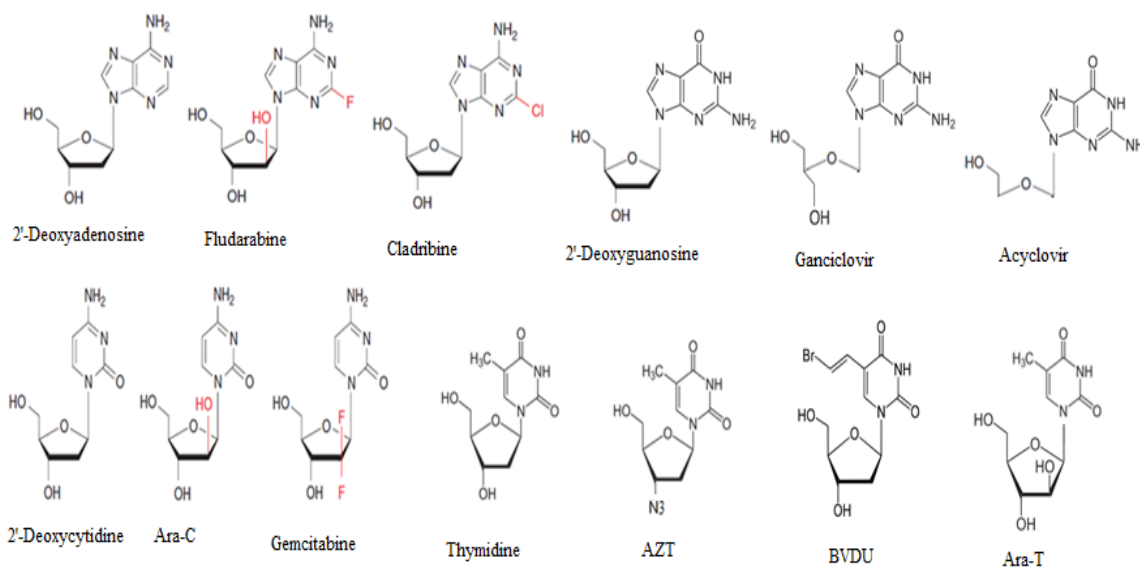
Fig. 1.15. Structures of purine and pyrimidine deoxyribonucleosides and their respective analogs. Abbreviations used: Ara-C, 1-β-D-arabinofuranosylcytosine or cytarabine; AZT, 3´-azido-2´,3´-dideoxythymidine or zidovudine; BVDU, bromovinyldeoxyuridine or brivudin; Ara-T, 1-β-D-arabinofuranosylthymidine.

Besides the essential role of dNKs in cancer and viral chemotherapy, deficiency or lack of dNKs activity leads to severe disorders like mitochondrial DNA depletion syndrome [219]. It is caused

by mutations in a number of genes including dGK and TK2. Out of all four human dNKs, TK2 is the least studied enzyme from its structural perspective. In its family, human TK2 is the only enzyme whose structure is still missing, and thus it is very important to explore its structure-function characterization.

1.2.1 Human mitochondrial thymidine kinase

1.2.1.1 Biological importance

Besides playing an important role in providing dNTPs for mitochondrial DNA replication and maintenance, TK2 activates numerous antiviral and anticancer nucleoside analogs, including analogs of thymidine, deoxyuridine and deoxycytidine. Examples of thymidine analogs are 3´-azido-2´,3´-dideoxythymidine (AZT), ara-T (arabinofuranosyl thymine), FLT (3´-fluoro-2´, 3´-deoxythymidine) and ribothymidine. Deoxyuridine analogs comprise 2´-difluoro-2´-deoxyuridine (dFdU), 5-Fluoro-2'-deoxyuridine (5-FdU), and 1-(2´-deoxy-2´-fluoro-β-D-arabinofuranosyl)-5-iodouracil (FIAU) [140, 152, 157, 187, 199, 231]. The deoxycytidine analogs phosphorylated by TK2 are dFdC (2', 2'-difluorodeoxycytidine, gemcitabine) and ara-C (1-β-D-arabinofuranosylcytosine). Similarly, deoxycytidine analogs with 5-substituition such as 5-(2-chloroethyl) and 5-(2-bromovinyl) are also accepted as substrates [152, 157]. The K_m values determined for AZT and FLT were 4.5 μM and 6.5 μM, respectively, with efficiencies (V_{max}/K_m) 2 and 3% to that of dThd [233]. AZT, 2´,3´-didehydro-2´,3´-dideoxythymidine (d4T), araT and FLT inhibited the phosphorylation of dThd. In contrast, the phosphorylation of dCyd was stimulated by AZT, d4T and FLT [233].

It has been reported that certain mutations in hTK2 cause a heterogeneous group of severe mitochondrial disorders of infancy and childhood, characterized by decreased mitochondrial DNA copy number, called myopathic mitochondrial DNA depletion syndrome (MDS) [229, 237, 239-241]. Patients suffering from myopathic MDS show progressive weakness, hypotonia, and areflexia, and die of respiratory failure before the age of 1 year (congenital form) or before 10 years (juvenile form) [234]. In addition, patients carrying TK2 mutations manifest neurological phenotypes, thus suggesting that TK2 deficiency results in neuronal dysfunction [242, 243]. Mutations which have been found in hTK2 obtained from different patients with MDS include I53M, T108M, Y112N, H121N, R130W, R130Q, R192K, and I212N [229, 234, 235, 238]. The

numbering used here is based on the full-length hTK2 sequence (265 aa, UniProt O00142-1), as TK2 sequences were numbered in the literature with shorter isoform of hTK2.

1.2.1.2 General characteristics

TK2 (ATP:thymidine 5´-phosphotransferase, EC 2.7.1.21) catalyzes the transfer of a γ-phosphate group from ATP (or other phosphate donor) to the 5´-OH group of a deoxyribonucleoside (dThd, dUrd or dCyd) [151, 157, 199]. The substrate specificity of TK2 overlaps with both TK1 by phosphorylating dThd and dUrd, and with dCK by phosphorylating dCyd [158]. TK2 is constitutively expressed in all tissues in proportion to the mitochondrial content of the cell types, and plays a critical role in providing dNTPs for the replication and maintenance of mitochondrial DNA [152, 159, 222]. TK2 contributes a small fraction to cellular TK levels in proliferating cells, however in nonproliferating cells it corresponds to the predominant fraction of TK activity [151]. The human TK2 gene is localized to chromosome 16q22 [153]. The primary structure of hTK2 is 46% identitical to Dm-dNK (*Drosophila melanogaster* deoxyribonucleoside kinase), 36% and 33% to human dCK and dGK, respectively. However, the hTK2 amino acid sequence is quite different from TK1. Previously, submitochondrial fractionation experiments had revealed deoxythymidine kinase activity in mitochondria [155]. TK2 has an N-terminal mitochondrial import signal, and its mitochondrial location is supported by the fact that the full-length mouse TK2 was targeted to mitochondria when tested in mammalian cells [150]. Similarly, the *in vitro* translation and translocation experiments with purified rat mitochondria endorsed that the N-terminal mitochondrial targeting signal in the full-length mouse TK2 (270 aa) directed its import into the mitochondrial matrix [140]. In addition, mutations in TK2 have been associated with the mitochondrial DNA depletion syndrome (MDS) [219].

Johansson and Karlsson [153] reported the cloning of hTK2 cDNA encoding a 234 amino acid protein, and recombinantly produced it in *E.coli*. Similarly, Wang et al [152] cloned a cDNA of hTK2 containing an ORF (open reading frame) of 232 residues and kinetically characterized it. TK2 phosphorylates dThd, dUrd, and dCyd using ATP or CTP as phosphate donors. It shows negative cooperativity for dThd and the anti-HIV drug AZT, but the phosphorylation of dCyd and dUrd follows Michaelis-Menten kinetics. The K_m values determined for dThd, dCyd, dFdC and dFdU were 4, 7, 66, and 29 µM, respectively [152, 199]. Unlike TK1, TK2 has relaxed enantioselectivity and phosphorylates both L- and D-enantiomers, for instance L-dThd and L-

dCyd are as efficiently phosphorylated as their D-enantiomers [46]. Moreover, other L-nucleoside substrates include L-BVDU, L-FMAU (2′-fluoro-5-methyl-β-L-arabinofuranosyluracil), and L-5-iodo-dU [220, 221]. The cDNAs of hTK2 reported were mostly incomplete as they do not contain the N-terminal mitochondrial targeting signal. Wettin et al [150] for the first time cloned a full-length mouse TK2, 270 amino acids in length, which was targeted to mitochondria when expressed as a fusion with C-terminal GFP in cancer cell lines. TK2 is predominantly found in mitochondria, but it may also be present in cytosol [139, 227, 228]. In fact, hTK2 shows multiple transcripts in most tissues, and the two isoforms (mitochondrial and cytosolic) may originate by alternative splicing of the alternative first exons in the hTK2 gene which consists of more than 10 exons [139]. The full-length mouse TK2 and its two N-terminally truncated forms were recombinantly produced in *E.coli*, and it was found that the truncation did not change the K_m values, but the V_{max} values were increased for truncated forms [140].

Barroso et al [139] when cloning and expressing hTK2 (234 aa) as a fusion with MBP (maltose binding protein) in *E.coli* reported that it extensively aggregated with very low recovery of the soluble fraction. Nevertheless, they reduced the nonspecific aggregation of MBP-hTK2 by coexpression of GroEL/ES chaperonins and ethanol supplementation to the growth media. The hTK2 was obtained in several oligomeric forms with the dimer being the dominant form, i.e. dimer > tetramer > hexamer. Similar kinetic properties of dimers and tetramers of hTK2 revealed that the basic functional unit is the dimer. [139]. It was demonstrated that the unliganded hTK2 had lower structural stability than the inhibitor (dTTP and dCTP)-bound enzyme, being more susceptible to aggregation, thermal denaturation, and limited proteolysis by trypsin [145]. It was found that ligand binding increases structural stability in the following order: dCyd < MgdCTP < dThd < dCTP < dTTP < MgdTTP. MgATP had rather destabilized hTK2. In addition, similar to NMP kinases, hTK2 was found to undergo distinct conformational changes upon binding of substrates and inhibitors as studied by far-UV circular dichroism, limited proteolysis, and intrinsic tryptophan fluorescence [145]. Some of the compounds such as 3′-C-branched *p*-methylphenyl and 3-CF₃-4-Cl-phenyl thiourea derivatives of β-dThd were found to be among the most potent inhibitors of TK2 [230]. Recombinantly produced hTK2 was purified with bound deoxyribonucleotides which are reported to be powerful inhibitors of the enzyme i.e. dTTP, dCTP and dATP [138, 139, 159, 199, 226]. Incubation of hTK2 with its substrate (dThd or

dCyd) resulted in the removal of the tightly bound inhibitory deoxyribonucleoside triphosphates. As the three-dimensional structure of hTK2 is unknown, a homology model was created using the SWISS MODEL computer algorithm applying Dm-dNK structure as a template [139]. The physiological substrates dThd and dCyd, the phosphate donor ATP, and the feedback inhibitors dTTP and dCTP were docked to the hTK2 model as shown below in Fig. 1.16.
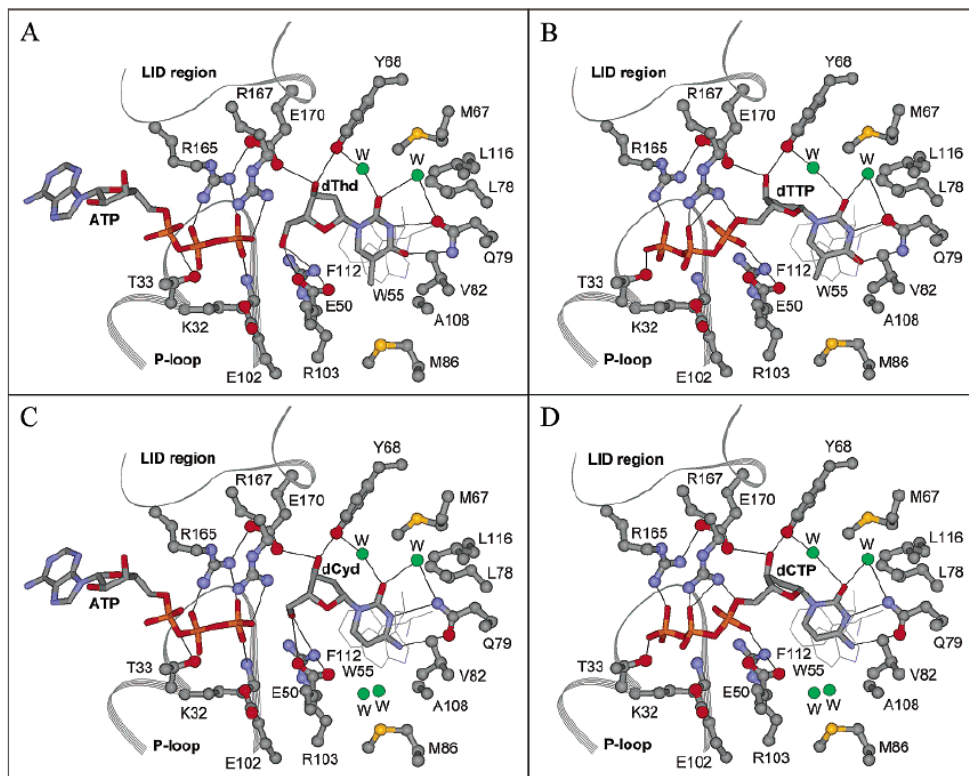
Fig. 1.16. Human TK2 model with docked deoxyribonucleosides and nucleotides. The hTK2 model with docked pyrimidine natural substrates dThd (A) and dCyd (C), the phosphate donor ATP (A and C), and the feedback inhibitors dTTP (B) and dCTP (D) are shown. The W55 on top of the nucleoside/nucleotide base and F112 behind the base are represented as wires. The P-loop (G26-T33), the arginine-rich site called LID (R161-E170), and water molecules (W, in *green*) are shown. Hydrogen bonds are represented as *black* lines [139].

It was observed that the P-loop (strand-turn-helix motif, G26-T33) and the arginine-rich LID region (R161-E170) interact with the phosphates of ATP or dTTP and dCTP explaining their tight binding in the inhibitor complexes [139]. The dNTPs may act as bisubstrate analogues for dNKs and bind at the enzyme active site with high specificity. The deoxyribonucleoside group of

dNTP binds at the deoxyribonucleoside binding site, and the triphosphate moiety fits into the phosphate donor site [139, 223-225]. The inhibition of TK2 by the end products, which most likely act as bi-substrate analogs, is depicted in Fig. 1.17.
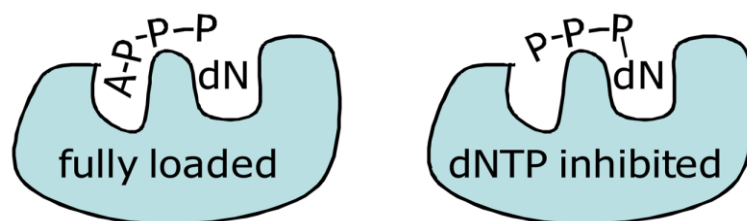
Fig. 1.17. Bisubstrate inihibition of TK2. The bisubstrate inhibition of TK2 by dNTP (deoxyribonucleoside triphosphate such as dCTP or dTTP) binding to the deoxyribonucleoside (dN) pocket is illustrated [232].

It was demonstrated that non-active site residues C189 and C264 of hTK2 were specifically glutathionylated by GSSG, but the process was reversed by the addition of dithiothreitol. In addition, the hydrogen peroxide-treated cells resulted in the reduction of TK2 activity and protein levels due to oxidative stress induced degradation of S-glutathionylated TK2 [179]. Two mutations in the hTK2 gene, H121N and I212N, were reported in patients with mitochondrial DNA depletion myopathy [229]. When these mutations were introduced in hTK2, the recombinantly produced mutants displayed altered kinetic properties [138]. The I212N mutant had less than 1% activity as compared to wild-type hTK2 with both dThd and dCyd. It showed 100-fold lower V_{max} with both dThd and dCyd, the K_m value for dThd was similar to that of wt-hTK2, but the K_m value for dCyd was 50 times increased as compared to wt-hTK2. The H121N mutant showed activity similar to wt-hTK2 when dThd was used as a substrate. Nevertheless, with dCyd as a substrate the activity was about 3-fold decreased, but the K_m was unchanged, and the V_{max} was about 30% of that of wt-hTK2. The H121N mutant lost negative cooperativity with dThd as a substrate [138].

1.2.2 *E.coli* guanosine-inosine kinase

Guanosine-inosine kinase (ATP:guanosine 5´-phosphotransferase, ecGSK, gsk, UniProt P0AEW6, EC 2.7.1.73) catalyzes the phosphorylation of guanosine and inosine to their corresponding monophosphates [149, 186]. The sodium salts of inosine 5´-monophosphate (5´-IMP) and guanosine 5´-monophosphate (5´-GMP) together with glutamate are used as flavor enhancer in food industry [250, 253]. The major industrial process of inosine 5´-monophosphate (5´-IMP) production is chemical phosphorylation of inosine [247]. But, Mori et al [248] demonstrated a novel process for producing 5´-IMP in two sequential bioreactions; the first was the fermentation of inosine by a mutant *Corynebacterium ammoniagenes*, and the second was ecGSK catalyzed conversion of inosine into IMP.

Inosine kinase was first identified in *Spirulina platensis* as a distinct enzyme protein [245]. The guanosine kinase activity has been reported in few organisms including *E. coli*, *Brevibacterium acetylicum*, *Exiguobacterium acetylicum*, *Streptococcus faecalis*, *Spirulina platensis*, *S. typhimurium* and *Trichomonas vaginalis*. Guanosine and inosine nucleosides, provided exogenously or synthesized endogenously, are converted to their corresponding monophosphates by two pathways. The minor pathway involves the phosphorylation of these two nucleosides by guanosine-inosine kinase. In the major pathway, purine nucleoside phosphorylase catalyzes the conversion of guanosine and inosines to their corresponding nucleobases and ribose-1-phosphate. Guanine and hypoxanthine are successively phosphoribosylated to their corresponding guanosine and inosine monophosphates by guanine and hypoxanthine phosphoribosyltransferases [149, 249-251].

The guanosine-inosine kinase gene was identified and mapped in *E.coli* chromosome to 11 min on the linkage map [186, 246]. The ecGSK gene which codes for a polypeptide of 433 amino acids with a molecular mass of 48,113 Da was cloned by Harlow et al [149]. It was overexpressed, and guanosine and inosine activities were determined. Mori et al [148] cloned and overexpressed a 434 amino acid ecGSK with a calculated molecular weight of 48.4 kDa. The enzyme was purified to homogeneity by ion exchange (DEAE Sepharose) and gel filtration (Sephacryl S-200) chromatography. The native protein was found to be a dimer. The PI determined by isoelectric focusing was 6.0. It was kinetically characterized. Guanosine, deoxyguanosine, inosine and xanthosine were phosphorylated by ecGSK, but adenosine, cytidine, deoxythymidine, and uridine were not accepted as phosphate acceptors. In addition to

ATP, UTP and CTP, dATP served as phosphoryl group donor. Unlike other nucleoside phosphotransferases, p-nitrophenyl phosphate did not act as a phosphate donor. ADP, AMP, GTP, acetylphosphate, tripolyphosphate, tetrapolyphosphate, PPi, and inorganic phosphate were also not accepted as phosphoryl donors. The K_m values for guanosine (Guo) and inosine (Ino) were 6.1 μM and 2.1 mM respectively. The V_{max} was 2.9 μmol/min/mg for Guo and 4.9 μmol/min/mg for Ino. The K_m values for ATP and dATP were 0.51 mM and 2.4 mM when Guo was used as a substrate. Similarly, K_m values for ATP and dATP using inosine as phosphate acceptor were 0.71 mM and 0.66 mM, respectively. The efficiency (V_{max}/K_m) for Guo phosphrylation was ~206 times higher than for Ino. The efficiency of UTP as a phosphoryl donor was 20% of that of ATP. The optimum pH values were 6.9 and 8.2 for the inosine and guanosine kinase reactions, respectively. The optimum temperature for the inosine kinase reaction was in the range of 26-39 ^{o}C, and for the guanosine kinase reaction it was 38 ^{o}C. Magnesiun (Mg^{2+}) and potassium (K^+) ions were required for the enzymatic activity. In contrast, the addition of Cu^{2+} or Zn^{2+} inhibited the ecGSK-catalyzed reactions. The enzyme activity was also inhibited by GDP and GTP. ecGSK follows an ordered Bi Bi mechanism in which Guo is the first substrate to bind, and GMP is the last product to be released. It was suggested that the guanosine kinase activity may be regulated by *in vivo* changes of nucleotide concentrations, and thus may have a role in modulating nucleotide levels [148, 252].

1.3 Aims of the present work and overview

Deoxyribonucleoside and nucleotide kinases, besides their physiological role in the synthesis of purine and pyrimidine nucleotides for RNA and DNA metabolism, are required for the metabolic activation of numerous antiviral and anticancer nucleoside analog prodrugs. Three enzymes including human guanylate kinase (hGMPK), human mitochondrial thymidine kinase (hTK2) and *E.coli* guanosine-inosine kinase (ecGSK) are peculiar in this regard, and are the least studied in their family of enzymes from structure-function perspectives. Therefore, my studies aimed to biochemically characterize them in order to provide a basis for their potential use in cancer chemotherapy.

Our preliminary goal was the cloning, expression and purification of these enzymes in catalytically active form. They were kinetically characterized. Nevertheless, due to certain limitations of the conventional NADH-dependent spectroscopic assay, it was highly demanded to

develop novel approaches for the detection and kinetic characterization of deoxyribonucleoside and nucleotide kinases. Thus, we developed two assays: a novel spectrophotometric and fluorometric enzyme-coupled assay for hGMPK, and the electrochemical detection of guanosine monophosphate with a quantum dot-based biosensor modified with hGMPK.

In order to explain the role of certain residues in enzyme catalysis and structural dynamics, a series of enzyme variants were generated by site-directed mutagenesis, and all mutants were kinetically characterized. In addition, mutant libraries were generated by directed evolution using error-prone PCR and screened for catalytically efficient mutants against antiviral and anticancer nucleoside analogs.

We optimized conditions for three-dimensional structure elucidation of these enzymes by X-ray crystallography and NMR, and determined the substrate-induced conformational changes by SAXS analysis.

We synthesized calcium carbonate microparticles of different shapes and sizes to test their enzyme loading capacities, and to select the best candidates as therapeutic enzyme carriers for their potential use in cellular targeting.

2. Materials and Methods

2.1 Materials

2.1.1 Plasmids

Human guanylate kinase (hGMPK), human mitochondrial thymidine kinase (hTK2), and *E.coli* guanosine-inosine kinase (ecGSK) were cloned into the following plasmids used for the production of proteins with different tags; NdeI and BamHI restriction sites served for ligation at the 5´ and 3´ ends of the insert, respectively.

Table 2.1 Plasmids used in this study

Plasmid	Description	Source
pET-14b	*E.coli* expression vector with N-terminal hexahistidine tag	Novagen (Cat. No. 69660-3)
pET-14bSUMOΔThr	*E.coli* expression vector with N-terminal hexahistidine-SUMO tag	pET-14b was modified by eliminating the thrombin cleavage site and introducing an additional N-terminal SUMO tag [84]
pGEX-RB	*E.coli* expression vector with N-terminal GST tag	R. Brundiers, Goettingen
pJC20HisN	*E.coli* expression vector with N-terminal decahistidine tag	T. Schüle, Goettingen
pJC20HisC	*E.coli* expression vector with C-terminal hexahistidine tag	T. Schüle, Goettingen
pET-14bEGFP-N	*E.coli* expression vector with N-terminal EGFP tag	C.S. Karamitros, et al. [86]
pK49	*E.coli* expression vector with N-terminal His$_{14}$-MBP-SUMObr tag	Steffen Frey and Dirk Görlich [82, 83]
pET-14bMBPcyt	*E.coli* expression vector with Δ23N cytosolic MBP version	Stephan Ort, Goettingen [84]
pET-14bMBPperi	*E.coli* expression vector with N-terminal MBP for periplasmic transport	Stephan Ort, Goettingen [84]
pET-14bSUMOPeriPep	*E.coli* expression vector with N-terminal SUMO periplasmic peptide leader	C.S. Karamitros, Goettingen [86]
pEGFP-N1	Mammalian expression vector with C-terminal EGFP tag	Theresa McSorley, et al. [85]
pEGFP-C1	Mammalian expression vector with N-terminal EGFP tag	Theresa McSorley, et al. [85]

2.1.2 Oligonucleotides

All oligonucleotides used for the amplification and sequence verification of hGMPK, hTK2, ecGSK and lpPOX (pyruvate oxidase from *Lactobacillus plantarum*) were synthesized by IBA GmbH, Goettingen, Germany. The synthesis scale of the single-stranded DNA oligos was from 0.01 to 0.05 µmol depending on the length of the primer. F and R mean forward and reverse oligos, respectively.

Table 2.2 Oligonucleotides used in this study

Oligos code	Sequence (5'-3')	Description
hGMPK-F	GGG AAT TCC ATA TGT CGG GCC CCA GGC CTG TGG	5´ NdeI site
hGMPK-R	CGC GGA TCC TTA GGC GCC GGT CCT TTG AGC TTT CTT G	3´ BamHI site
hGMPK-R no stop	CGC GGA TCC GGC GCC GGT CCT TTG AGC TTT CTT G	3´ BamHI site with no stop codon
hGMPK S37A-F	CTT TGG CTT CAG CGT GGC CCA TAC CAC GAG GAA C	Mutagenic primer for S37A point mutation
hGMPK S37A-R	GTT CCT CGT GGT ATG GGC CAC GCT GAA GCC AAA G	Mutagenic primer for S37A point mutation
hGMPK S37Y-F	CTT TGG CTT CAG CGT GTA CCA TAC CAC GAG GAA CCC	Mutagenic primer for S37Y point mutation
hGMPK S37Y-R	GGG TTC CTC GTG GTA TGG TAC ACG CTG AAG CCA AAG	Mutagenic primer for S37Y point mutation
hGMPK S37C-F	CTT TGG CTT CAG CGT GTG CCA TAC CAC GAG GAA CC	Mutagenic primer for S37C point mutation
hGMPK S37C-R	GGT TCC TCG TGG TAT GGC ACA CGC TGA AGC CAA AG	Mutagenic primer for S37C point mutation
hGMPK S37P-F	CAT CTT TGG CTT CAG CGT GCC TCA TAC CAC GAG GAA CCC GAG	Mutagenic primer for S37P point mutation
hGMPK S37P-R	CTC GGG TTC CTC GTG GTA TGA GGC ACG CTG AAG CCA AAG ATG	Mutagenic primer for S37P point mutation
hGMPK Y81F-F	GTT CTC GGG GAA CCT GTT TGG CAC GAG CAA GGT G	Mutagenic primer for Y81F point mutation
hGMPK Y81F-R	CAC CTT GCT CGT GCC AAA CAG GTT CCC CGA GAA C	Mutagenic primer for Y81F point mutation
hGMPK T83S-F	GAA CCT GTA TGG CTC GAG CAA GGT GGC	Mutagenic primer for T83S point mutation
hGMPK T83S-R	GCC ACC TTG CTC GAG CCA TAC AGG TTC	Mutagenic primer for T83S point mutation
hGMPK T83A-F	GGA ACC TGT ATG GCG CGA GCA AGG TGG CG	Mutagenic primer for T83A point mutation
hGMPK T83A-R	CGC CAC CTT GCT CGC GCC	Mutagenic primer for T83A

Oligos code	Sequence (5'-3')	Description
	ATA CAG GTT CC	point mutation
hGMPK N42P-F	GTC CCA TAC CAC GAG GCC TCC GAG GCC CGG CGA G	Mutagenic primer for N42P point mutation
hGMPK N42P-R	CTC GCC GGG CCT CGG AGG CCT CGT GGT ATG GGA C	Mutagenic primer for N42P point mutation
pET-14bSUMO-F	GGA GGA TAA CGA TAT TAT TGA GG	pET-14bSUMOΔThr sequencing primer (starts from the C-terminus of the SUMO tag)
pET-14b-R	ATT AAC CCT CAC TAA AGG GA	pET-14b sequencing primer (3´)
pGEX-RB-F	GGG CTG GCA AGC CAC GTT TGG TG	pGEX-RB sequencing primer (5´)
pGEX-RB-R	CCG GGA GCT GCA TGT GTC AGA GG	pGEX-RB sequencing primer (3´)
pK49-F	CAG ACC CCG GAT GAA CTG GAG	pK49 sequencing primer (5´)
pK49-R	CTG GAT CTA TCA ACA GGA GTC CA	pK49 sequencing primer (3´)
lpPOX-F	GGG AAT TCC ATA TGG TTA TGA AAC AAA CAA AAC AAA CTA AC	5´ NdeI site
lpPOX-R	CGC GGA TCC TTA AAA CCC ACC TGT CCA ATT TG	3´ BamHI site
hTK2-F	GGG AAT TCC ATA TGC TGC TGT GGC CTC TGC GTG G	5´ NdeI site of full-length sequence-optimized hTK2
hTK2-R no stop	CGC GGA TCC GGG CAG TGT TTA CGG TTT TCC GGG G	3´ BamHI site of full-length sequence-optimized hTK2 with no stop codon
hTK2-R	CGC GGA TCC TTA CGG GCA GTG TTT ACG	3´ BamHI site of full-length sequence-optimized hTK2
hTK2-Δ44N-F	GGG AAT TCC ATA TGC AAG AGA AAG AAA AAA AAA GCG	For truncation of 44 amino acids from N-terminus of hTK2
hTK2-Δ50N-F	GGG AAT TCC ATA TGA GCG TGA TCT GTG TGG AAG GC	For truncation of 50 amino acids from N-terminus of hTK2
hTK2-Δ8C-R	CGC GGA TCC TTA GGT CAG AAT ACG GTC ACG GTT TTG C	For truncation of 8 amino acids from C-terminus of hTK2
hTK2-Δ25C-R	CGC GGA TCC TTA GTG ATG GTC CGC TTC GAT TAC C	For truncation of 25 amino acids from C-terminus of hTK2
ecGSK-F	GGG AAT TCC ATA TGA AAT TTC CCG GTA AAC GTA AAT CC	5´ NdeI site
ecGSK-R no stop	CGC GGA TCC ACG ATC CCA GTA AGA CTC TTC C	3´ BamHI site with no stop codon
ecGSK-R	CGC GGA TCC TTA ACG ATC CCA GTA AGA CTC TTC C	3´ BamHI site
ecGSK-Δ30N-F	GGG AAT TCC ATA TGA CCA GCG CTG CCT GGG TAG TG	For truncation of 30 amino acids from N-terminus of ecGSK
ecGSK-Δ21C-R	CGC GGA TCC TTA TGA ATG CTG GTT CAG TAC CTG ATA GC	For truncation of 21 amino acids from C-terminus of ecGSK

2.1.3 *Escherichia coli* strains

The following expression and non-expression strains of *E.coli* were used.

Table 2.3 *E. coli* strains

Strain	Genotype	Source
XL1-Blue	$F^-::Tn10/proA^+B^+\,lacI^a\Delta(lacZ)M15/recA1endA1\,gyrA96(Nal^r)thi^-,\;hsdR17(rK^-mK^+)supE44\;relA1\;lac$	Bullock, W.O., et al. [95]
DH5α	$F^-/endA1\;\;hsdR17(rk^-mk^+)\;\;supE44\;\;thi\text{-}l\;\;recA1\;gyrA(Nal^r)\;\;relA1\;\;\Delta(lacZYA\text{-}argF)U169\;\;(m80\;lacZ\Delta M15\,)$	Gibco BRL (Eggenstein)
BL21(DE3)	F^- *ompT*, *hsd*S(r$_B$ m$_B$) *dcm gal* λ(DE3)	Studier, F.W., et al. [96]
BL21(DE3)pLysS	F^- *ompT hsd*S(rB⁻, mB⁻) *dcm gal* λ(DE3) [pLysS Cm^r]	Studier, F.W., et al. [96]
C41(DE3)	Same genotype as BL21(DE3) with an uncharacterized mutation that increases the overexpression of toxic proteins	Miroux, B. and Walker, J.E. [97]
Origami B(DE3)	F^- *ompT hsd*S$_B$(r$_B^-$ m$_B^-$) *gal dcm lacY1 ahpC* (DE3) *gor522*:: Tn*10 trxB* (Kan^R, Tet^R)	Novagen
Rosetta-gami B(DE3)	F^- *ompT hsd*S$_B$ (r$_B^-$ m$_B^-$) *gal dcm lacY1 ahpC* (DE3) *gor522*::Tn*10 trxB* pRARE (Cam^R, Kan^R, Tet^R)	Novagen
KY895	F^-, *tdk-1*, *ilv*	Christian Monnerjahn, Goettingen Igarashi, K., et al. [320]

2.1.4 Enzymes

KAPAHiFi DNA polymerase was purchased from Peqlab Biotechnologie GmbH (Erlangen, Germany). Horseradish peroxidase, pyruvate kinase/lactic dehydrogenase were from Sigma-Aldrich (St. Louis, USA). High Fidelity Phusion DNA polymerase, *Taq* DNA polymerase, restriction endonucleases and T4 DNA ligase were from New England BioLabs GmbH

(Frankfurt/Main, Germany). SUMO-protease from *Saccharomyces cerevisiae* was recombinantly produced in *E. coli* C41 [118].

2.1.5 Kits

The NucleoSpin Gel and PCR Clean-up, NucleoSpin Plasmid DNA purification and NucleoBond Xtra Midi kits were purchased from Macherey-Nagel GmbH & Co. KG (Düren, Germany). Qiagen Plasmid Midi kit was from Qiagen GmbH (Hilden, Germany). KAPAHiFi PCR kit was from Peqlab Biotechnologie GmbH (Erlangen, Germany).

2.1.6 Chemicals

Chemicals used in this study were obtained from different suppliers including Merck, AppliChem and Sigma-Aldrich. Adenosine 5′-triphosphate (ATP), adenosine 5′-diphosphate (ADP), guanosine 5′-monophosphate (GMP), 2′-deoxyguanosine 5′-monophosphate (dGMP), 6-thioguanine, 6-thioguanosine, D-glucose-$^{13}C_6$, ammonium-^{15}N chloride, pyruvic acid, bromophenol blue, α-Lactose, chloramphenicol, Coomassie Brilliant Blue G 250, tris(2-carboxyethyl)phosphine hydrochloride (TCEP), Na_2SO_4, Dulbecco's Modified Eagle Medium (DMEM), RPMI 1640, 3-(4,5-Dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT), (E)-5-(2-Bromovinyl)-2′-deoxyuridine (BVDU), 5-Fluorouracil, azidothymidine (AZT), ganciclovir, acyclovir, cladribine, (β-D-Arabinofuranosyl)cytosine (Ara-C) and 5-Fluoro-2′-deoxyuridine were purchased from Sigma–Aldrich (St. Louis, USA). Gemcitabine was from Synchem UG & Co. KG (Germany). Agar, agarose, yeast extract, peptone (from casein), $MnCl_2$, ethidium bromide, D-glucose, ampicillin, β-mercaptoethanol and dithiothreitol (DTT) were from AppliChem GmbH (Darmstadt, Germany). Na_2HPO_4, NaH_2PO_4, K_2HPO_4, KH_2PO_4, $MgCl_2$, NaOH, HCl, sodium dodecyl sulfate (SDS), NaCl, triton X-100, glycerol and glycine were from Merck (Darmstadt, Germany). dNTPs master mix was from Fermentas. 1Kb GeneRuler DNA Ladder and PageRuler Unstained Protein Ladder were from Thermo Scientific (USA). Protino Ni-IDA for affinity chromatography purification was from Macherey-Nagel GmbH & Co. KG (Düren, Germany). HEPES, acrylamide/bisacrylamide, isopropyl β-D-1-thiogalactopyranoside (IPTG), EDTA, inhibitor cocktail plus and imidazole were from Carl Roth GmbH + Co. KG (Karlsruhe, Germany). Nicotinamide adenine dinucleotide reduced (NADH) and phosphoenolpyruvate were from Roche Diagnostic GmbH (Manheimm, Germany). Adenosine-

5´-[(β, γ)-imido] triphosphate (AMP-PNP), 6-thioGMP, Ap5G and 6-mercaptopurine-riboside-5´-monophosphate were from Jena Bioscience GmbH (Jena, Germany). NH_4Cl was from J.T.Baker B.V.-Deventer-Holland. Amplex red was from Cayman Chemical Company. Tris-(hydroxymethyl) aminomethane was from VWR Chemicals (Leuvan, Belgium). L-glutamine and fetal calf serum (FCS) were purchased from PAA Laboratories GmbH (Cölbe, Germany). G418 solution was from Carl Roth GmbH (Karlsruhe, Germany). FuGene1 HD Transfection Reagent and Cell Proliferation Reagent WST-1 were purchased from Roche (Mannheim, Germany). Anti-GFP IgG antibody was gifted by Dr. Dieter Schmitt (MPI for Biophysical Chemistry, Goettingen). IgG antibodies conjugated to horseradish peroxidase were purchased from DiaNova (Hamburg, Germany).

2.1.7 Consumables

General glassware were from DURAN Group GmbH (Germany), polypropylene centrifuge tubes, serological pipettes, tissue culture plates, filter tips, tissue culture flasks, microtubes, and tissue culture dishes were from SARSTEDT AG & Co (Nümbrecht, Germany). Bottle top filters were from NALGENE (New York, USA). PCR tubes were from Sigma-Aldrich. Disposable plastic cuvettes were from Carl Roth GmbH+Co.KG (Germany). Slide-A-Lyzer mini dialysis units and disposable polypropylene columns were from Thermo Scientific (Rockford, USA). Vivaspin ultrafiltration spin columns and syringe filters were from Sartorius Stedim Biotech GmbH (Goettingen, Germany). Filter papers and nitrocellulose membranes were from Schleicher & Schuell (Dassel, Germany).

2.1.8 General Instruments

FP-8300 fluorescence spectrometer and V-650 UV/VIS spectrophotometer were from JASCO (Germany). UVIKON 943 UV/VIS spectrophotometer was from BIO-TEK (Neufahrn, Germany). NanoDrop UV/VIS spectrophotometer was from PeqLab Biotechnologie GmbH (Germany). ÄKTA prime plus was from GE Healthcare Life Sciences (Uppsala, Sweden). TPersonal thermocycler was from Biometra GmbH (Goettingen, Germany). Centrifuges used in this study were Eppendorf 5415D (Hamburg, Germany), SORVALL RC4 and SORVALL RC 6+ from Thermo Scientific (Osterode, Germany), and Heraeus Megafuge 1.0R from Kendro (Germany). Innova 4230 Incubator Shaker was from New Brunswick Scientific (Edison, USA).

Branson Sonifier 250 was from G. Heinemann (Schwäbisch Gmünd, Germany). UV transilluminators, 312 nm and 366 nm were from Bachofer (Reutlingen, Germany). Leica TCS SP5 and Leica DMIRB were from Leica Microsystems (Germany). The electrophoresis power supply was from RENNER GmbH (Darmstadt, Germany). DNA gel chambers were custom-made by the technical service department of the Max Planck Institute for Biophysical Chemistry (Goettingen, Germany). Mini-Protean precast gels and Mini Trans-Blot electrophoretic cell chambers were from BioRad (Richmond, USA). X-ray film developer: Gevamatik 60 was from AGFA (Hannover, Germany). The autoclave used was Systec VX-150 (Wettenberg, Germany). Analytical balance was from Sartorius (Goettingen, Germany). Elutrap electroelution system was from Schleicher and Schuell (Dassel, Germany).

2.1.9 Culture media

2.1.9.1 Bacterial culture media

LB (Luria-Bertani) medium: Tryptone/peptone (10 g/l), yeast extract (5 g/l) and NaCl (5 g/l). The pH of LB medium was adjusted to 7.5 with NaOH. It was autoclaved at 121 °C for 20 min. For LB plates, agar with final concentration of 1.5% (w/v) was added prior to autoclaving. The autoclaved medium was allowed to cool to 55 °C; for bacterial growth selection, the corresponding antibiotic (ampicillin 100 µg/ml, chloramphenicol 20 µg/ml, and kanamycin 30 µg/ml working concentrations) was added. The medium and agar plates were stored at 4 °C.

TB (Terrific Broth) medium: The following ingredients were added to 800 ml millipore water: Tryptone/peptone (12 g), yeast extract (24 g) and glycerol (4 ml). The volume was adjusted to 900 ml with millipore water and sterilized by autoclaving. The media was allowed to cool to room temperature, and the final volume of 1000 ml was made by adding 100 ml of 0.17 M KH_2PO_4 and 0.72 M K_2HPO_4.

LBE-5052 autoinducing medium: Trypton/peptone (10 g/l), yeast extract (5 g/l), glycerol (5 g/l), glucose (0.5 g/l), lactose (2 g/l), Na_2SO_4 (0.7 g/l), NH_4Cl (2.5 g/l), H_2O to make it 900 ml, mixed well and autoclaved. The following solutions were added before use: 1 ml of sterile 2 M $MgSO_4$, 1 ml of filter-sterilized 1000x metals mix, and 100 ml of filter-sterilized 50 mM potassium phosphate mix.

50 mM potassium phosphate mix (100 ml): 10 ml of 1 M KH_2PO_4, 40 ml of 1 M K_2HPO_4 and 50 ml H_2O.

1000x metals mix (100 ml): 1 ml of 1 M $MnCl_2$, 1 ml of 1 M $ZnSO_4$, 1 ml of 0.2 M $CoCl_2$, 1 ml of 0.2 M $NiCl_2$, 46 ml H_2O. All metal solutions were mixed together and filter-sterilized.

Then, 50 ml of 0.1 M $FeCl_3$ solution (in 0.1 M HCl) was mixed with the sterilized metal solution (50 ml) to make it 100 ml of 1000x metals mix. Upon 1000-fold dilution, the final concentrations of metals were 10 µM Mn, 10 µM Zn, 2 µM Co, 2 µM Ni and 50 µM Fe. The solution was stored at room temperature.

M9 minimal growth medium (1 L): 200 ml of 5X M9 salts, 20 ml of D-glucose (20 g/100 ml) (0.2 µm filter sterilized)[*], 1 ml of 2 M $MgSO_4$ (autoclaved), 0.1 ml of 1 M $CaCl_2$ (autoclaved), 1 ml of thiamine (1 mg/ml) (0.2 µM filter-sterilized), and 766 ml of H_2O (autoclaved).

5X M9 salts (autoclaved): KH_2PO_4 (15 g/l), $Na_2HPO_4.7H_2O$ (64 g/l), NaCl (2.5 g/l), and NH_4Cl[*] (pH adjusted to 7.2 with NaOH) (5.0 g/l)

[*]For isotopic labelling, [13]C D-glucose and [15]NH_4Cl (Sigma-Aldrich) were substituted for unlabeled D-glucose and NH_4Cl, respectively. The corresponding antibiotics (ampicillin 100 µg/ml and chloramphenicol 20 µg/ml working concentrations) were added to the M9 minimal media for plasmid selection.

2.1.9.2 Mammalian cell culture medium

DMEM (Sigma-Aldrich) supplemented with 10% FCS (PAA), 1% L-glutamine and G418 (Carl Roth) was used for routine cell culture. The medium without antibiotics was used for transfection purposes.

2.1.10 General buffers

Lysis buffer A: 50 mM HEPES (or Na_2HPO_4), pH 8.0, 300 mM NaCl, 0.44 mM EDTA 0.5% Triton X-100, 5 mM DTT (added before use), and protease inhibitors cocktail (added before use)

Lysis buffer B: 50 mM HEPES, pH 8.0, 500 mM NaCl, 4.4 mM $MgCl_2$, 0.44 mM EDTA, 0.5% Triton X-100, 1 M urea, 5 mM DTT, and 1X cocktail of protease-inhibitors

Wash buffer A: 50 mM HEPES (or Na_2HPO_4), pH 8.0, 300 mM NaCl, 10 mM imidazole, 5% glycerol, and 2 mM DTT (added before use)

Wash buffer B: 50 mM HEPES (or Na_2HPO_4), pH 8.0, 300 mM NaCl, 5 mM imidazole 5% glycerol, and 2 mM DTT (added before use)

Wash buffer-1: 50 mM HEPES, pH 8.0, 290 mM NaCl, 4.4 mM $MgCl_2$, 0.44 mM EDTA, 0.1% Triton X-100, 10 mM imidazole, 10% glycerol, and 3 mM DTT

Wash buffer-2: 50 mM HEPES, pH 8.0, 10 mM $MgCl_2$, 0.44 mM EDTA, 0.1% Triton X-100, 10% glycerol, 5 mM ATP, and 3 mM DTT

Wash buffer-3: 25 mM HEPES, pH 7.5, 4.4 mM $MgCl_2$, 0.44 mM EDTA, 0.1% Triton X-100, 10% glycerol, and 2 mM DTT

Elution buffer A: 50 mM HEPES (or Na_2HPO_4), pH 8.0, 150 mM NaCl, 250 mM imidazole, 10 % glycerol, 0.1% Triton X-100, and 2 mM DTT (added before use)

Dialysis buffer A: 50 mM HEPES (or Na_2HPO_4), pH 8.0, 300 mM NaCl, 10 % glycerol, and 1 mM DTT (added before use)

Gel filtration buffer A: 50 mM HEPES (or Na_2HPO_4), pH 7.4, 300 mM NaCl, 5% glycerol, and 1 mM DTT/TCEP (added before use)

Buffer-A (2X): 200 mM Tris/HCl, pH 7.5, 200 mM KCl, and 20 mM $MgCl_2$

Phosphate-buffered saline (PBS) (1X): 137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, and the pH was adjusted to 7.4 with HCl.

Laemmli sample buffer (2X): 65.5 mM Tris/HCl, pH 6.8, 2.1% (w/v) SDS, 26.3% (w/v) glycerol, 0.01% (w/v) bromophenol blue, and 50 µl β-mercaptoethanol per 950 µl

DNA sample loading dye (10X): 50% glycerol, 0.25% bromophenol blue, and 0.25% xylene cyanol FF in 1X TAE buffer

TAE (Tris-acetate-EDTA) buffer (1X): 40 mM Tris (pH 7.6), 20 mM acetic acid, and 1 mM EDTA

TBE (Tris-borate-EDTA) buffer (1X): 89 mM Tris (pH 7.6), 89 mM boric acid, and 2 mM EDTA

TE (Tris-EDTA) buffer (autoclaved): 10 mM Tris/HCl (pH 8.0) and 1 mM EDTA

Competent wash buffer (pH 5.8): 15 ml of 1 M CH_3COOK (pH 7.5) , 0.75 g of $CaCl_2$, 6.1 g of RbCl, 4.95 g of $MnCl_2$, 75 ml of glycerol, H_2O up to a final volume of 500 ml, pH adjusted with 0.2 M of CH_3COOH. The buffer was filter-sterilized and prechilled on ice before use.

Competent freezing buffer (pH 6.8): 10 ml of 0.5 M Mops (pH 6.8), 5.5 g of $CaCl_2$, 0.61 g of RbCl, 100 ml of glycerol, H_2O up to final volume of 500 ml, pH adjusted with 0.2 M NaOH. The buffer was filter-sterilized and prechilled on ice before use.

2.1.11 Bioinformatic tools

Discovery studio 4.0 visualizer: For visualization and analysis of three-dimensional structures of proteins

PyMOL molecular graphics system: For visualization and analysis of three-dimensional structures of proteins

Gnuplot 5.0: For plotting functions and data points in both two- and three-dimensional plots

BLAST (NCBI): For basic local alignment of nucleotide and amino acid sequences

Clustal W2: For multiple sequence alignment of nucleotide and amino acid sequences

2.2 Methods

2.2.1 Sterilization methods

Culture media, buffer solutions, millipore water, microcentrifuge tubes, wooden toothpicks and pipette tips were sterilized by autoclaving at 121 °C for 20 min. All glassware was sterilized by heating at 200 °C in an oven (Memmert) for 4 h. Heat-labile solutions such as IPTG, 6-thioguanine, and antibiotics were filter-sterilized using 0.2 µm filters (Sartorius).

2.2.2 Preparation of *E. coli* culture glycerol stocks

A single colony of a construct was picked from the LB agar plate to inoculate 1-3 ml of liquid LB containing the appropriate antibiotic. It was incubated at 37 °C with constant shaking (200–250 rpm) preferably 12–16 h overnight. 1 ml of the culture was centrifuged, and the supernatant was discarded. The cells were resuspended in 0.5 ml of fresh LB, and glycerol was added to a final concentration of 30% (v/v). The cell suspension was gently vortexed, frozen in liquid nitrogen, and then stored at -80 °C. To recover bacteria from the glycerol stock for plasmid purification, formation of competent cells, or protein overexpression, some of the frozen bacteria were scrapped from the top with a toothpick and streaked on an LB agar plate. They were grown overnight at 37 °C to get isolated colonies.

2.2.3 Preparation of *E. coli* competent cells

A single colony of *E.coli* cells (XL1-Blue or DH5α) was inoculated into 5 ml LB in 50 ml flask and incubated overnight at 37 ^{o}C with vigorous shaking.

For preparation of chemically competent cells, fresh 50 ml LB medium was inoculated with 0.5 ml of the overnight culture in 250 ml flask and incubated at 37 ^{o}C until the OD_{600} reached 0.3-0.5. The cells were transferred to a sterile Falcon tube, chilled on ice for 15 min, and then harvested by centrifugation at 3,000 x g for 10 min at 4 ^{o}C. The supernatant was discarded, and the cells were rinsed with 10 ml competent wash buffer (see section 2.1.10). The buffer was removed, and the cells were resuspended in 10 ml competent wash buffer and incubated on ice for 15 min. The cells were recovered by centrifugation at 2,300 x g at 4 ^{o}C for 5 min, and the supernatant was discarded. The cell pellets were resuspended in 3 ml competent freezing buffer (see section 2.1.10) and aliquoted in microcentrifuge tubes (100-500 µl/tube). The competent cells were flash-frozen in liquid nitrogen and stored at -80 ^{o}C until use for transformation.

For preparation of electro-competent cells, 500 ml LB was inoculated with the overnight culture and grown at 37 ^{o}C until an OD_{600} of 0.4-0.6. Cells were then cooled on ice and collected in sterile bottles by centrifugation at 3,000 x g for 15 min at 4 ^{o}C. The pellet was washed twice with 500 ml ice-cold sterile water and twice with 200 ml ice-cold sterile 10% glycerol. The final cell pellet was resuspended in 5 ml ice-cold sterile 10% glycerol, aliquoted in cold microtubes, flash-frozen in liquid nitrogen and stored at -80 ^{o}C.

2.2.4 Cloning of human guanylate kinase

2.2.4.1 PCR amplification

The 591 bp open reading frame (ORF) of human guanylate kinase (UniProt entry Q16774, hGMPK, GUK1, or GMK), was amplified via polymerase chain reaction (PCR) using the template DNA obtained from Oliver Spangenberg [81]. NdeI and BamHI sites were incorporated into the forward (hGMPK-F) and reverse (hGMPK-R) primers, respectively (Table 2.2). The Phusion High-Fidelity DNA polymerase (New England BioLabs) was used according to the manufacturer's instructions. The 50 µl PCR reaction mixture contained 1X Phusion HF buffer, 200 µM dNTPs, 0.5 µM forward and reverse primers, 10 ng template DNA and 1.0 unit of Phusion HF DNA polymerase. Tm (melting temperature) values for the oligonucleotides were

calculated using the online NEB Tm calculator based on thermodynamic data from Breslauer et al. (1986) and the salt correction outlined in Owczarzy et al. (2004) [NEB webpage]. Thermocycling conditions were set to initial denaturation at 98 ^{o}C for 30 s, followed by 25 cycles of denaturation at 98 ^{o}C for 7 s, primers annealing at 67 ^{o}C for 20 s, extension at 72 ^{o}C for 15 s, and a final extension at 72 ^{o}C for 7 min.

2.2.4.2 Agarose gel electrophoresis

To separate and visualize the hGMPK PCR product for correct size, it was analyzed by agarose gel electrophoresis. Agarose gel of 1% (w/v) in TAE buffer (see section 2.1.10) was used. The PCR product was mixed with the 10X DNA loading dye. The PCR sample and 1 Kb GeneRuler DNA Ladder (Thermo Scientific) were loaded into the wells, and the gel was run at 75 volts power for about 2 h. The gel was stained in the appropriate volume of 0.5 µg/ml ethidium bromide (EtBr) for 15 min. After staining, the gel was briefly rinsed with Millipore water to remove any residual staining solution. To visualize and excise the hGMPK DNA band, the gel was placed on a transilluminator with UV light of 366 nm or 312 nm. The gel picture was taken, and the 591 bp hGMPK DNA band was excised with a clean scalpel for DNA extraction.

2.2.4.3 DNA extraction from agarose gel

The hGMPK DNA from the gel slice was extracted by using NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). According to the manufacturer's protocol, the gel slice was solubilized in the buffer NT1 (provided with the kit) and loaded on the NucleoSpin Gel and PCR Clean-up column. The column was washed with buffer NT3 (provided with the kit), and finally the hGMPK DNA was eluted in the TE buffer (see section 2.1.10). The purified hGMPK DNA sample was subjected to restriction enzyme digestion and ligation reactions.

2.2.4.4 Ligation reaction

For creating compatible sticky ends for ligation, the hGMPK DNA sample was treated with 10 units of NdeI and BamHI-HF in 1X CutSmart buffer at 37 ^{o}C in a water bath for 1 h. The enzyme digest was cleaned by the NucleoSpin Gel and PCR Clean-up kit according to the manufacturer's protocol. The purified hGMPK DNA fragment with sticky NdeI and BamHI ends was ligated into pET-14bSUMOΔThr, pGEX-RB, pEGFP-N1, and pEGFP-C1 vectors using T4 DNA ligase (Fig. 2.1). The molar ratio of 1:3 vector to insert was used. In a 20 µl reaction volume,

a.
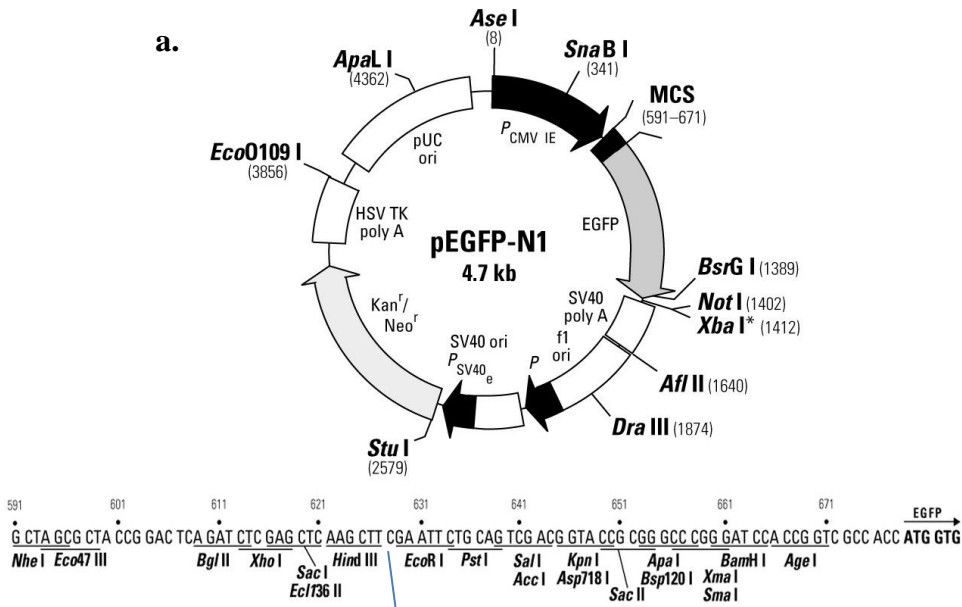
pEGFP-N1
4.7 kb

NdeI restriction site (628 **CATATG** 633) was introduced in between HindIII and EcoRI sites.

b.

pEGFP-C1
4.7 kb
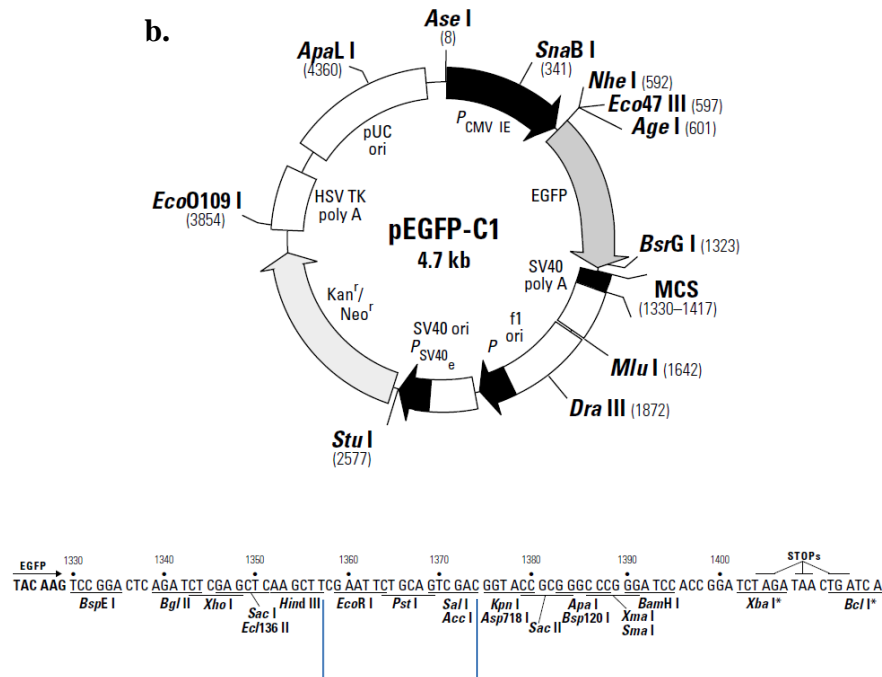
NdeI restriction site (1357 TCA **CAT ATG** GCT 1368) was introduced in between HindIII and KpnI sites.

c.

d.
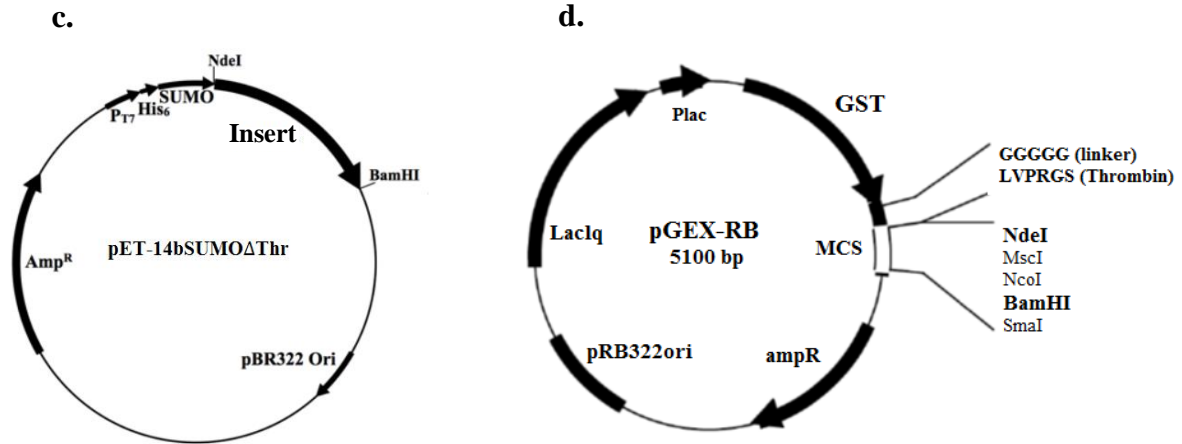
Fig. 2.1. Vector maps. (a) pEGFP-N1 is a mammalian cloning vector. An NdeI restriction site (CATATG) was introduced in between HindIII and EcoRI sites. NdeI and BamHI sites were used for ligation of insert followed by a C-terminal EGFP tag. (b) pEGFP-C1, a mammalian cloning vector was modified to introduce an NdeI restriction site in between HindIII and KpnI sites. The insert was ligated using NdeI and BamHI sites introducing an EGFP tag at the N-terminus of the insert. (c) pET-14bSUMOΔThr, and (d) pGEX-RB are *E.coli* expression plasmids with His$_6$-SUMO (small ubiquitin-related modifier) and GST (Glutathione-S-Transferase) tags, respectively. In both plasmids, the tag is N-terminal to the ligated insert.

50 ng of linear plasmid, 18 ng of hGMPK DNA fragment, 1X T4 DNA ligase buffer and 400 units of T4 DNA ligase were mixed. The reaction mixture was incubated at 16 $^{\circ}$C overnight. The ligation mixture was used to transform *E.coli* competent cells.

2.2.4.5 Transformation of chemically competent cells

The XL1-Blue chemically competent cells were thawed on ice, and 50 μl was transferred to 1.5 ml microcentrifuge tube. About 5 μl of the ligation mixture was added to the competent cells and mixed gently by flicking the tube a few times. The competent cells/DNA mixture was incubated on ice for 30 min. Heat shock was given at 42 $^{\circ}$C for 35 s, and the tube was transferred to ice for 2 min. 250 μl of LB medium was added to the tube and was incubated at 37 $^{\circ}$C with vigorous shaking at 1000 rpm for 1 h. 25 μl of the cell suspension was spread on a pre-warmed ampicillin selection plate and was incubated at 37 $^{\circ}$C overnight. The colonies obtained on the LB agar plate were screened for positive constructs by colony PCR.

2.2.4.6 Colony PCR

Colony PCR was used to determine the presence or absence of hGMPK DNA insert in pET-14bSUMOΔThr constructs by using the transformed XL1-Blue colonies from the LB agar plates. The hGMPK specific primers were used to target the insert DNA to ensure the specificity and correct size of the inserted DNA. Alternatively, vector-specific primers were used for screening of multiple constructs simultaneously. Six colonies were selected from the ampicillin-containing LB agar plate. Each 50 µl reaction mixture contained 1X standard Taq reaction buffer, 200 µM dNTPs, 0.2 µM forward and reverse primers, a small amount of each colony, and 1.25 units of the standard Taq DNA polymerase. Thermocycling conditions were as follows: initial denaturation at 95 °C for 5 min, 30 cycles of denaturation at 95 °C for 25 s, annealing at 67 °C for 30 s and extension at 68 °C for 1 min, and a final extension at 68 °C for 7 min. The PCR products were tested by agarose gel electrophoresis. The colonies which manifested the presence of correct size inserts were considered as positive constructs. Positive colonies were used to inoculate 5 ml LB medium containing 100 µg/ml of ampicillin for plasmid purification by mini or midi prep.

2.2.4.7 *E. coli* plasmid purification

For small-scale preparation of highly pure hGMPK[pET-14bSUMOΔThr], the NucleoSpin Plasmid kit was used (yield < 25 µg of plasmid). The 5 ml XL1-Blue culture preparation and plasmid purification were performed according to the manufacturer's protocol. The purified plasmid was tested by restriction digestion with NdeI and BamHI-HF followed by agarose gel electrophoresis to ensure the presence of hGMPK insert, and ultimately sent to Seqlab (Goettingen, Germany) for DNA sequence verification. The final construct includes an N-terminal hexahistidine tag, followed by the SUMO (small ubiquitin-related modifier; SUMO family protein SMT3 of 101 residues) tag, which was used to improve heterologous protein solubility and stability [98]. The concentration of hGMPK[pET-14bSUMOΔThr] was determined by absorbance at 260 nm.

2.2.4.8 Determination of DNA concentration and purity

The DNA concentration of the hGMPK sample was calculated from its UV absorbance at 260 nm where an absorbance of 1 (1 cm path length) is equivalent to 50 µg DNA/ml. The absorbance

was measured on a JASCO V-650 UV-Vis spectrophotometer in a 0.5 ml quartz cuvette with a path length of 1 cm. The measured absorbance was in between 0.1 and 0.7 in order to be in the linear part of Lambert-Beer´s law. Alternatively, the DNA concentration was measured by NanoDrop UV/VIS spectrophotometer. The plasmid purity was also checked by UV spectroscopy. A ratio of A260/A280 between 1.80–1.90 and A260/A230 around 2.0 indicates pure plasmid DNA. An A260/A280 ratio above 2.0 is a sign for too much RNA in the preparation, an A260/A280 ratio below 1.8 indicates protein contamination. The quality of the construct was tested by agarose gel electrophoresis. For overexpression and purification, the hGMPK[pET-14bSUMOΔThr] was transformed into BL21(DE3)pLysS for tight regulation and expression control under the T7 promoter.

2.2.5 Expression and purification of hGMPK

For improved production of hGMPK, His$_6$-SUMO- or GST tag-containing vectors were used. The maps of both vectors are shown in Fig. 2.1. Both constructs were transformed into the *E.coli* expression strain BL21(DE3)pLysS for tight regulation of expression. The hGMPK was expressed and purified as a fusion with His$_6$-SUMO or GST tags.

2.2.5.1 Expression and purification of His-tagged hGMPK

One liter of lactose-containing auto-inducing media was inoculated with the starter culture of BL21(DE3)pLysS carrying the hGMPK[pET-14bSUMOΔThr] plasmid. The culture was incubated at 37 °C with rapid shaking until the optical density at 600 nm reached about 0.7, and was then transferred to 21 °C for overnight induction. Cells were harvested by centrifugation at 4,000 rpm for 30 min at 4 $^{\circ}$C. For purification of hGMPK by affinity chromatography, the cell pellet was re-suspended in lysis buffer A (see section 2.1.10), and lysed by sonication. The lysate was cleared by centrifugation at 10,000 x g for 1 h at 4 °C and subjected to batch/gravity-flow purification by affinity chromatography. One gram of Protino Ni-IDA (Macherey-Nagel, Düren Germany) resin was added to the supernatant in a centrifugation tube, and the suspension was agitated on an orbital shaker for 3 h at 4 °C. The resin was washed two times with 40 ml of wash buffer A (see section 2.1.10) by incubating on an orbital shaker for 25 min in a cold room at 4 $^{\circ}$C. Similarly, the sample was washed two times with 40 ml wash buffer B (see section 2.1.10) for 25 min incubation on an orbital shaker at 4 °C. The resin was transferred to an empty

chromatography column (5 ml column bed, Thermo Scientific), and the protein was eluted with elution buffer A. This yielded 110 mg of His$_6$-SUMO fusion protein, as determined by the Bradford dye-binding assay. The His$_6$-SUMO tag was cleaved by SUMO-protease (1:100 molar ratios) at room temperature for 30 min. To remove imidazole from the sample, it was passed through a PD10 column (Sephadex-G25) using dialysis buffer A (see section 2.1.10). The His$_6$-SUMO tag was removed from the sample by adding 500 mg of Protino Ni-IDA resin and incubated for 30 min on an orbital shaker at 4 °C. The resin suspension was passed through an empty column (5 ml column bed, Thermo Scientific) with a filter frit to collect the purified hGMPK in the flow-through. To purify the enzyme to homogeneity, it was loaded onto a Superdex 75 10/300 GL gel filtration column (GE Healthcare) pre-equilibrated with gel filtration buffer A (section 2.1.10). The protein was eluted as a monomer as indicated by the elution profile of marker proteins (BioRad Gel Filtration Standard). The monomer peak was pooled, concentrated to 38 mg/ml, aliquoted, and stored at -80 °C. Protein concentration was determined by the Bradford dye-binding assay.

2.2.5.2 Expression and purification of GST-tagged hGMPK

For growing a starter culture, 50 ml LB having ampicillin and chloramphenicol antibiotics was inoculated with a single colony of hGMPK[pGEX-RB]-containing BL21(DE3)pLysS cells. It was grown overnight at 37 °C with vigorous shaking. The 50 ml starter culture was added to 950 ml of TB (terrific broth) medium containing the two antibiotics as mentioned above. The main culture was grown at 37 °C with 250 rpm shaking until the OD$_{600}$ reached ~0.7. For induction of protein expression, 0.5 mM IPTG was added to the culture medium and incubated at 21 °C overnight. The cells were harvested by centrifugation at 5,000 x g for 20 min at 4 °C. The expression level of GST-hGMPK was tested by heating a fraction of cell pellets in Laemmli sample buffer at 95 °C for 5 min and loading it on 12% SDS-PAGE. Once a high level expression was confirmed, the cell pellet was resuspended in lysis buffer A (see section 2.1.10) and sonicated for cell disruption. The lysate was centrifuged at 10,000 x g for 1 h at 4 °C, and the supernatant containing the soluble overexpressed GST-hGMPK was subjected to batch/gravity-flow purification by using Glutathione Sepharose 4B. All purification steps were carried out at 4 °C in a cold room. For 1000 ml culture, 10 ml volume of 50% Glutathione Sepharose 4B (Macherey-Nagel, Düren Germany) was pre-equilibrated in wash buffer B (see section 2.1.10).

The cleared supernatant was added to the equilibrated gel of Glutathione Sepharose 4B. The suspension was incubated on an orbital shaker for 3 h at 4 ^{o}C to allow proper binding to the matrix. A disposable polypropylene column (5 ml bed volume from Thermo Scientific) with a filter frit was fixed on a column holder, and the suspension was added to it. After the gel settled at the bottom of column, it was washed with 30 bed volumes of wash buffer B by gravity flow. The column outlet was closed with a cap, and 1.5 ml wash buffer B supplemented with 1 unit/ml thrombin (Serva Electrophoresis) was added into the column. The inlet of the column was closed, and the suspension was allowed to mix on an orbital shaker overnight at 4 ^{o}C. Thrombin cleaved the GST-hGMPK fusion protein releasing hGMPK protein into the solution and leaving the GST tag bound to the Glutathione Sepharose 4B. The buffer was allowed to flow through the column, and the eluate containing pure hGMPK was collected. The elution step was repeated at least twice, and the collected eluates were pooled. Protein concentration was determined by the Bradford dye-binding assay, and purity of the sample was tested on 12% SDS-PAGE. To get the enzyme in highly pure and homogeneous form, gel filtration chromatography was performed as described above in section 2.2.5.1.

The activity of hGMPK was measured by the NADH-dependent spectroscopic assay, and by a novel electrochemical detection assay.

2.2.6 NADH-dependent spectroscopic assay

The activity of hGMPK was determined by the standard NADH-dependent enzyme-coupled assay using a JASCO V-650 UV-Vis spectrophotometer [99]. The formation of ADP and GDP by hGMPK was coupled to two additional reactions catalyzed by pyruvate kinase (PK) and lactate dehydrogenase (LDH), respectively. As shown in the reaction scheme below, each mole of phosphoryl group transferred from ATP produces two moles of NDPs, and consequently two moles of NADH are oxidized to NAD$^+$ [100]. The absorbance was monitored at 340 nm because NADH absorbs light at 340 nm whereas NAD$^+$ does not. The time-dependent decrease in absorbance at 340 nm associated with NADH oxidation was measured spectrophotometrically.

$$\text{ATP + GMP} \xrightarrow{\text{hGMPK}} \text{ADP + GDP}$$

$$\text{ADP + GDP + 2PEP} \xrightarrow{\text{PK}} \text{ATP + GTP + 2 Pyruvate}$$

$$\text{2 Pyruvate + 2NADH + 2H}^+ \xrightarrow{\text{LDH}} \text{2 L-Lactate + 2NAD}^+$$
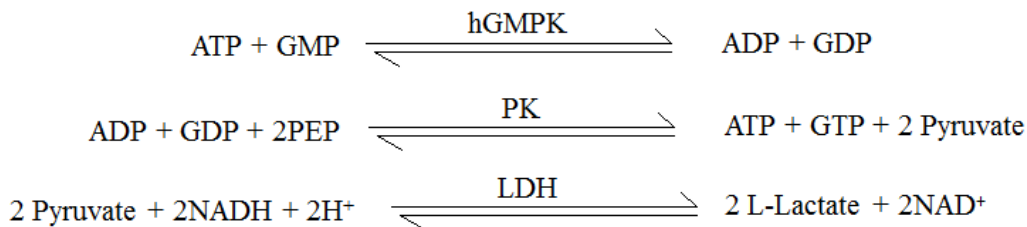
All measurements were performed at 25 ^{o}C in buffer A (see section 2.1.10) containing 100 mM Tris, pH 7.5, 100 mM KCl, and 10 mM $MgCl_2$. The hGMPK was used in 18 nM concentration in a reaction volume of 1 ml. For steady-state kinetics, the concentration of the physiological substrate, GMP, was varied in the range of 0-20 K_m. Enzyme activity was calculated as given below.

$$\text{Volume activity} = \frac{\Delta A_{340}/\Delta t}{\varepsilon \times d} \qquad \text{[U/ml]}$$

$\Delta A_{340}/\Delta t$ = change in absorbance per unit time
ε = molar extinction coefficient of NADH at 340 nm is 6.22 mM^{-1}cm^{-1}
d = light path (cm)

$$\text{Specific activity} = \frac{\text{volume activity (U/ml)}}{\text{concentration of enzyme (mg/ml)}} \qquad \text{[U/mg]}$$

$$k_{cat} = \frac{\mu mol \text{ (substrate)}}{\mu mol \text{ (enzyme) x s}} \qquad \text{[s}^{-1}\text{]}$$

Turnover number (k_{cat}) of the enzyme can be calculated from the specific activity if the molecular weight of the enzyme is known (e.g. hGMPK ~22 kDa).

All data points were the means of duplicate or triplicate measurements. Specific activity or k_{obs} (steady-state rates) values were plotted against the respective substrate concentrations by Gnuplot 5.0 software using the Michaelis-Menten equation.

$$V = \frac{V_{max}[S]}{K_m + [S]} \qquad\qquad (1)$$

The kinetic parameters, K_m and k_{cat}, were calculated from the plot.

2.2.7 Electrochemical detection assay

A new approach has been developed to detect nucleotide kinase-catalyzed reactions based on the light-triggered electrochemical sensing of NADH in a three-step coupled-assay as shown schematically below. We demonstrated a proof of biosensor for hGMPK immobilized on CdS/ZnS quantum-dot modified gold electrode.

In this work, which was done in collaboration with Prof. Dr. Wolfgang Parak and collaborators from Philipps University of Marburg, modified gold electrodes were used with self-assembled monolayer of stilbenedithiol as a substrate for spin coating of the semiconducting CdS/ZnS quantum dots (QDs) [87-89]. The CdS/ZnS QDs serve as switch for light-controlled detection of NADH which is an electron carrier. Electrochemical experiments were performed with a homemade potentiostat in 1 ml electrochemical cell with three-electrode arrangement consisting of the QDs working electrode, an Ag/AgCl reference electrode and a platinum wire as the counter electrode. Light pulses were produced periodically from a light source of Xenon arc lamp (emission spectrum, λ_{em} = 300-700 nm) to illuminate working electrodes from a fixed distance. All measurements were performed at room temperature in 100 mM HEPES buffer pH 7.5 containing 100 mM KCl and 20 mM MgCl$_2$. The following constituents of the coupled assay were used in fixed amounts: 18 nM of hGMPK, 4 mM ATP, 2 mM PEP (phosphoenolpyruvate), 12 units of PK (pyruvate kinase), 1.2 mM NADH, and 15 units of LDH (lactate dehydrogenase). The first step in the three step-assay (this reaction scheme is identical to that of the NADH-dependent spectroscopic assay, see section 2.2.6) was catalyzed by hGMPK in the presence of two helper enzymes in the reaction mixture, pyruvate kinase (PK) and lactate dehydrogenase

(LDH), resulted into the oxidation of NADH to NAD$^+$. The photocurrent "I" was recorded during light on-periods at fixed bias potential, U = +50 mV, in the absence and presence of hGMPK with different concentrations (50-1600 μM) of GMP substrate. Each experiment was performed in triplicate. It was observed that the change in photoelectric current was directly proportional to the GMP concentration.
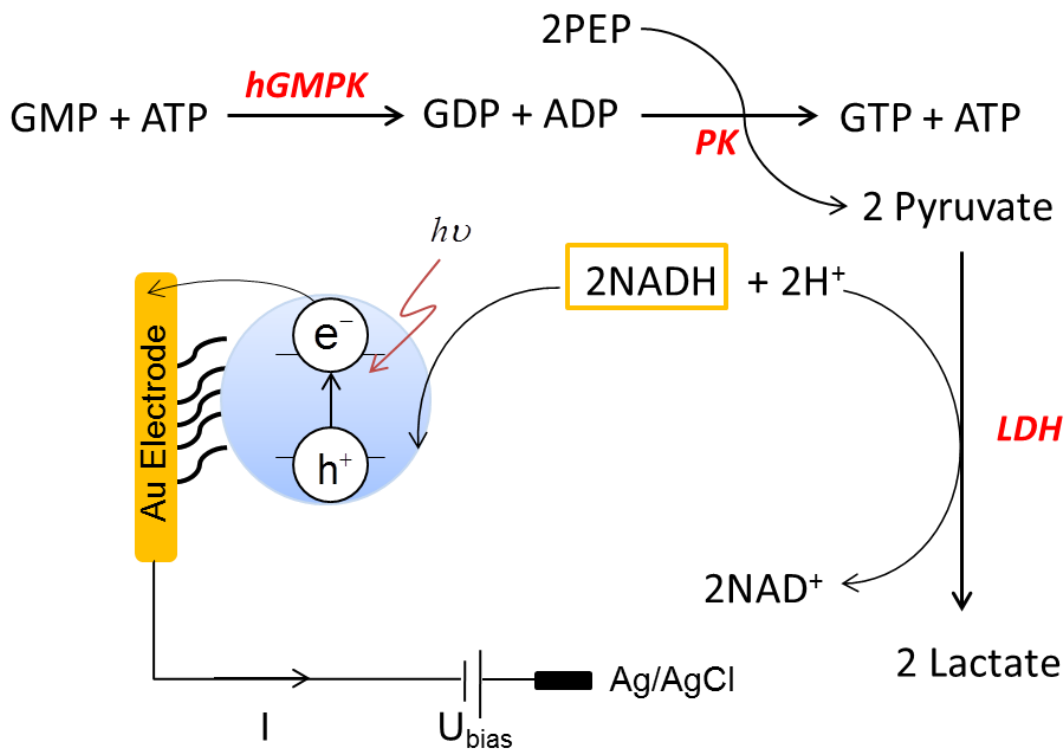
Fig. 2.2. Schematic representation of the electrochemical detection assay. GMP which is the substrate for hGMPK was detected indirectly by coupling it to NADH in the third redox reaction. NADH is an electron carrier and was sensed electrochemically in the reaction mixture at a constant bias voltage U applied to the Au (gold) electrode versus the Ag/AgCl reference electrode. The Au electrode was coated with CdS/ZnS (cadmium sulfide and zinc sulfide) quantum dots (QDs) via StDT (stilbenedithiol). The illumination of light on QDs generates electron hole pairs and causes the transfer of electrons from NADH to QDs which was detected as oxidation-dependent photocurrent. However, the addition of hGMPK to the reaction mixture depletes NADH in a GMP concentration-dependent way and therefore the photocurrent also changes. The CdS/ZnS QDs interlayer between the electrode and the redox system is used for the light triggered readout of the electron transfer reaction with the electrode [87, 88].

2.2.8 A novel spectrophotometric and fluorometric enzyme-coupled assay for hGMPK

We developed a spectrophotometric and fluorometric assay for deoxyribonucleoside and nucleotide kinases which can be used in absorbance mode as well as in fluorescence mode. The assay is based on the use of the nonfluorescent compound Amplex Red which is enzymatically oxidized to resorufin that has excellent absorbance as well as fluorescence properties. We used hGMPK and its physiological substrate GMP to optimize the assay. There are four coupled steps in the assay as shown below. In the first step, GMP is phosphorylated to GDP by hGMPK using ATP as a phosphoryl group donor. In the second step GDP and ADP are converted to GTP and ATP by pyruvate kinase (PK) using phosphoenolpyruvate (PEP) as a phosphate donor resulting in the formation of pyruvate. In the third step, pyruvate is converted by pyruvate oxidase (lpPOX) producing hydrogen peroxide (H_2O_2), CO_2, and acetyl phosphate. In the fourth step, H_2O_2 is used by horseradish peroxidase (HRP) to oxidize Amplex Red to resorufin which has higher absorbance at approximately 570 nm and gives fluorescence at excitation wavelength, $\lambda_{ex} = 568$ nm, and emission wavelength, $\lambda_{em} = 584$ nm [116].

$$\text{ATP + GMP} \xrightarrow{\text{hGMPK}} \text{GDP + ADP}$$

$$\text{GDP + ADP + 2PEP} \xrightarrow{\text{PK}} \text{2 Pyruvate + ATP + GTP}$$

$$\text{2 Pyruvate + 2Pi + 2O}_2 \xrightarrow{\text{lpPOX, TPP, FAD}} \text{2H}_2\text{O}_2 + \text{2CO}_2 + \text{2 Acetyl phosphate}$$

$$\text{2H}_2\text{O}_2 + \text{2 Amplex Red} \xrightarrow{\text{HRP}} \text{2 Resorufin}$$
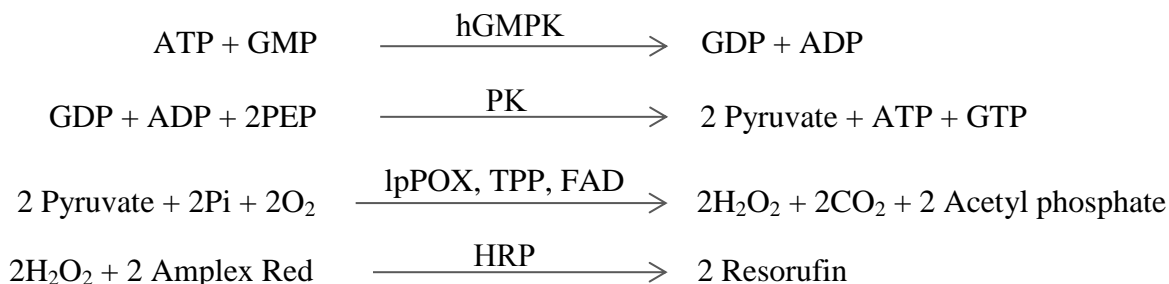
For establishing the assay, pyruvate oxidase (lpPOX) from *Lactobacillus plantarum* was cloned, and recombinantly produced in *E.coli* as discussed below in section 2.2.8.1 and 2.2.8.2. The assay was optimized in the absorbance mode at 570 nm. All four reactions in the coupled assay were optimized in a stepwise manner starting from reaction four, then combining reactions 3 and 4, afterwards 2, 3 and 4, and finally all four reactions. In our assay conditions, the activities of the three auxiliary enzymes (PK, lpPOX and HRP) were kept higher than the enzyme of interest which catalyzes the first reaction. All measurements were performed at 25 ^{o}C in buffer-H (50 mM HEPES pH 7.2, 10 mM MgCl$_2$, 50 mM KCl and 25 mM potassium phosphate). The final reaction mixture contained 2 mM ATP, 5 nM hGMPK, 2 mM PEP, 1 unit of pyruvate kinase

(PK), 1.2 µM of pyruvate oxidase (lpPOX), 0.2 mM thiamine pyrophosphate (TPP), 10 µM flavin adenine dinucleotide (FAD), 25 µM Amplex Red (resorufin, $\varepsilon = 5.4$ x 10^4 $M^{-1}cm^{-1}$ at 570 nm) and 0.1 µM HRP in 200 µl total volume. The linearity of the reaction velocity in dependence of the hGMPK concentration was tested, and was found to be fulfilled in the concentration range of 0.5 nM to 18 nM without being limited by other reactions. For kinetic measurements, 5 nM of hGMPK was used, and the concentration of GMP was varied up to 20 K_m. Steady-state turnover rates (k_{obs}) were calculated and the values were fit to the Michaelis-Menten equation (see equation 1 in section 2.2.6) using Gnuplot 5.0 software. To validate our assay, the kinetics of hGMPK was also determined by the standard NADH-dependent enzyme-coupled assay (see section 2.2.6). It was found that the kinetic parameters determined by both assays were in good agreement. Thus, our new assay is highly authentic and has the advantage to be applicable both in the absorbance and in the fluorescence mode.

2.2.8.1 Cloning of pyruvate oxidase

The plasmid containing the gene for pyruvate oxidase from *Lactobacillus plantarum* (lpPOX) was obtained from Prof. Dr. Kai Tittmann (Georg August University Goettingen, Germany). The open reading frame (ORF) of lpPOX (UniProt accession number P37063, pyruvate oxidase, POX or pox5) which is 1809 bp was amplified via polymerase chain reaction (PCR). NdeI and BamHI sites were incorporated in the two oligonucleotides (forward 5´-GGGAATTCCATATGGTTA TGAAACAAACAAAACAAACTAAC-3´ and reverse 5´-CGCGGATCCTTAAAACCCACC CTGTCCAATTTG-3´) targeting the 5′ and 3′ ORF ends, respectively. The PCR product was gel-purified, digested with NdeI and BamHI-HF (New England Biolabs), and ultimately ligated into the pJC20HisN vector using T4 DNA ligase (New England Biolabs). The ligation mixture was used to transform XL1-Blue cells. Positive clones were identified by colony PCR and by restriction digestion with NdeI and BamHI-HF enzymes. The entire gene insert was sequence verified (Seqlab, Goettingen). The final construct includes an N-terminal decahistidine tag for affinity purification. For the overexpression of lpPOX, BL21(DE3)pLysS cells were used.

2.2.8.2 Expression and purification of pyruvate oxidase

A starter culture was prepared by inoculating 20 ml LB with a single colony of BL21(DE3) pLysS carrying the lpPOX-containing plasmid and incubated at 37 °C with rapid shaking (200-

250 rpm). One liter of lactose-containing auto-inducing media was inoculated with the starter culture. It was incubated at 37 °C with rapid shaking until the optical density at 600 nm reached about 0.7, and was then transferred to 21 °C for overnight auto-induction. Cells were harvested, and the pellet was resuspended in lysis buffer (100 mM potassium phosphate pH 6.0, 300 mM NaCl, 0.1 mM thiamine pyrophosphate (TPP), 0.01 mM FAD, 1 mM $MgSO_4$, 5 mM DTT, and cocktail of protease-inhibitors), and lysed by sonication. The lysate was cleared by centrifugation at 10,000 x g for 1 h at 4 °C and subjected to batch purification by affinity chromatography. One gram of Protino Ni-IDA resin (Macherey-Nagel) was added to the supernatant in a centrifugation tube, and the suspension was agitated on an orbital shaker for 1 h at 4 °C. The resin was washed four times with 40 ml of buffer A (100 mM potassium phosphate pH 6.0 containing 300 mM NaCl, 0.1 mM TPP, 0.01 mM FAD, 1 mM $MgSO_4$, 10 mM imidazole, and 1 mM DTT) each time with 25 min incubation on an orbital shaker at 4 °C. The resin was transferred to an empty chromatography column (5 ml column bed, Thermo Scientific), and the protein was eluted with buffer B (100 mM potassium phosphate pH 6.0, 150 mM NaCl, 0.1 mM TPP, 0.01 mM FAD, 1 mM $MgSO_4$, 250 mM imidazole, and 1 mM DTT). All fractions containing high amount of pyruvate oxidase were pooled yielding 50 mg of His_{10}-lpPOX protein, as determined by the Bradford dye-binding assay. Imidazole was removed from the sample by dialysis against buffer C (100 mM potassium phosphate pH 7.0, 150 mM NaCl, 0.1 mM TPP, 0.01 mM FAD, and 1 mM $MgSO_4$) using 2 kDa MWCO membrane. The enzyme was further purified to homogeneity by gel filtration chromatography using a Superdex 200 column (GE Healthcare) pre-equilibrated with buffer C. It eluted as a homo-tetramer as indicated by the elution profile of marker proteins (BioRad gel filtration standard). The purified protein was concentrated by 3 kDa MWCO ultracentrifugal filter, aliquoted, and stored at -80 °C. Protein concentration was determined by the Bradford dye-binding assay and its purity was tested by 12 % SDS-PAGE.

2.2.9 hGMPK-catalyzed reactions in polyelectrolyte containers of various shapes and sizes

2.2.9.1 Synthesis of calcium carbonate particles

Calcium carbonate ($CaCO_3$) particles of different shapes (rhomboidal, ellipsoidal, spherical) were synthesized by rapid mixing of equal volumes of $CaCl_2$ and Na_2CO_3 solutions added in specific molar ratios (1:1, 1:10 or 10:1) in the presence of varied amounts of ethylene glycol from 66.6 % to 80%. The reaction mixture was stirred at 500 rpm for 30 min, 60 min, or 90 min

at room temperature. The pH of $CaCl_2$ and Na_2CO_3 solutions were 7.4 and 9.5, respectively. After precipitation, the particles were recovered by centrifugation at 2,000 rpm for 2 min. The supernatant was discarded, and the particles were washed 3 times in deionized water and in 50% ethanol to remove residual ethylene glycol and salts. The particles were dried in an oven at 70 °C for 1-2 h and were stored at room temperature.

2.2.9.2 Protein loading

Proteins can be loaded into particles either by the coprecipitation method or by physical adsorption [136]. In the coprecipitation method, a specific amount of protein is dissolved in the $CaCl_2$ solution prior to mixing with Na_2CO_3 solution for particle synthesis. When the precipitation process is allowed to occur, the protein molecules are entrapped within the growing microparticles [131]. In contrast, in the physical adsorption method, the preformed particles are resuspended in the protein solution and put on a shaker at 600 rpm for ~15 min. The protein molecules are adsorbed in the pores of $CaCO_3$ particles.

2.2.9.3 Capsule fabrication

The synthetic polyelectrolytes PAH (Poly(allylamine hydrochloride)) and PSS (Poly(sodium 4-styrenesulfonate)) were dissolved in 20 mM Tris buffer pH 7.4 containing 150 mM NaCl at a final polymer concentration of 2 mg/ml [137]. Fabrication of microcapsules is shown schematically in Fig. 2.3. 1 ml of PAH was added to about 10 mg of particles, mixed by flicking few times, and agitated in the ultrasonic bath for 1 min. The suspension was then incubated by constant shaking at 1,000 rpm for 15 min. It was centrifuged at 2,000 rpm for 2 min, and the supernatant was discarded. Microparticles were washed three times in the Tris buffer. Then next layer of PSS was deposited in the same way as PAH. The particles were coated with the desired number of polyelectrolyte layers. To obtain hollow capsules, the $CaCO_3$ core was dissolved by dialysis against 20 mM Tris buffer pH 7.4 containing 150 mM NaCl and 20-50 mM EDTA. The hollow microcapsules with loaded hGMPK were used for activity measurements (see section 2.2.6) and for further analysis. Microcapsules were stored at 4 °C for long term use.
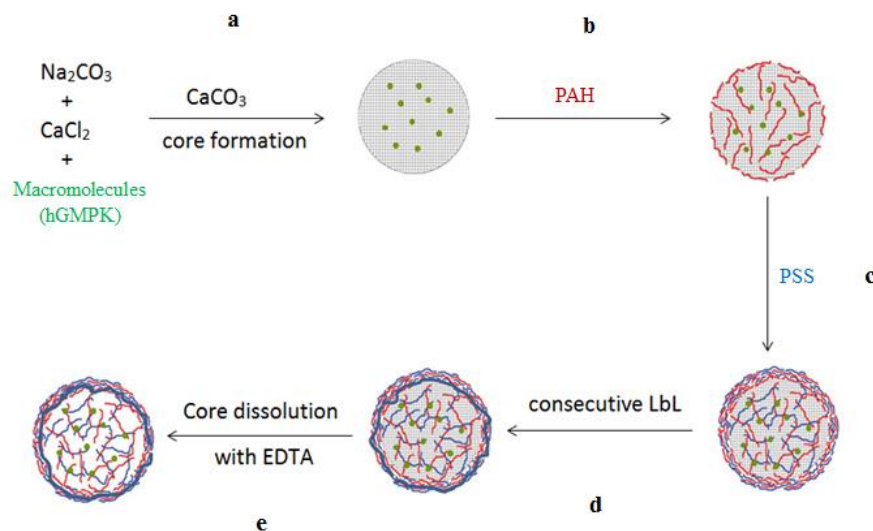
Fig. 2.3. Scheme for fabrication of microcapsules and their packaging with macromolecules (hGMPK). (a) The protein was coprecipitated during the synthesis of calcium carbonate ($CaCO_3$) microparticles from calcium chloride ($CaCl_2$) and sodium carbonate (Na_2CO_3); (b-d), Microparticles were encapsulated in polyelectrolyte layers (PAH and PSS) in a consecutive way by applying the LbL (layer-by-layer) technique. (e) The $CaCO_3$ core was dissolved by treating the microcapsules with EDTA.

2.2.10 Site-directed mutagenesis of hGMPK

The *E.coli* expression plasmid containing the hGMPK insert, hGMPK[pET-14bSUMOΔThr], was used as a template for the introduction of eight amino acid substitutions at different positions in the active site (S37A, S37Y, S37C, S37P, Y81F, T83S and T83A) and in the hinge region (N42P) of the hGMPK molecule by the QuikChange site-directed mutagenesis procedure [101] using KAPAHiFi DNA polymerase (Fig. 2.4). Sixteen mutagenic oligonucleotides containing the desired point mutations were designed for this purpose, and were synthesized by IBA GmbH (Goettingen). All oligonucleotides were 25-45 bases in length, Tm greater or equal to 78 ^{o}C, and GC contents in the range of 40-60%. Each 50 µl reaction mixture contained 20 ng of hGMPK[pET-14bSUMOΔThr] template (~5.5 kb), 1X KAPAHiFi Fidelity or GC buffer, 0.3 mM dNTP Mix, 0.3 µM forward and reverse primers, and 1 unit of KAPAHiFi DNA polymerase. The cycling parameters used were initial denaturation at 95 ^{o}C for 2 min, followed by 16 cycles of denaturation at 95 ^{o}C for 20 s, annealing of primers at 57 ^{o}C for 30 s, and extension at 72 ^{o}C for 3 min and 20 s. The final extension was at 72 ^{o}C for 10 min. To check for sufficient amplification, 10 µl of the PCR product was loaded on 1% agarose gel. Once verified the

amplified product, 1 µl (10 units) of DpnI restriction enzyme was added to the remaining 40 µl PCR product, mixed well, and incubated at 37 °C for 2 h. The DpnI enzyme only digests the methylated template DNA (non-mutated), and does not cleave the PCR-amplified non-methylated plasmid DNA (mutated). To remove the restriction enzyme and other contamination from the PCR product, it was cleaned by passing through the NucleoSpin Gel and PCR Clean-up columns according to the manufacturer's protocol. For transformation of competent XL1-Blue cells, 5 µl of the purified PCR product was used according to the protocol as described in section 2.2.4.5. After transformation, the XL1-Blue competent cells repair the nicks in the mutated plasmid. 50 µl of the transformed cells were spread on pre-warmed ampicillin-containing selection plates and were incubated at 37 °C overnight. For plasmid purification, two colonies were picked from each agar plate (out of 8 plates) and were used to inoculate 10 ml LB medium containing ampicillin, and incubated at 37 °C overnight with vigorous shaking at 250 rpm. The mutated plasmids were purified by NucleoSpin Plasmid kit according to the manufacturer's protocol. To verify the desired eight mutations in the hGMPK inserts, one set of the corresponding plasmids were sent to Seqlab (Goettingen, Germany) for DNA sequencing. For protein expression and purification, the sequence-verified eight mutants of hGMPK were transformed into BL21(DE3)pLysS cells and were streaked on ampicillin- and chloramphenicol-containing agar plates. The plates were incubated at 37 °C overnight. All mutants were expressed in autoinducing media, and the proteins were purified by affinity chromatography using Protino Ni-IDA resin as mentioned in section 2.2.5.1. Their purity was tested by 12 % SDS-PAGE, and the enzymes were kinetically characterized by using the NADH-dependent enzyme-coupled assay as described in section 2.2.6.
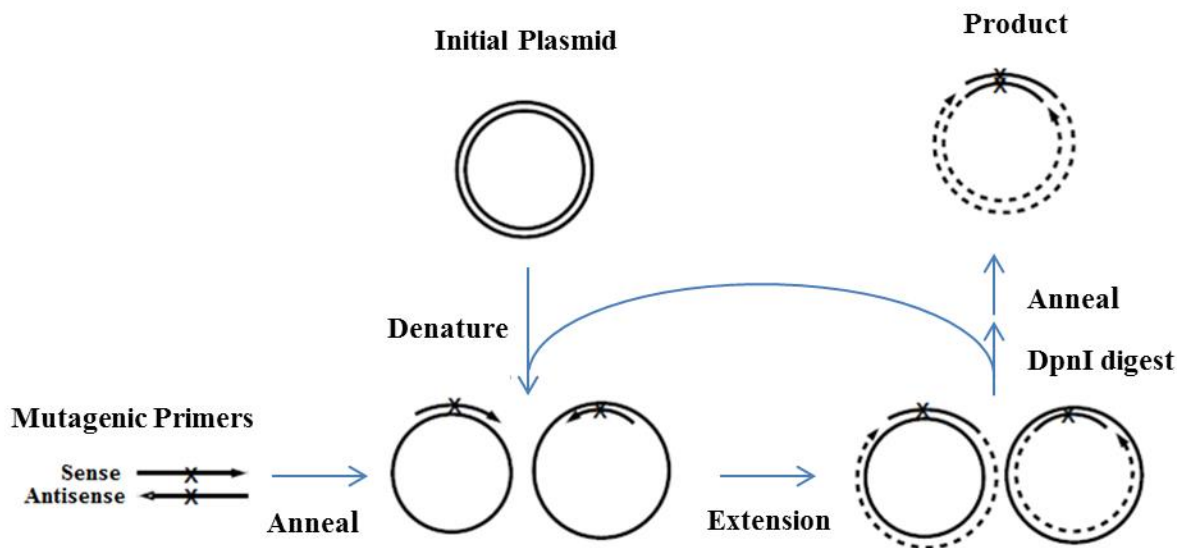
Fig. 2.4. QuikChange site-directed mutagenesis. Schematic overview of the QuikChange site-directed mutagenesis method, as described in the KAPA HIFi and Stratagene QuikChange protocol [101].

2.2.11 Optimizing the production of isotopically labeled (^{15}N, ^{15}N/^{13}C) human GMPK for NMR structure elucidation

Structure elucidation by NMR requires a large quantity and high quality of isotopically labeled recombinant protein (^{15}N, ^{15}N/^{13}C), the production of which is costly and time-consuming.

A low yield of hGMPK was observed in minimal growth media using the glycerol stock made with a selected colony. Therefore, the double-colony selection procedure was used to optimize the yield which is one of the most important factors for high-level protein production using high-cell-density bacterial expression methods [102]. Using this optimized protocol the hGMPK yield was increased 8-fold utilizing the unlabeled M9 minimal growth medium (see section 2.1.9.1). The hGMPK[pET-14bSUMOΔThr] construct was transformed into the BL21(DE3)pLysS cells as described in section 2.2.4.5, and an ampicillin (100 μg/ml)/chloramphenicol (20 μg/ml)-containing agar plate was streaked and incubated at 37 °C overnight. 5 ml LB containing ampicillin/chloramphenicol was inoculated with a single colony from the agar plate and cells were grown up to an optical cell density at 600 nm (OD_{600}) of 0.7. Approximately 5 μl of the cell suspension was spread onto an agar plate with the corresponding antibiotics and incubated at 37 °C overnight. For 1st colony selection, four colonies were selected from the LB agar plate, and

each colony was inoculated into 5 ml LB containing ampicillin and chloramphenicol. They were incubated overnight at 37 $^{\circ}$C with vigorous shaking at 250 rpm. 20 ml LB was inoculated with the 2 ml overnight cultures and incubated at 37 $^{\circ}$C shaken at 250 rpm. Upon reaching OD_{600} of 0.7, aliquots of 10 µl from each flask was stored at 4 $^{\circ}$C for streaking purposes, and the rest of the cells were harvested by centrifugation at 5,000 x g for 25 min. The cell pellets were washed in 20 ml M9 salt solution (see section 2.1.9.1) to remove the LB medium left over and pelleted by centrifugation. For overexpression in minimal medium, each cell pellet was resuspended in 5 ml (4-fold smaller volume) of M9 minimal medium and grown for 1 h at 37 $^{\circ}$C. After 1 h incubation, protein expression was induced by adding 1 mM IPTG, and the cultures were left on vigorous shaking for 4 h at 37 $^{\circ}$C. After 4 h induction, 200 µl of cell suspension was spun down at 6,000 x g for 5 min, and the supernatant was discarded. The cell pellets were resuspended in 100 µl Laemmli sample buffer (see section 2.1.10) and heated at 95 $^{\circ}$C for 5 min. SDS-PAGE was carried out for all four samples to check their levels of hGMPK expression. Only the colony which displayed the highest level of expression was selected for 2nd round of selection. The 10 µl aliquot of highly expressed colony was used to streak an agar plate for overnight incubation at 37 $^{\circ}$C. The procedure followed for the 2nd selection was the same as for the 1st selection. A permanent glycerol stock was made for the chosen colony from the 2nd selection round and was used for any future overexpression of single-labeled (^{15}N) and double-labeled (^{15}N and ^{13}C) production of hGMPK in labeled M9 minimal growth media.

The optimized protocol developed for obtaining higher yield of isotope-labeled hGMPK was used to grow an overnight culture of hGMPK-expressing cells in 200 ml LB containing ampicillin and chloramphenicol at 37 $^{\circ}$C shaken at 250 rpm [103]. The overnight 200 ml LB culture was added to 3,800 ml LB containing selection antibiotics making a final 4-litre culture. It was grown at 37 $^{\circ}$C with rapid shaking until OD_{600} of ~0.7 was attained. The cells were harvested by centrifugation at 5,000 x g for 30 min. For excluding all nitrogen and carbon sources, the cells were washed and pelleted using M9 salt solution. The cell pellets were resuspended in 1 liter isotopically labeled (^{15}N labeled, or ^{15}N and ^{13}C double-labeled) minimal growth medium at cell concentration fourfold (4X) greater relative to the cultures grown in LB medium. The M9 minimal medium employs ^{13}C glucose for carbon labeling, and ^{15}N ammonium chloride for ^{15}N labeling. The minimal growth medium was simplified by replacing the commercially available Basal Vitamins Eagle Media [103] with thiamine vitamin only. The cells

were incubated at 37 ^{o}C on vigorous shaking for 1 h to allow the recovery of growth and clearance of unlabeled metabolites. After 1 h, the protein expression was induced by adding 1 mM IPTG and incubating the cells for 4 h at 37 ^{o}C on constant shaking. The cells were harvested by centrifugation at 5,000 x g for 30 min and were either stored at -20 ^{o}C, or directly processed for protein purification as described in section 2.2.5.1.

2.2.12 Enhancing cytotoxicity of 6-thioguanine by expressing human GMPK

2.2.12.1 Cell-culture and stable HEK293 cell line

To culture HEK293 cells, the standard complete DMEM medium was used which is DMEM supplemented with 10% heat-inactivated FCS and L-glutamine and maintained at 37 °C in humidified atmosphere containing 5% CO_2. The HEK293 cells were seeded into 10 cm dishes and transfected with hGMPK[pEGFP-C1] and hGMPK[pEGFP-N1] using FuGene1 HD transfection reagent (Roche Mannheim, Germany) according to the manufacturer's instructions. The standard complete DMEM medium was exchanged to complete DMEM supplemented with 900 µg/ml of G418. The cells were observed after 7–10 days under an inverted light-microscope, and transferred to 12-well plates. After 2–3 days, the stable colonies were monitored using a fluorescence-microscope and further sub-cultured. The established stable cell lines were maintained in complete DMEM medium containing 300 µg/ml of G418 [85].

2.2.12.2 MTT cell proliferation/survival assay

For 6-thioguanine (6-TG) dose-response analysis, HEK293 stable cell lines were seeded in 24 well plates. 6-thioguanine (stock solution; 1.5 mM in sterile cell culture medium) in the range of 0.001–1000 µM was added in triplicate for each concentration. After 48 h, the MTT proliferation assay was performed according to the supplier's instructions (Sigma-Aldrich). MTT stock solution was prepared by dissolving MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) in RPMI-1640 without phenol red to a final concentration of 5 mg/ml. The stock solution was filtered through 0.2 µm filter and was added to each well being assayed, equal to one-tenth of the original culture volume, and incubated for 30 min to 3 h. At the end of the incubation period, the medium was removed, and the converted dye was solubilized in acidic isopropanol (0.04-0.1 N HCl in absolute isopropanol). Absorbance of converted dye was measured at a wavelength of 570 nm with background subtraction at 630–690 nm. The

measurements were performed by using a JASCO V-650 UV-Vis spectrophotometer. Results obtained at different concentrations of 6-TG were compared to the wells without 6-TG (control), and the percentage cell survival was plotted against concentration of 6-thioguanine.

2.2.12.3 Confocal microscopy

Cells were grown on Lab-Tek Chamber slides (Thermo Fisher Scientific), fixed with 4% paraformaldehyde for 10 min at room temperature, and incubated with DAPI (300 nM) for 2 min at room temperature. The cover-slips were mounted in Vectashield medium (Vector Laboratories). The images were taken with a 40x objective on a Leica SP5 confocal microscope.

2.2.12.4 Western-blot analysis

The stable HEK293 cells cultured in 24-well plates were washed with 1X PBS and lysed in lysis buffer A (see section 2.1.10). Cells were immediately scraped off the plate, transferred to microcentrifuge tubes and kept on ice. For complete cell lysis, the extract was sonicated 3 times for 10-15 s to shear the DNA and reduce sample viscosity. The lysates were spun down at 16,000 x g for 20 min in a 4 ^{o}C pre-cooled centrifuge. Supernatants were transferred to fresh microcentrifuge tubes and placed on ice. Protein concentration was determined by the Bradford dye-binding assay. For SDS-PAGE, 20 µg of each sample was mixed with equal volume of 2X-Laemmli sample buffer and heated at 95 ^{o}C for 5 min. The samples (about 20 µg of protein) were loaded on a 12.5% SDS-gel and run at 100-150 V for about 1 h. To transfer the proteins from SDS-gel to a nitrocellulose membrane, the transfer sandwich was assembled and the cassette was placed in the transfer tank. An ice block was placed on one side in the tank. Transfer buffer (25 mM Tris pH 8.3, 190 mM glycine and 20% methanol) was added to the transfer tank, and protein transfer was performed at 10 mA constant current overnight in a cold room. The blot was rinsed in Millipore water and stained in Ponceau S solution (0.2% (w/v) Ponceau S and 5% glacial acetic acid) to check the transfer quality. Ponceau S stain was rinsed off by washing three times in TBST (Tris-buffered saline with Tween 20) solution (20 mM Tris pH 7.5, 150 mM NaCl and 0.1% Tween 20). The nitrocellulose membrane was incubated in blocking buffer (5% (w/v) skimmed milk in TBST) for 1 h at room temperature. The membrane was incubated in the primary antibody solution (anti-GFP IgG antibody in 1% milk in TBST) for 2–3 h at room temperature. The blot was washed three times in TBST and incubated in the horseradish

peroxidase-conjugated anti-rabbit IgG in 1% milk for 1 h at room temperature. It was rinsed 3-5 times in TBST. The membrane was incubated for one minute in a 1:1 mixture (1 to 2 ml each) of the two ECL solutions (Roche) and the chemiluminescent signals were captured by the Lumi-Imager workstation (Boehringer).

2.2.13 Expression and purification of human mitochondrial thymidine kinase

The synthetic and codon-optimized DNA of the long isoform (UniProt O00142-1, 265 amino acids) of human mitochondrial thymidine kinase (hTK2) was obtained from GeneArt (Life Technologies). It had a prospective 38-amino acid N-terminal signal sequence (predicted by MITOPROT proteomics tool) for translocation into mitochondria. In order to remove the N-terminal signal peptide and to study the effect of truncation on the localization, solubility, and activity of the enzyme, six truncated forms were generated using PCR amplification. NdeI and BamHI cloning sites were incorporated at the 5´ and 3´ ends of all fragments, respectively. Open reading frames (ORF), and the corresponding oligonucleotides used for the amplification of full-length hTK2 and its truncated forms are shown in Table 2.2 (section 2.1.2) and Table 2.4.

Table 2.4 Truncated forms of hTK2

hTK2	Description	Oligonucleotides
hTK2	Synthetic full-length hTK2, 795 bp ORF, 265 amino acids (aa) in length	hTK2-F and hTK2-R
hTK2-Δ44N	N-terminal 44 amino acids were truncated, 666 bp, 222 aa	hTK2-Δ44N-F and hTK2-R
hTK2-Δ44N/Δ8C	N-terminal 44 and C-terminal 8 amino acids were truncated, 642 bp, 214 aa	hTK2-Δ44N-F and hTK2-Δ8C-R
hTK2-Δ44N/Δ25C	N-terminal 44 and C-terminal 25 amino acids were truncated, 591 bp, 197 aa	hTK2-Δ44N-F and hTK2-Δ25C-R
hTK2-Δ50N	N-terminal 50 amino acids were truncated, 648 bp, 216 aa	hTK2-Δ50N-F and hTK2-R
hTK2-Δ50N/Δ8C	N-terminal 50 and C-terminal 8 amino acids were truncated, 624 bp, 208 aa	hTK2-Δ50N-F and hTK2-Δ8C-R
hTK2-Δ50N/Δ25C	N-terminal 50 and C-terminal 25 amino acids were truncated, 573 bp, 191 aa	hTK2-Δ50N-F and hTK2-Δ25C-R

The synthetic full-length hTK2 containing plasmid was used as a template in all cases. The Phusion High-Fidelity DNA polymerase (New England BioLabs) was used according to the manufacturer's instructions. PCR amplification, restriction digestion and ligation reactions were carried out as described in sections 2.2.4.1 to 2.2.4.4. For improved expression and solubility, PCR fragments were cloned in a variety of *E.coli* expression plasmids having different fusion tags (Table 2.5). The positive clones of the six truncated hTK2 variants were confirmed by restriction enzyme digestion and sequence analysis (Seqlab, Goettingen). As it was found that hTK2 has little solubility and mostly forms inclusion bodies under native folding conditions when expressed in *E.coli* cells, different expression and purification conditions were tested for optimizing the solubility of hTK2. For that purpose, various *E.coli* expression strains and several experimental parameters were used as shown in Table 2.6 below.

After a series of experiments, a protocol was optimized for the expression and purification of hTK2 in soluble and active form as described here in detail. A starter culture was prepared by inoculating 50 ml LB (having appropriate antibiotics) with a single colony of each hTK2 construct and incubated overnight at 37 ^{o}C with rapid shaking (250 rpm). The starter culture was added to 950 ml LB containing the required antibiotics and incubated at 25 ^{o}C with rapid shaking at 250 rpm until the OD_{600} reached ~0.65. For protein induction, 0.4 mM IPTG was added, and the culture was incubated at 18 ^{o}C for 16 h on constant shaking. The cells were harvested by centrifugation at 4,000 rpm for 30 min at 4 ^{o}C and the pellet was resuspended in lysis buffer B (see section 2.1.10). The cells were lysed by sonication, and the lysate was cleared by centrifugation at 10,000 x g for 1 h at 4 °C. The purity and solubility of hTK2 was tested by SDS-PAGE which indicated improved solubility. For affinity chromatography purification, 500 mg Protino Ni-IDA (Macherey-Nagel, Düren, Germany) was added to the supernatant and mixed on an orbital shaker for 3 h at 4 ^{o}C in the cold room. The resin was washed two times with 40 ml of wash buffer-1 (see section 2.1.10) in a Falcon tube which was rotating on an orbital shaker for 30 min at 4 ^{o}C. The resin was then resuspended in wash buffer-2 (see section 2.1.10) and incubated on an orbital shaker at 37 ^{o}C for 15-30 min. This step was necessary to remove any tightly bound *E.coli* chaperones (e.g. GroEL). Finally, the resin was washed with wash buffer-3 (section 2.1.10) for 30 min on an orbital shaker at 4 ^{o}C. The 60 kDa fusion tag "His_{14}-MBP-SUMObr" was cleaved by treating the fusion enzyme with SUMO-protease (yeast SUMO-

protease to protein molar ratio, 1:100) for 2-3 h at room temperature. After on-beads cleavage, the enzyme was released into the solution.

Table 2.5 *E.coli* expression plasmids used for cloning of hTK2

Plasmid	Fusion tag	Cleavage site	Cloning site
pET-14b	His$_6$	LVPR/GS (Thrombin)	NdeI-BamHI
pET-14bSUMOThr	His$_6$-SUMOThr	SUMO-GG/ (SUMO protease or thrombin)	NdeI-BamHI
pET-14bSUMOΔThr	His$_6$-SUMO	SUMO-GG/	NdeI-BamHI
pET-14bSUMO-PeriPep	His$_6$-SUMO-PL	PL(22 aa)/ (*E.coli* protease)	NdeI-BamHI
pET-14bMBPcyt	His$_6$-MBPcyt	LVPR/GS	NdeI-BamHI
pET-14bMBPperi	His$_6$-MBPperi	LVPR/GS	NdeI-BamHI
pK49	His$_{14}$-MBP-SUMObr	SUMO-GG/ (SUMO protease)	NdeI-BamHI
pGEX-RB	GST	LVPR/GS	NdeI-BamHI
pJC20HisN	His$_{10}$	DDDDK/ (Enterokinase)	NdeI-BamHI
pJC20HisC	His$_6$ (C-terminal)	None	NdeI-BamHI

Table 2.6 Conditions for improved solubility of hTK2

E.coli expression strains	Conditions used for improving solubility and yield of the recombinant hTK2
- BL21(DE3) - BL21(DE3)pLysS - C41(DE3) - Origami B(DE3) - Rosetta-gami B(DE3)	- Liquid media used: LB, TB and autoinducing media - Incubation temperature tested: 16 $^{\circ}$C, 18 $^{\circ}$C, 21 $^{\circ}$C, 25 $^{\circ}$C & 37 $^{\circ}$C - IPTG concentrations used.: 0.1 mM, 0.25 mM, 0.4 mM, 0.5 mM & 1 mM - Duration of induction: 2 h, 3 h, 6 h, 16 h & 24 h - Detergents/reagents used for improving solubility: Triton X-100, urea, CHAPS, sarkosyl & L-arginine

For high purity, hTK2 was further purified by anion exchange chromatography using DEAE-Sepharose FF (GE Healthcare Life sciences, 1 ml prepacked column) connected to the 'ÄKTAprime plus' system (GE Healthcare Life Sciences, Uppsala, Sweden). Buffer A (25 mM HEPES pH 7.5, 4.4 mM $MgCl_2$, 5% glycerol, and 2 mM DTT) was used for column equilibration and binding of hTK2 to the resin. For elution of hTK2, a salt gradient of 0-1 M NaCl was applied in a total volume of 25 ml with flow rate of 0.5 ml/min. Fractions of 250 μl were collected, and their purity was tested by SDS-PAGE. Aliquots with highly pure hTK2 were pooled, concentrated by Amicon ultracentrifugal filter units (3 kDa MWCO). Using this protocol, a final yield of 4-8 mg/l culture was obtained. The activity was determined by the NADH-dependent spectroscopic assay (see section 2.2.6).

2.2.14 Intracellular localization of hTK2

Most mitochondrial proteins are first fully synthesized as precursor proteins in the cytosol and then translocated into mitochondria by a post-translational mechanism. One or more signal sequences direct all mitochondrial precursor proteins to their appropriate mitochondrial subcompartments. The full-length hTK2 (265 aa) has a putative 38-amino acid N-terminal signal sequence that directs its transport into the mitochondrial matrix space. For determining its subcellular localization in mammalian cells such as HEK293 cells, it was cloned into a mammalian vector pEGFP-N1 (4.7 kb). The two truncated forms of hTK2 which lack the N-

terminal signal sequence i.e., hTK2-Δ44N and hTK2-Δ44N/Δ8C were cloned into another mammalian plasmid pEGFP-C1 (4.7 kb). The full-length hTK2 cloned into pEGFP-N1 was expressed as a fusion to the N-terminus of EGFP, while the two N-terminally truncated forms cloned into the MCS of pEGFP-C1 were expressed as fusions to the C-terminus of EGFP. All three plasmids were used to transfect HEK293 cell lines (see section 2.2.12.1), and after 24 h incubation at 37 ^{o}C the cells were fixed for analysis on a confocal microscope (see section 2.2.12.3). The confocal images indicated that the full-length hTK2 was translocated into mitochondria whereas the two truncated forms without the N-terminal signal sequences remained in the cytosol.

2.2.15 Immunodetection of hTK2 by Western blot

For the immunodetection of hTK2 in mammalian cells, polyclonal antibodies were produced against hTK2-Δ44N by Seqlab (Goettingen) by immunizing rabbits. HEK293 cells were transfected (see section 2.2.12.1) with full-length hTK2[pEGFP-N1]), and two truncated constructs i.e., hTK2-Δ44N[pEGFP-C1] and hTK2-Δ44N/Δ8C[pEGFP-C1]. After 24 h incubation at 37 ^{o}C, the cells were analyzed for the expression of hTK2 as EGFP fusion proteins. The cells were lysed and subjected to Western blotting (see section 2.2.12.4) using the polyclonal antibodies-containing serum for detection.

2.2.16 Directed evolution and screening of hTK2 mutants with enhanced activity towards phosphorylation of antiviral and anticancer nucleoside analogs

Directed evolution is a powerful tool for engineering enzymes to enhance their activity, and to explore their structure-function relationships. It is an important way to search for a large combination of sequences for rare molecules displaying specific predetermined functions. Although there are many ways to introduce genetic diversity, error-prone PCR is one of the most common methods for creating combinatorial libraries of a single gene [90]. DNA libraries of hTK2 mutants were generated using random mutagenesis by error-prone PCR. It was aimed to evolve hTK2 for increased activities against anticancer and antiviral nucleoside analogs such as gemcitabine and azidothymidine (AZT). One of the protocols [91] introduces an average error rate of ~3.5% per nucleotide per PCR reaction. Following that protocol, a total of 16 error-prone PCR reactions were carried out using conditions that reduce the fidelity of Taq DNA polymerase

during DNA synthesis. The whole procedure of mutagenesis and the two-step screening in a thymidine kinase-deficient KY895 *E. coli* strain is shown schematically below.

Random mutagenesis of hTK2-Δ44N/Δ8C and screening of tk-deficient *E.coli* cells expressing hTK2 mutants. Testing for efficiency against anticancer and antiviral nucleoside analogs

	Original hTK2-Δ44N/Δ8C gene
	16 Error-prone PCR reactions
	"n" number of mutant hTK2 DNA fragments
	Mutant hTK2 fragments ligated into pGEX-RB vector, "n" number of clones
	Transformation of "n" clones into KY895 (tk-deficient *E.coli* strain)
	Plating transformants on TK selection agar (2% peptone, 0.5 % NaCl, 0.2% glucose, 10 µg/ml 5-fluorodeoxyuridine, 2 µg/ml deoxythymidine, 12.5 µg/ml ribouridine, 1.5% agar and 100 µg/ml ampicillin)
	O/N incubation of the selected colonies in LB/Amp media (>200 mutants)
	Streaking M9 plates containing nucleoside analogs (Gemcitabine, AZT, BVDU, GCV, Ara-T etc) & determined LD$_{100}$

hTK2-Δ44N/Δ8C[pGEX-RB] was used as a template, and the PCR-amplified DNA fragments were treated with NdeI and BamHI-HF restriction enzymes to produce sticky ends. All fragments were cloned into the pGEX-RB vector via NdeI and BamHI cloning sites. The clones were used to transform a thymidine kinase-deficient (F⁻, tdk⁻, 1-ilv) *E.coli* strain, KY895 [92]. The transformed cells were plated onto TK selection plates (2% peptone, 0.5 % NaCl, 0.2% glucose, 10 µg/ml 5-fluorodeoxyuridine, 2 µg/ml deoxythymidine, 12.5 µg/ml ribouridine, 1.5% agar and 100 µg/ml ampicillin) for primary selection screening based on TK complementation efficiency

of mutants [92-94]. The TK-selection plates were incubated at 37 $^{\circ}$C for 24 h. The colonies per plate were counted and re-streaked on TK-selection plates to confirm TK-complementation and phenotypes. About 200 colonies were selected and used for the purification of plasmids having hTK2-Δ44/Δ8C mutant inserts. They were labeled and used for secondary selection screening based on M9 plates containing the following nucleoside analogs: gemcitabine, AZT, Ara-C, GCV, BVDU and Ara-T. One-liter of M9 minimal growth medium (see section 2.1.9.1) was supplemented with 15 g of agar, 4 mg of Ile and Val. About 200 µl of the nucleoside analog solutions with varied concentrations were spread on the M9 agar plates. The plates were streaked with the mutant colonies from the 1st screening. All M9 agar plates were incubated at 37 $^{\circ}$C for 24 h, and the sensitivity of mutants to the nucleoside analogs was identified in the form of no growth. The LD$_{100}$ (the lowest concentration of a nucleoside analog that causes 100% lethality of KY895 strain transformed with hTK2 mutant in pGEX-RB vector) was calculated for all 200 colonies.

2.2.17 Expression and purification of *E.coli* guanosine-inosine kinase

The *E.coli* guanosine-inosine kinase (UniProt P0AEW6-1, gsk, ecGSK, 434 aa) phosphorylates guanosine and inosine nucleosides to their respective monophosphates. The ORF (open reading frame) of full-length ecGSK (1302 bp) and its three truncated forms were PCR-amplified from *E.coli* XL1-Blue genomic DNA using Phusion High-Fidelity DNA polymerase (New England BioLabs). The truncated forms generated were ecGSK-Δ30N (N-terminal 30 aa truncated, 1215 bp), ecGSK-Δ21C (C-terminal 21 aa truncated, 1239 bp), ecGSK-Δ30N/Δ21C (N-terminal 30 aa and C-terminal 21 aa truncated, 1152 bp). Oligonucleotides used for the ecGSK amplification are listed in Table 2.2 (section 2.1.2): the forward oligo ecGSK-F and reverse oligo ecGSK-R were used for the amplification of full-length ecGSK; oligo ecGSK-Δ30N-F and oligo ecGSK-R for amplifying ecGSK-Δ30N; oligo ecGSK-F and oligo ecGSK-Δ21C-R for amplifying ecGSK-Δ21C; oligo ecGSK-Δ30N-F and oligo ecGSK-Δ21C-R for amplifying the double truncated construct, ecGSK-Δ30N/Δ21C. In all cases, NdeI and BamHI sites were incorporated into the forward and reverse primers, respectively. The four amplicons were treated with NdeI and BamHI-HF restriction enzymes to create sticky ends and were ligated into pET-14bSUMOΔThr having an N-terminal His$_6$-SUMO tag. The sequence-verified (Seqlab, Goettingen) recombinant plasmids were used to transform BL21(DE3)pLysS cells for expression. All four constructs were

expressed and purified by affinity chromatography using Protino Ni-IDA resin (see section 2.2.5.1), anion exchange chromatography using DEAE Sepharose FF column, and finally by gel filtration chromatography using Superdex 200 column (GE Healthcare). Purity of the purified ecGSK was tested by 12% SDS-PAGE and the protein concentration was determined by the Bradford assay. The enzymatic activity was determined by the NADH-dependent spectroscopic assay (see section 2.2.6). The highly pure and homogeneous ecGSK protein was used for crystallization experiments and for limited proteolysis. Actually, the truncated forms generated for ecGSK as described above were based on the results of limited proteolysis, IUPred prediction and secondary structure prediction analysis. Limited proteolysis was used to identify and remove the highly flexible (e.g. loops) or unfolded regions from ecGSK to optimize the sample for crystallization [313]. In this method, ecGSK (~0.7 mg/ml) in separate microtubes was treated with different proteases such as chymotrypsin, Gluc-C, subtilisin, and thermolysin added in 1/10, 1/100, and 1/1000 dilutions. All samples were incubated at 20 $^{\circ}$C for 30 min. The reactions were stopped by adding Laemmli sample buffer and heated at 95 $^{\circ}$C for 5 min. SDS-PAGE was performed to identify the cleavage products.

3. Results

3.1 Biochemical characterization of human guanylate kinase

3.1.1 Expression, purification, and kinetic characterization of wild-type hGMPK and site-specific mutants

To investigate the structural and functional properties of human guanylate kinase (hGMPK), the ORF (open reading frame) of full-length wild-type hGMPK (591 bp) was PCR-amplified from the template DNA gifted by Oliver Spangenberg [81]. It was previously reported that hGMPK, when purified as a recombinant protein produced in *E.coli* cells, was found to be catalytically inactive [16, 29, 34]. Therefore, our primary goal was to obtain hGMPK in catalytically active and highly pure form. For that purpose, it was cloned into the [pET-14bSUMOΔThr] expression vector. In order to explain the critical role of certain residues in catalysis and domain movements, eight mutants were generated by QuikChange site-directed mutagenesis using the hGMPK-containing pET-14bSUMOΔThr plasmid as a template. These mutants were S37A, S37C, S37Y, S37P, N42P, Y81F, T83A, and T83S. The wild-type hGMPK and all mutant constructs were sequence-verified, and were transformed into the *E.coli* expression strain BL21(DE3)pLysS. The overexpression was induced in 1 liter lactose-containing auto-inducing media incubated overnight at 21 ^{o}C. Cells were harvested by high-speed centrifugation at 4 ^{o}C, and the pellets were lysed by sonication. Lysates were cleared by centrifugation and subjected to purification by affinity chromatography using Protino Ni-IDA resin. All purification steps were performed at 4 ^{o}C in the cold room. The purified fusion proteins carried N-terminal His$_6$-SUMO tags which were cleaved by SUMO-protease added in 1:100 molar ratio of protease:protein and incubated for 30 min at room temperature. After cleavage, the His$_6$-SUMO tag was removed from the respective hGMPK samples by adding Protino Ni-IDA resin to selectively bind the tag. Purity of the purified wild-type and mutant hGMPKs was tested by 12% SDS-PAGE (Fig. 3.1). To get the hGMPK in homogeneous form, it was further passed through a gel filtration column (Superdex 75 10/300 GL, GE Healthcare). It eluted as a monomer when compared to the chromatogram of Bio-Rad gel filtration standard (Fig. 3.2). Protein yield from 1 liter-culture was in the range of 10-15 mg.

The wild-type hGMPK and all eight mutants were characterized by steady-state kinetics using the NADH-dependent enzyme coupled assay [84, 99, 100]. The GMP substrate concentration

was used in the range of 10-2500 μM keeping the hGMPK concentration constant at 18 nM. Turnover rates (k_{obs}) were calculated, and the values were fit to the Michaelis-Menten equation (equation 1, section 2.2.6) using the command-driven interactive function plotting program Gnuplot 5.0. The kinetic parameters, K_m and k_{cat}, were calculated for all constructs from the plots (Fig. 3.3a-h); the data are summarized in Table 3.1. Seven mutants out of eight had decreased catalytic efficiency as compared to the wild-type hGMPK. Only one mutant, S37Y, was found to be catalytically inactive. Nevertheless, the recombinantly produced S37Y was a soluble and stable protein with no aggregation observed during its expression and purification. The loss of activity could be due to the disturbed microenvironment at the active site for binding the GMP substrate when serine was substituted with the aromatic tyrosine residue.
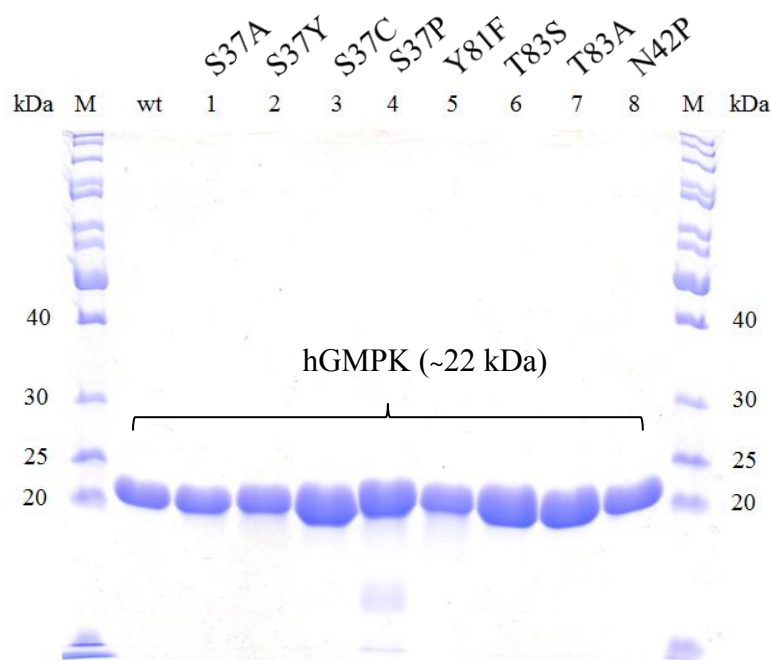
Fig. 3.1. SDS-PAGE of wild-type hGMPK and site-specific mutants. The wild-type hGMPK and its eight mutants were purified by affinity chromatography using Protino Ni-IDA resin (Macherey-Nagel). The purity was tested by 12 % SDS-PAGE. Lanes: M, marker proteins (PageRuler Unstained Protein Ladder from Thermo Scientific); wt, wild-type hGMPK (~22 kDa); 1-8, S37A, S37Y, S37C, S37P, Y81F, T83S, T83A, and N42P mutants, respectively.
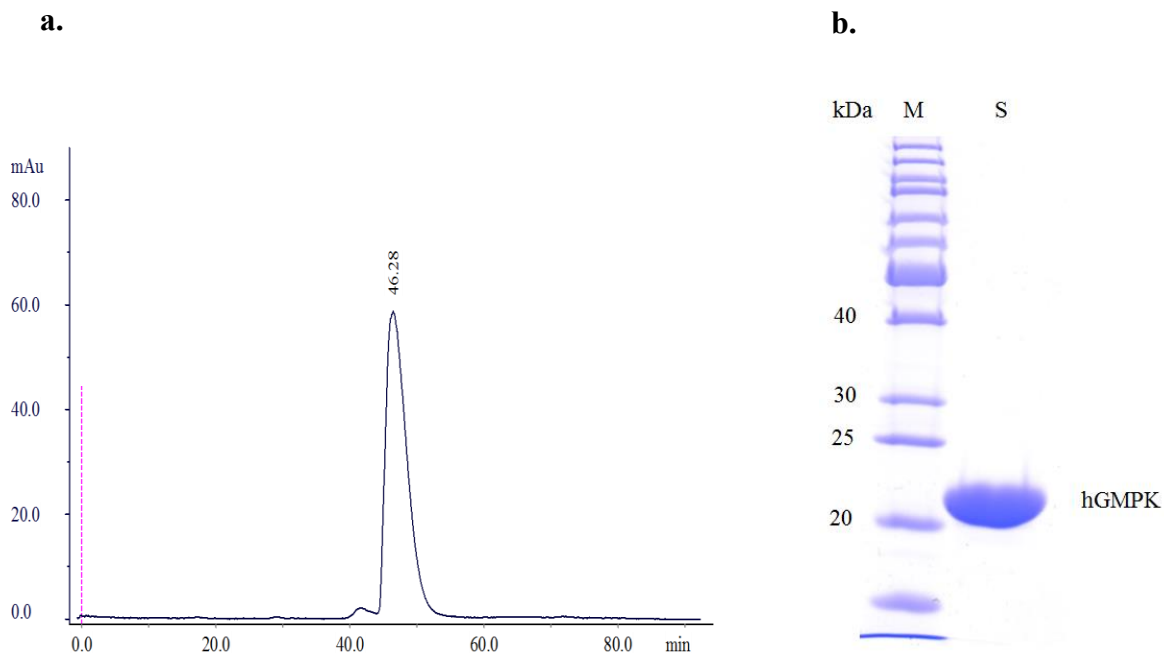
Fig. 3.2. Monomeric form of hGMPK. (a) The protein eluted at a retention time (t_R) of 46.28 min when loaded on a Superdex 75 column (10/300 GL, 24 ml, GE Healthcare) using the buffer 50 mM HEPES pH 7.4, 300 mM NaCl, 5 % glycerol, and 1 mM DTT. The t_R of 46.28 min corresponds to the monomeric form (~22 kDa) of hGMPK when compared to the chromatogram of Bio-Rad's gel filtration standard (data not shown). (b) 12 % SDS-PAGE of hGMPK after gel filtration chromatography. M, marker proteins (PageRuler Unstained Protein Ladder from Thermo Scientific) and S, hGMPK.
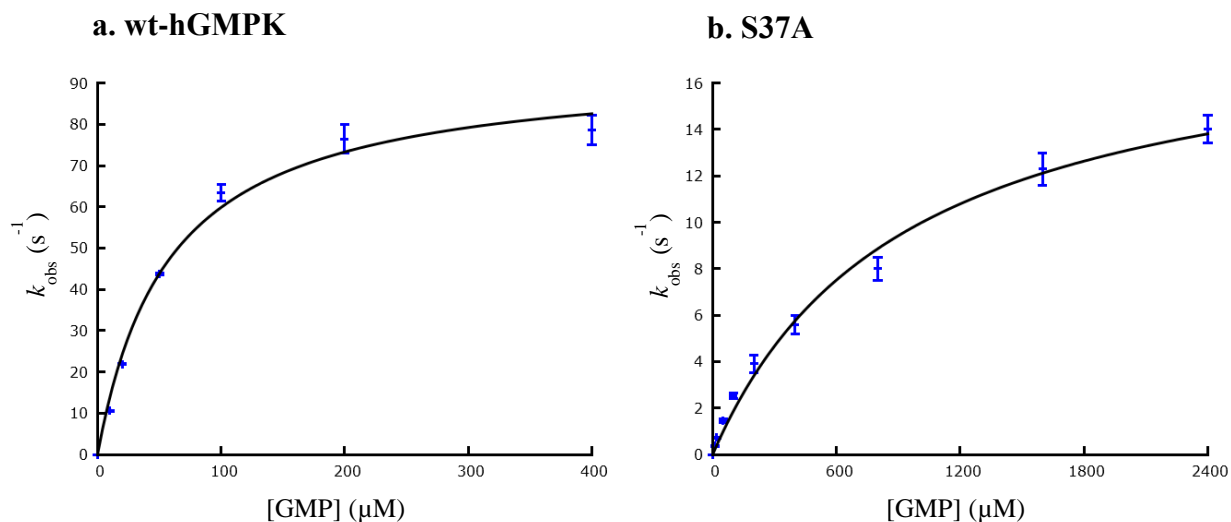
c. S37C
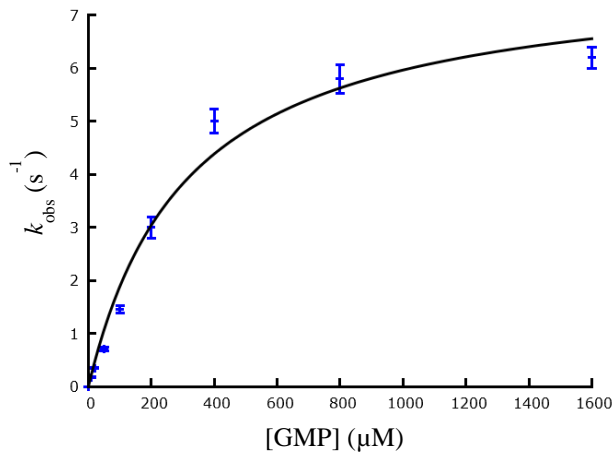
d. S37P
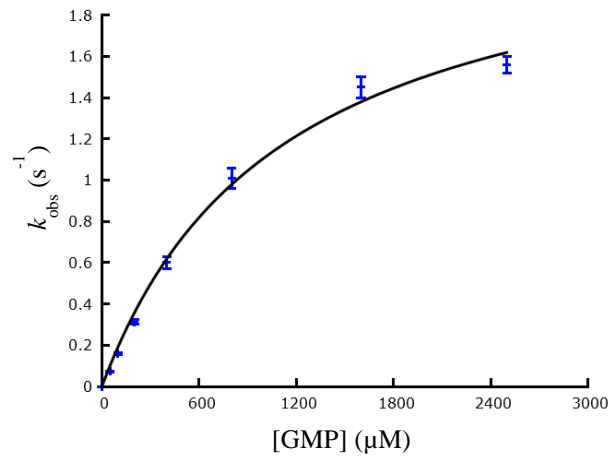
e. N42P
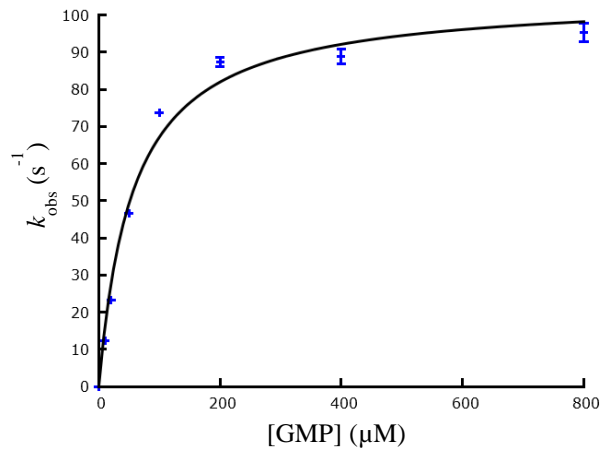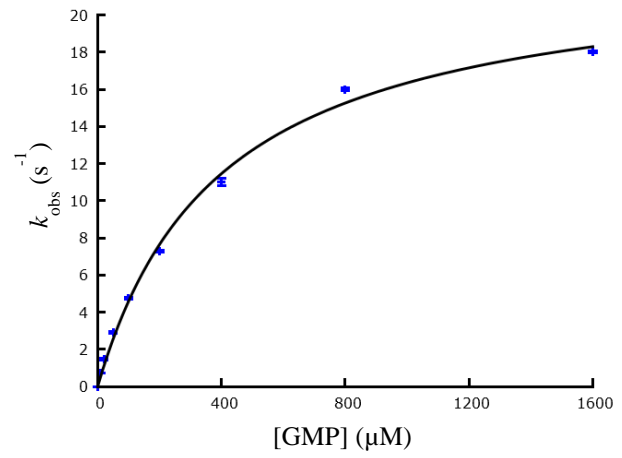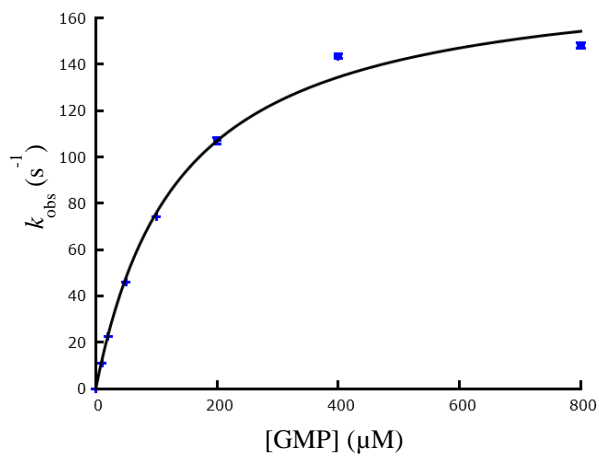
f. Y81F

g. T83S

h. T83A
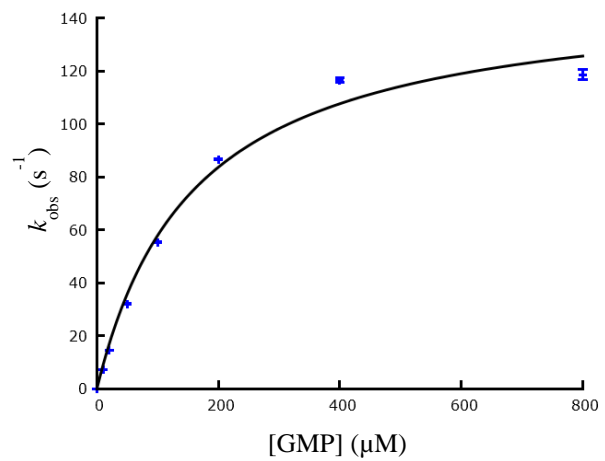
Fig. 3.3. Steady-state kinetic plots for wild-type hGMPK and site-specific mutants (a-h). V/E versus [GMP] plot for (a) wild-type hGMPK and (b-h) mutants: b, S37A; c, S37C; d, S37P; e, N42P; f, Y81F; g, T83S and h, T83A. All measurements were performed at 25 °C in 1 ml of 100 mM Tris-HCl buffer pH 7.5, containing 100 mM KCl and 10 mM MgCl$_2$. The enzyme concentration used was 18.4 nM in each case (0.4 µg of hGMPK with a molecular weight of ~22 kDa). Turnover rates (k_{obs}) are expressed as a function of the GMP concentration. Error bars indicate standard deviation of duplicate measurements. The hyperbolic plots were prepared by the Gnuplot 5.0 software by non-linear regression using the Michaelis-Menten equation (equation 1, section 2.2.6).

Table 3.1 Steady-state kinetic parameters for wild-type hGMPK and site-specific mutants. The kinetic parameters were calculated from the Michaelis-Menten plots given above. Values are represented as means ± standard deviation of duplicate measurements. nd, activity was not detected.

Enzyme	K_m (µM) for GMP	k_{cat} (s^{-1})	k_{cat}/K_m (M^{-1} s^{-1})
wt	25 ± 0.31	79 ± 0.02	316 x 10^4 ± 0.06
S37A	580 ± 0.51	14 ± 0.001	2.4 x 10^4 ± 0.002
S37C	260 ± 0.71	7 ± 0.002	3.0 x 10^4 ± 0.003
S37P	505 ± 0.85	2 ± 0.0001	0.3 x 10^4 ± 0.0001
S37Y	nd	nd	nd
N42P	50 ± 0.67	101 ± 0.003	202 x 10^4 ± 0.004
Y81F	290 ± 1.5	18 ± 0.005	6.0 x 10^4 ± 0.002
T83S	130 ± 0.78	171 ± 0.01	132 x 10^4 ± 0.01
T83A	110 ± 0.93	119 ± 0.006	108 x 10^4 ± 0.006

3.1.2 Substrate-induced conformational changes in hGMPK studied by small angle X-ray scattering

To determine the four conformational states designated as open form (hGMPK$_{apo}$), two partially closed forms (hGMPK$_{PC}$), and completely closed forms (hGMPK$_{closed}$), we analyzed the low

resolution structures of hGMPK by small angle X-ray scattering (SAXS) in the presence of different ligands. The SAXS measurements were done in collaboration with Prof. Simone Techert and Rohit Jain from the Structural Dynamics of (Bio)chemical Systems group at MPI-bpc, Goettingen.

For structural analysis of different conformations, SAXS measurements were performed at the third-generation cSAXS beamline (Paul Scherrer Institute, Swiss Light Source, Switzerland) in the presence of enzyme substrates and substrate analogs; GMP, AMP-PNP (non-hydrolyzable ATP-analog), GMP/AMP-PNP, GMP/ATP, and Ap5G (bi-substrate analog). The measured data was analyzed for conformational changes. There was no significant change in the activity of the enzyme after SAXS measurements (Table 3.2). Guinier analysis and the corresponding R_g (radius of gyration) values are shown in Table 3.3. As expected, the $hGMPK_{apo}$ molecule was bigger (~21 Å) in size than $hGMPK_{closed}$ (~19 Å) and the two partially closed forms (19.7 Å & 20.2 Å). It is also obvious that the product formation upon addition of ATP and GMP caused a small increase (~0.7 Å) in the $hGMPK_{RO}$ (reopened form) size as compared to that of closed forms. In agreement with the Guinier analysis, pair distance distribution function $P(r)$ for different conditions indicated that the D_{max} (maximum diameter from the pair-distance distribution function $P(r)$) of hGMPK decreased from 68 Å in the open form to ~55 Å in the closed form, and increased from ~55 Å to 63 Å in the reopened form after the enzymatic reaction (Table 3.3, and Appendix Fig. 1 & 2). Based on these observations, we conclude that the closed forms of hGMPK have more compact conformations than all other forms in solution:

$$hGMPK_{closed} > hGMPK_{PC} \geq hGMPK_{RO} > hGMPK_{apo}$$

The scattering shapes of hGMPK in different conformations generated by GASBOR (program for *ab initio* reconstruction of protein structure) are shown in Fig. 3.4, and Appendix Fig. 2. The three-dimensional surface reconstruction of $hGMPK_{apo}$ in the absence of ligands is a globular domain with two projections originating from opposite sides of the molecule, large P1 and small P2. These two projections move with respect to the center of the enzyme molecule and become less conspicuous in the closed and reopened forms (Fig. 3.4, and Appendix Fig. 2). The superposition of closed and reopened forms of hGMPK onto its open form confirms the substrate-induced conformational changes due to domain movements (Fig. 3.4). Major visible changes in the closed and reopened forms occur in the P1 and P2 regions (Fig. 3.4).

Table 3.2 Kinetic parameters of hGMPK at 25 °C

Substrate	K_m (µM)	k_{cat} (s^{-1})[a]	k_{cat}/K_m (M^{-1} s^{-1})
GMP	25	79	316 x 10^4
ATP	95	79	83 x 10^4

[a] The k_{cat} was calculated using the equation $V_{max} = k_{cat}/[E]$ where [E] is total enzyme concentration and is based on one active site per monomer. Assay conditions are described in Materials and Methods, section 2.2.6.

Table 3.3 Structural parameters for hGMPK in unliganded and nucleotide-bound forms

hGMPK	Conformation	Guinier plot						Indirect Fourier Transformation		
		Manual		PRIMUS				P(r) function		
		R_g (Å)	I_0	R_g (Å)	I_0	R_g (Å)	L (Å)	D_{max} (Å)	R_g (Å)	I_0
Apo-form	Open	21 ± 2.5	4.3	21 ± 2.5	0.14	10.1 ± 0.14	63.8	68	21 ± 0.06	1.4
GMP + AMP-PNP	Closed	18.8 ± 2.4	4.3	18.8 ± 2.4	0.15	8.74 ± 0.18	59.2	55	18.76 ± 0.05	1.4
AP5G	Closed	19.0 ± 2.6	4.2	19.0 ± 2.6	0.14	7.93 ± 0.21	60.2	57	18.45 ± 0.06	1.4
GMP + ATP	Partially reopened	19.7 ± 2.7	4.3	19.7 ± 2.7	0.14	7.7 ± 0.19	62.8	63	19.2 ± 0.06	1.4
AMP-PNP	Partially closed	20.2		20.8		9.8	63.5	NA		
GMP	Partially closed	19.7		20		8.8	62.2	NA		

a.

P1

P2

hGMPK$_{apo}$

b.

hGMPK$_{apo}$ *vs.* hGMPK$_{closed1}$

c.

hGMPK$_{apo}$ *vs.* hGMPK$_{closed2}$

d.

hGMPK$_{apo}$ *vs.* hGMPK$_{RO}$
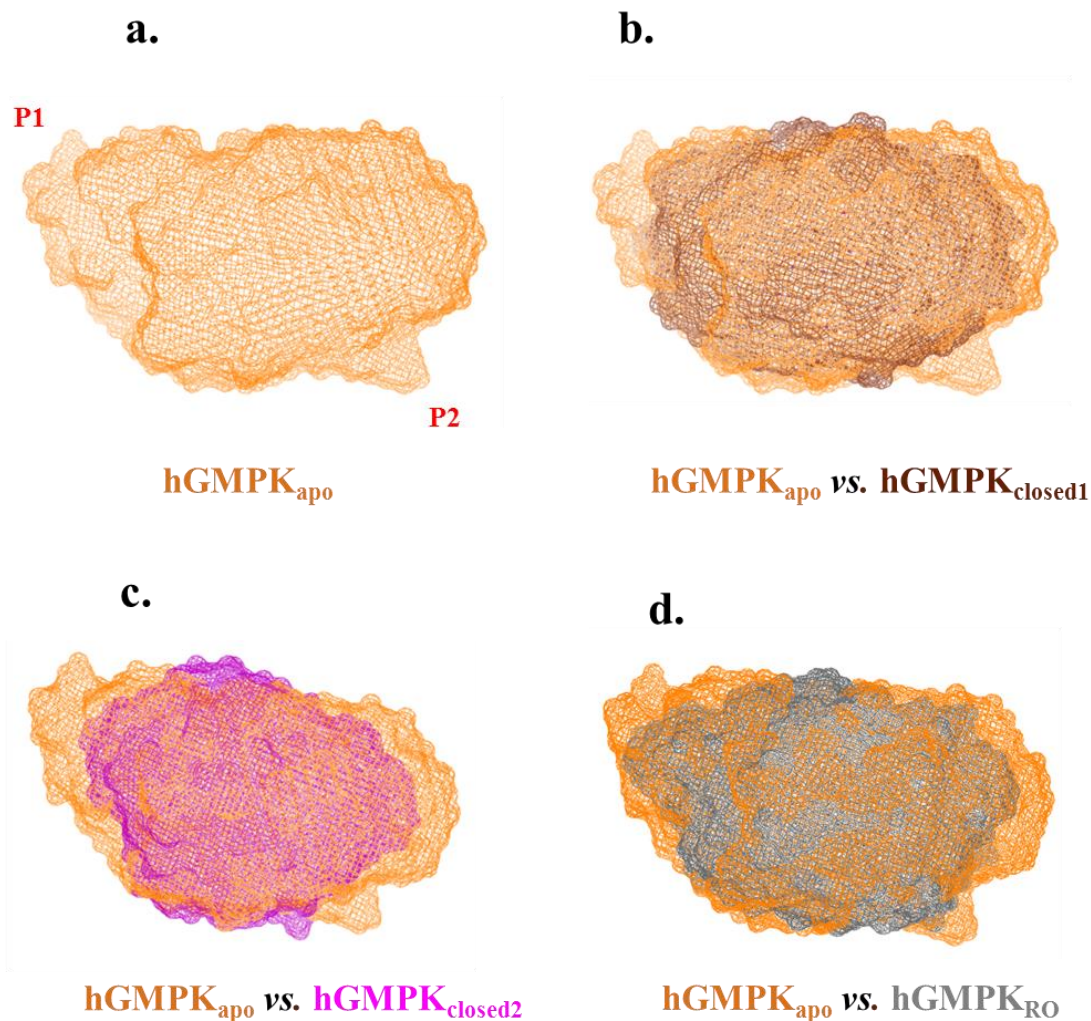
Fig. 3.4. Superposition of the three-dimensional surface reconstruction of hGMPK$_{apo}$ on three other conformational forms of hGMPK. The open form (hGMPK$_{apo}$) of hGMPK was overlaid on the two closed forms (hGMPK$_{closed1}$ and hGMPK$_{closed2}$) and one reopened form (hGMPK$_{RO}$) using PyMOL to identify the structural regions that undergo main conformational changes upon binding of ligands. hGMPK$_{apo}$ is the unliganded form, hGMPK$_{closed1}$ is with bound GMP and AMP-PNP (non-hydrolyzable ATP analog), hGMPK$_{closed2}$ is with bound Ap5G (bi-substrate analog), and hGMPK$_{RO}$ is the form in the presence of GMP and ATP. P1 and P2 are the two projections (structural regions) that move upon binding of nucleotides.

The overall fold of hGMPK is very similar to other members of the NMP-kinases family, in particular to mouse and yeast GMPKs [10, 11, 117]. As the two mammalian enzymes share high amino acid sequence identity (88%), an *in silico* homology model was constructed for hGMPK

based on mGMPK as a template to highlight its structural features (Fig. 3.5). It has three dynamic structural regions called NMP-binding region (NMP-BR), CORE and LID regions interconnected by four hinges.
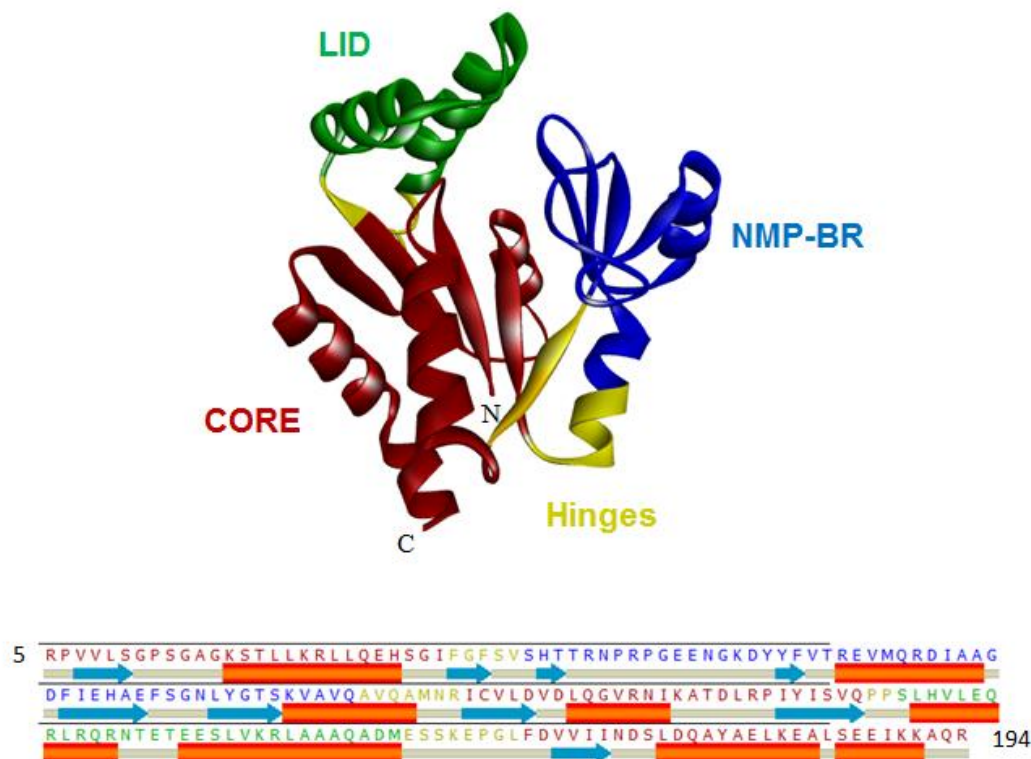
Fig. 3.5. Ribbon diagram of hGMPK. The homology model for hGMPK was constructed using SWISS-MODEL (ExPASy server) based on the closely related (88% identical) mGMPK crystal structure (PDB 1LVG). The important structural regions designated as NMP-binding region (NMP-BR, *blue*), CORE (*red*) and LID (*green*), and the corresponding amino acid sequences are color coded. These three structural regions are interconnected by four dynamic hinges which are *yellow* color coded. The predicted secondary structures are 8 α-helices (*orange* bars) and 9 β-strands (*blue* arrows). The amino and carboxy termini are denoted by N and C, respectively.

To identify these three structural regions in the low resolution SAXS models of hGMPK, the SAXS structures of hGMPK were overlaid on the crystal structure of mGMPK$_{closed}$ (PDB 1LVG), and regions were identified after manually aligning their surface topology (Fig. 3.6).

a.

P1

HINGES

NMP-BR

CORE

LID

P2

hGMPK$_{apo}$ (SAXS)
mGMPK$_{ADP+GMP}$ (PDB-1LVG)

b.

HINGES

NMP-BR

CORE

LID

hGMPK$_{AMP-PNP+GMP}$ (SAXS)
mGMPK$_{ADP+GMP}$ (PDB-1LVG)

c.

HINGES

NMP-BR

CORE

LID

hGMPK$_{Ap5G}$ (SAXS)
mGMPK$_{ADP+GMP}$ (PDB-1LVG)

d.

HINGES

P1

NMP-BR

CORE

LID

hGMPK$_{GMP+ATP}$ (SAXS)
mGMPK$_{ADP+GMP}$ (PDB-1LVG)
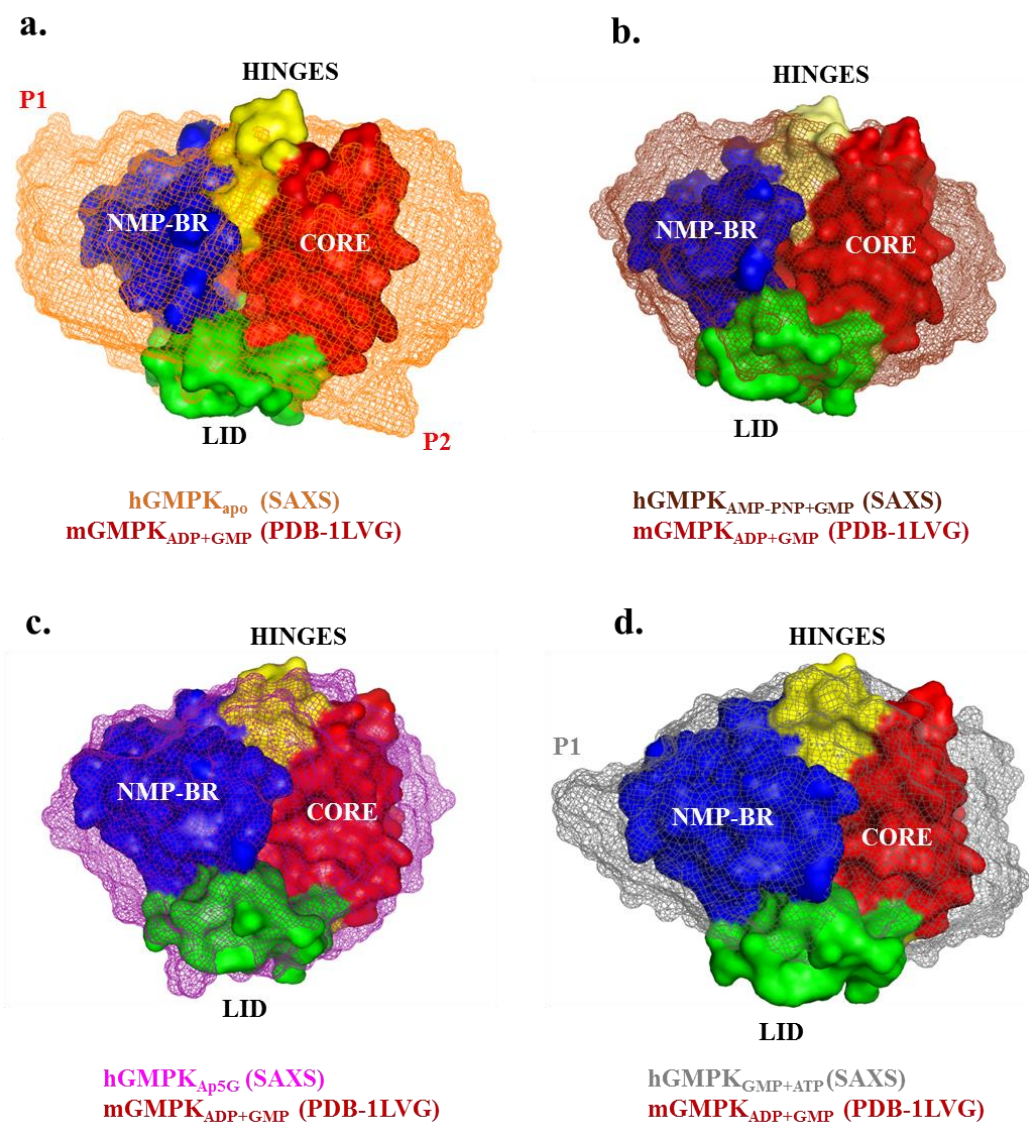
Fig. 3.6. Comparison of the hGMPK SAXS structures with mGMPK crystal structure. Different SAXS models of human guanylate kinase (hGMPK, mesh representation) were overlaid on the crystal structure (PDB 1LVG) of mouse guanylate kinase (mGMPK, surface representation) using PyMOL (a-d). mGMPK$_{ADP+GMP}$ is a closed form with bound ADP and GMP, hGMPK$_{apo}$ is unliganded form (open form), hGMPK$_{AMP-PNP+GMP}$ is a closed form with bound AMP-PNP (non-hydrolyzable ATP analog) and GMP, hGMPK$_{Ap5G}$ is a closed form with bound Ap5G (bi-substrate analog), and hGMPK$_{GMP+ATP}$ is the reopened form attained after adding GMP and ATP substrates to hGMPK. The three distinctive structural regions in the mGMPK are color coded, and the two structural regions which undergo major conformational changes in hGMPK are labeled as P1 and P2 (a-d).

110

The P1 region in hGMPK aligns with the NMB-BR of mGMPK, whereas P2 aligns with the LID region. These two regions are highly prominent in the $hGMPK_{apo}$ (Fig. 3.6a) and become least conspicuous in the closed forms of hGMPK (Fig. 3.6b & c) due to their movements towards the center of the molecule for binding the substrates. Providing ATP and GMP to hGMPK, the reaction proceeds to produce GDP and ADP reversibly till equilibrium is established. When this catalytic reaction was allowed to occur, only P1 reappeared in the $hGMPK_{RO}$ (Fig. 3.6d), suggesting that the major movement is brought about by P1 (Fig. 3.6a-d).

3.1.3 Structural characterization of hGMPK by NMR

NMR spectroscopy is a powerful tool for determining the three-dimensional structure of proteins. Proteins are much closer to their natural state in NMR studies, which analyze proteins in solution, than they are in the crystalline form used for X-ray crystallography. Proteins which fail to crystallize like hGMPK as reported by Sekulic et al [11] are ideally studied by NMR for their structure-function characterization. The hGMPK which is a monomeric protein has the molecular mass of about 22 kDa (197 amino acids) and is within the size range (5-25 kDa) of proteins amenable to NMR studies. Nevertheless, structure elucidation by NMR requires a large quantity (0.5-1.5 mM in 260-550 µl volume) and high quality (>95% purity) of isotopically (^{15}N, ^{13}C) labeled recombinant protein, the production of which is a costly and time-consuming aspect of NMR studies. Therefore, I optimized a protocol for improved production of labeled hGMPK for its NMR structure determination. The work on NMR structure characterization of hGMPK has been done in collaboration with Dr. Donghan Lee and Dr. Thomas Michael Sabo from the Department of NMR-based Structural Biology, MPI-bpc Goettingen.

3.1.3.1 Optimizing the yield of isotope-labeled (^{15}N, ^{15}N/^{13}C) hGMPK

A low yield of hGMPK (~1 mg/l culture) was observed when expressed in M9 minimal growth medium (see section 2.1.9.1) using a selected colony from its glycerol stock. One of the solutions to the problem was to use freshly transformed colonies for overexpression. However, this did not guarantee the reproducibility of higher yield. Sometimes it resulted in higher expression, other time very low expression. Such a situation becomes critical when dealing with the overexpression of isotopically labeled proteins. To solve this problem, I combined the protocols mentioned by Cai et al [105], Marley et al [103], and Sivashanmugam et al [102] with several

modifications to achieve higher yield of single- (^{15}N) and double-labeled (^{15}N and ^{13}C) hGMPK. Our optimized protocol was based on the use of both the double-colony selection and the high cell-density method. The double-colony selection procedure was used to optimize the higher yield which is one of the most important factors for high-level protein production using high-density bacterial expression methods [102, 103].

3.1.3.1.1 Double-colony selection

I combined double-colony selection with the high-density cell expression for the production of ^{15}N- and ^{13}C-labeled hGMPK. To avoid the loss of costly ^{13}C-glucose and ^{15}N NH$_4$Cl, the method was first optimized using unlabeled M9 minimal growth medium (see section 2.1.9.1) and then applied to the recombinant production of single- and double-labeled hGMPK. I simplified the M9 minimal medium composition by omitting the commercially available "Basal Vitamins Eagle Media" and replacing it with only thiamine (1 µg/ml) ingredient [103]. No adverse effect was observed on the final yield of hGMPK. The hGMPK[pET-14bSUMOΔThr] construct was transformed into BL21(DE3)pLysS cells and streaked on an agar plate. One of the colonies from the agar plate was used to inoculate LB medium for growing up to OD$_{600}$ ~0.7, and was then spread on an agar plate for overnight incubation at 37 ^{o}C. For 1st colony selection, four colonies were randomly selected and grown in enriched LB media up to the OD$_{600}$ of ~0.7, and then the cultures were switched to fourfold smaller volume of M9 minimal media for induction with 1 mM IPTG for 4 h at 37 ^{o}C. Cells were harvested by centrifugation, and their level of hGMPK expression was tested by 12% SDS-PAGE (Fig. 3.7).

One of the colonies (Lane-1, Fig. 3.7) displayed a comparatively higher level of hGMPK expression and was therefore used for the 2nd round of selection. An aliquot of colony-1 was streaked on an agar plate for overnight incubation at 37 ^{o}C. Five colonies were chosen from the agar plate for hGMPK expression using M9 minimal media. The procedure followed for the 2nd selection was the same as for the 1st selection. After expression in M9 minimal medium, cell pellets of the five selected colonies were subjected to SDS-PAGE (Fig. 3.8).
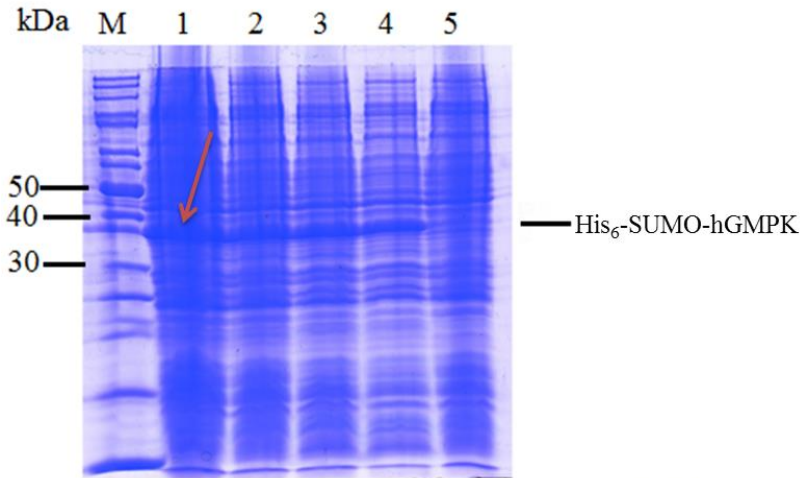
Fig. 3.7. 12% SDS-PAGE of hGMPK for the 1ˢᵗ colony selection. The expression was induced in M9 minimal medium (0.4% glucose and 0.1% NH_4Cl) using 1 mM IPTG incubating at 37 °C for 4 h. Lanes: M, marker proteins (PageRuler Unstained Protein Ladder, Thermo Scientific); 1-4, four different colonies picked up from the 1ˢᵗ colony selection plate for testing their level of hGMPK expression as a fusion with N-terminal His_6-SUMO tag (~37 kDa); 5, control sample without IPTG-induction. Colony-1(Lane-1), due to its high level of hGMPK expression, was selected for the next step of the 2ⁿᵈ colony selection.
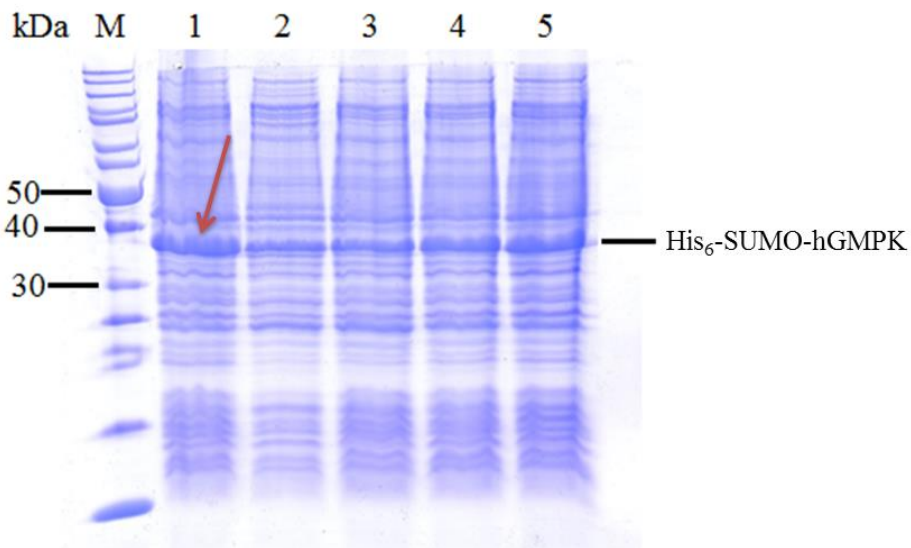
Fig. 3.8. 12% SDS-PAGE of hGMPK-expressing clones for the 2ⁿᵈ colony selection. The expression was induced in M9 minimal medium (0.4% glucose and 0.1% NH_4Cl) using 1 mM IPTG at 37 °C for 4 h. Lanes: M, marker proteins; 1-5, five different colonies from the 2ⁿᵈ colony selection plate to check their levels of hGMPK expression as a fusion with N-terminal His_6-SUMO tag (~37 kDa). Colony-1(Lane-1) was selected for the next step of the 3ʳᵈ colony selection.

If we compare the expression levels of the five colonies from Fig. 3.8, it is obvious that the colony corresponding to lane-1 has higher expression level for hGMPK; thus it was streaked on an agar plate for the 3rd selection step. Six colonies were picked up from the 2nd selection plate. The expression procedure followed was the same as for the 1st selection process. After M9 minimal expression for 4 h at 37 ^{o}C in the presence of 1 mM IPTG, the colonies were tested for their level of hGMPK expression (Fig. 3.9).
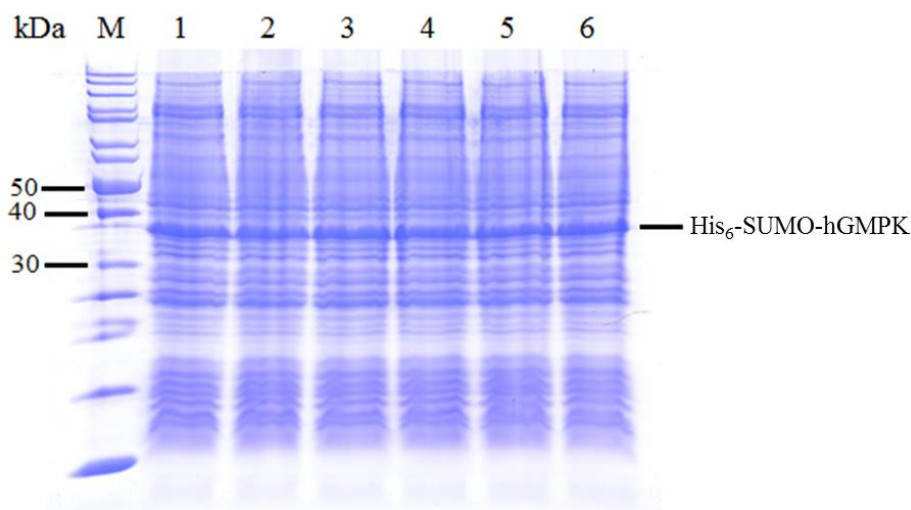
Fig. 3.9. 12% SDS-PAGE of hGMPK for 3rd colony selection. The expression was induced in M9 minimal medium (0.4% glucose and 0.1% NH$_4$Cl) using 1 mM IPTG at 37 ^{o}C for 4 h. Lanes: M, marker proteins; 1-6, six different colonies from the 3rd selection plate. All colonies had almost identical levels of hGMPK after expression under similar conditions.

The SDS-PAGE profile of all six colonies was almost identical which means that no further improvement could be achieved after the 2nd selection process. Therefore, 30% glycerol stock of the selected colony was prepared in LB and was stored at -80 ^{o}C for future use. In the next step, the high-expression colony was tested for the production of single (^{15}N) and double-labeled (^{15}N and ^{13}C) hGMPK production using the high cell-density procedure given below.

3.1.3.1.2 High cell-density method

Protein yield is proportional to the cell-density of bacterial expression [102]. To attain high cell-density, a rich medium like LB was used for the initial growth of *E.coli* cells to reach the mid-logarithmic phase (OD_{600} ~0.7) as shown schematically (Fig. 3.10).

Starting culture
(4 liters LB)

$37\,^{\circ}C$

OD_{600} ~0.7

Switch medium
(1 liter M9 minimal medium
^{13}C-glucose and ^{15}N NH_4Cl)

$37\,^{\circ}C$, 1 h

IPTG induction (1 mM)
OD_{600} ~2.4

$37\,^{\circ}C$, 4 h

Harvest cells
OD_{600} ~4.2

Protein purification
(Affinity chromatography using Protino Ni-IDA resin and
gel filtration chromatography using Superdex 75 column)
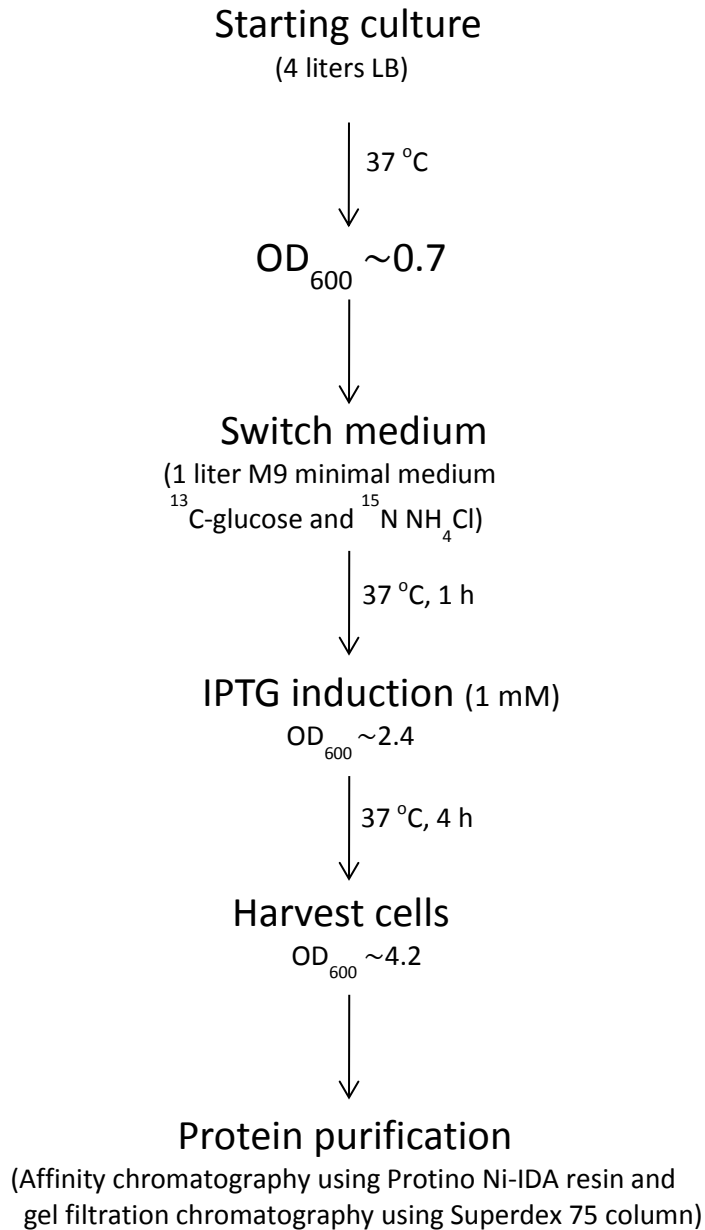
Fig. 3.10. Schematic representation of the high cell-density method

At this point, the cultures were centrifuged, and the cells were resuspended in a fourfold smaller volume of M9 minimal medium as compared to the starting LB medium thereby attaining approximately four times higher OD_{600} ~2.4 [103]. After switching to M9 minimal growth medium, the bacterial cells were incubated at 37 ^{o}C for 1.0–1.5 h before IPTG-induction. This incubation period was essential for two reasons: First, to allow cells to clear unlabeled metabolites while preserving the labeled ingredients for protein synthesis after IPTG-induction; second, to let the cells adjust to the minimal medium conditions because the bacterial growth during the exchange period may slow down. After adding 1 mM IPTG, cells were incubated at 37 ^{o}C for 4 h induction with rapid shaking at 250 rpm. The final OD_{600} obtained was ~4.2 yielding an appreciable amount of hGMPK (~8 mg/l culture). Additionally, different experimental parameters were tested (data not shown) such as various incubation temperatures (21 ^{o}C, 25 ^{o}C, 30 ^{o}C, 37 ^{o}C), IPTG induction time (3 h, 4 h, 5 h, 6 h, overnight), IPTG concentrations (0.25 mM, 0.5 mM, and 1 mM). With all these modifications, I successfully optimized the protocol for high production of labeled hGMPK.

Using the optimized protocol, ~8 mg of single/double-labeled (^{13}C and ^{15}N) hGMPK was reproducibly obtained from one liter of M9 minimal medium. This amount is 8 times higher than the traditional unmodified protocols.

3.1.3.2 The ^{15}N-^{1}H HSQC spectra of ^{15}N-labeled hGMPK$_{apo}$ (open form) and ^{15}N-labeled hGMPK$_{GMP}$ (partially closed form)

The 1 mM ^{15}N-labeled hGMPK (~22 mg/ml) was titrated against increasing concentrations of the GMP substrate ranging from 3 mM to 18.6 mM. A significant change in the positions of residues was observed clearly indicating that the hGMPK molecule changes its conformation from a dilated open form (or apo-form) to a more compact partially-closed form (Fig. 3.11a). The magenta color in the Fig. 3.11a represents the apo-form of hGMPK whereas the cyan color represents the partially closed-form of the enzyme.

3.1.3.3 The 3D HNCA and 3D ^{15}N-resolved NOESY spectra of ^{13}C/^{15}N-labeled hGMPK

About 177 residues out of total 197 amino acids of hGMPK have been identified from the 3D NMR spectra analysis. Of these, five C-terminal amino acids are shown in (Fig. 3.11b). Work is in progress for solving the three-dimensional structure of hGMPK.

a.

b.
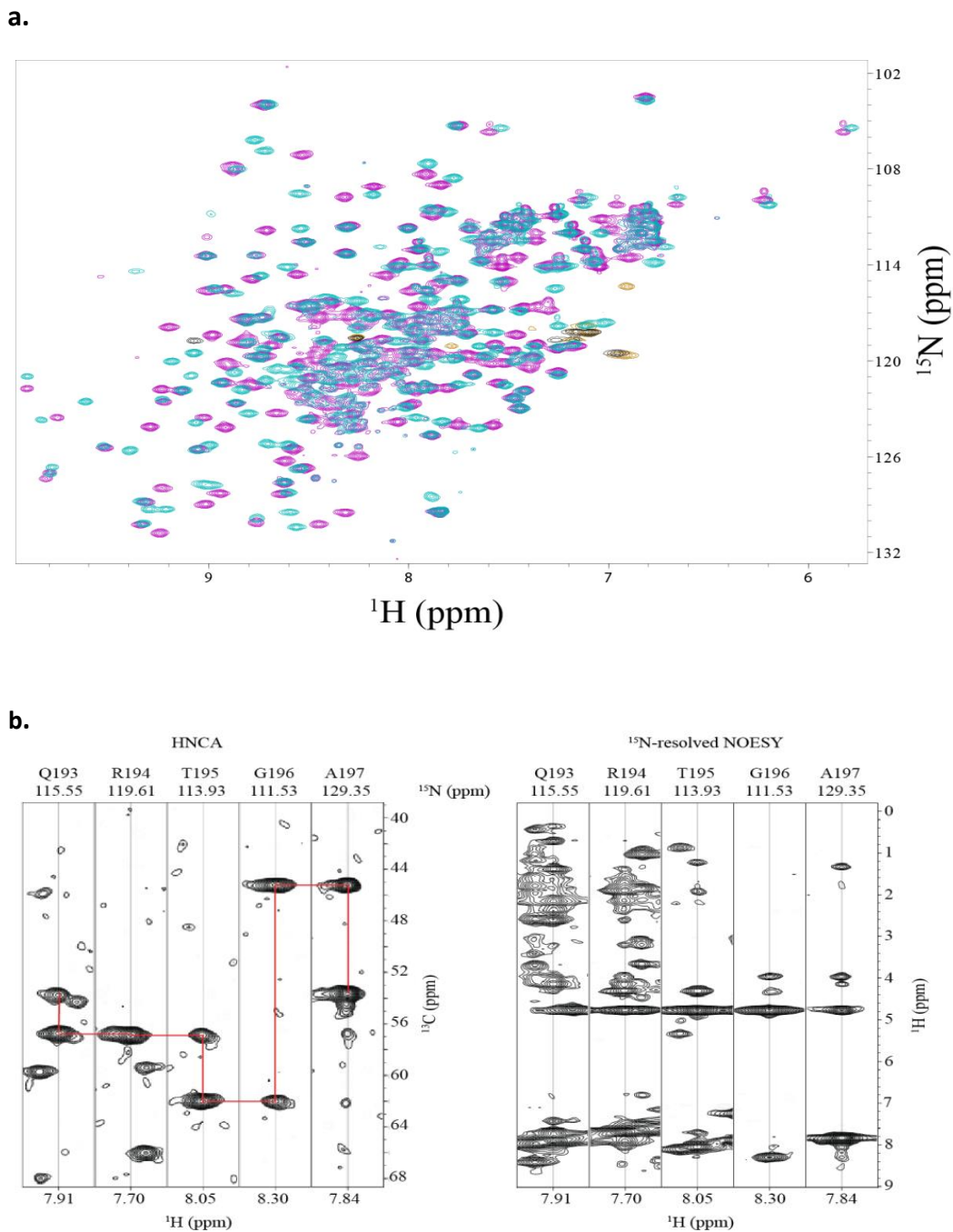
Fig. 3.11. **^{15}N-^{1}H HSQC , 3D HNCA and 3D ^{15}N-resolved NOESY spectra of hGMPK.** (a) ^{15}N-^{1}H HSQC spectra of 1 mM ^{15}N-labeled hGMPK$_{apo}$ (*magenta*) and 0.6 mM ^{15}N-labeled hGMPK and 18.6 mM GMP (*cyan*) recorded at a proton frequency of 900 MHz and a temperature of 25 ^{o}C. Buffer conditions were 25 mM phosphate, pH 7.2, 10% D$_2$O, 100 mM NaCl, and 1 mM TCEP. (b) Strips from a 3D HNCA and 3D ^{15}N-resolved NOESY spectra of 1 mM ^{13}C/^{15}N-labeled hGMPK recorded at a proton frequency of 800 MHz and a temperature of 25 ^{o}C. Buffer conditions were 25 mM phosphate, pH 7.2, 10% D$_2$O 100 mM NaCl, and 1 mM TCEP. The spectra were processed with NMRPipe software [314].

3.1.4 Enhanced cytotoxicity of the antileukemic drug 6-thioguanine by expressing hGMPK in HEK293 cells

6-Thioguanine is an FDA-approved antileukemic drug [107]. It is a prodrug that is activated intracellularly via the purine salvage pathway including the phosphorylation step catalyzed by nucleoside monophosphate kinase to produce 6-thioGDP from 6-thioGMP [18, 108]. After further phosphorylation to 6-thiodGTP, it is incorporated into DNA by DNA polymerases and exerts its cytotoxic effects in several ways, most importantly by stimulating the DNA mismatch repair system, producing DNA-damaging reactive oxygen species, and by causing mitochondrial dysfunction [17, 19, 109, 110]. Previously, it was reported that a conditional guanylate kinase deficient *E.coli* strain (TS202A(DE3)) when genetically complemented with wild-type mouse GMPK showed sensitivity to 6-TG. However, certain mutations in the GMPK caused resistance to the drug [16, 108]. This indicated the role of GMPK in the metabolic activation of 6-TG when tested in *E.coli*.

We decided to use HEK293 cells as a model system as previously described for ganciclovir-induced cell killing [111], and tested hGMPK in order to explore whether it would sensitize HEK293 cells in the presence of the medically most relevant prodrug 6-thioguanine. This question was important because if wild-type hGMPK increases the sensitivity of cells towards 6-TG, then the next step would be testing an engineered enzyme to further reduce high doses of the drug required for cell killing.

For that purpose, I generated two fusion constructs of hGMPK with N-terminal EGFP (EGFP-hGMPK[pEGFP-C1]) and C-terminal EGFP (hGMPK-EGFP[pEGFP-N1]) as shown below:

1. hGMPK with C-terminal EGFP (pEGFP-N1):
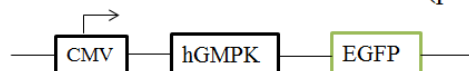
CMV — hGMPK — EGFP

2. hGMPK with N-terminal EGFP (pEGFP-C1):

CMV — EGFP — hGMPK

3. Control (pEGFP-C1):

CMV — EGFP

3.1.4.1 Stable HEK293 cell lines expressing EGFP-hGMPK and hGMPK-EGFP

The two constructs of hGMPK with N-terminal and C-terminal EGFP fusion tags were transfected into HEK293 cells to generate stable cell lines (see section 2.2.12.1). The C-terminal EGFP construct gave rise to very low transfection efficiency of hGMPK as compared to the N-terminal EGFP. As the N-terminal EGFP had no adverse effects on the activity of the recombinantly produced hGMPK (using hGMPK[pET-14bEGFP-N] construct, data not shown), we decided to use the N-terminal EGFP fusion construct. The EGFP-hGMPK protein as observed by confocal microscope was highly expressed (Fig. 3.12a). Its stable cell lines were generated. The pEGFP-C1 plasmid expressing only EGFP was used as a control. Once stable transfection was achieved, expression levels of the whole cell lysates of EGFP-hGMPK and EGFP were analyzed by immunoblotting using polyclonal antibody against GFP (Fig. 3.12b). The immunoblot showed similar expression levels of EGP-hGMPK and EGFP.

a.

EGFP   EGFP/DAPI overlay

EGFP-hGMPK   EGFP-hGMPK/DAPI overlay
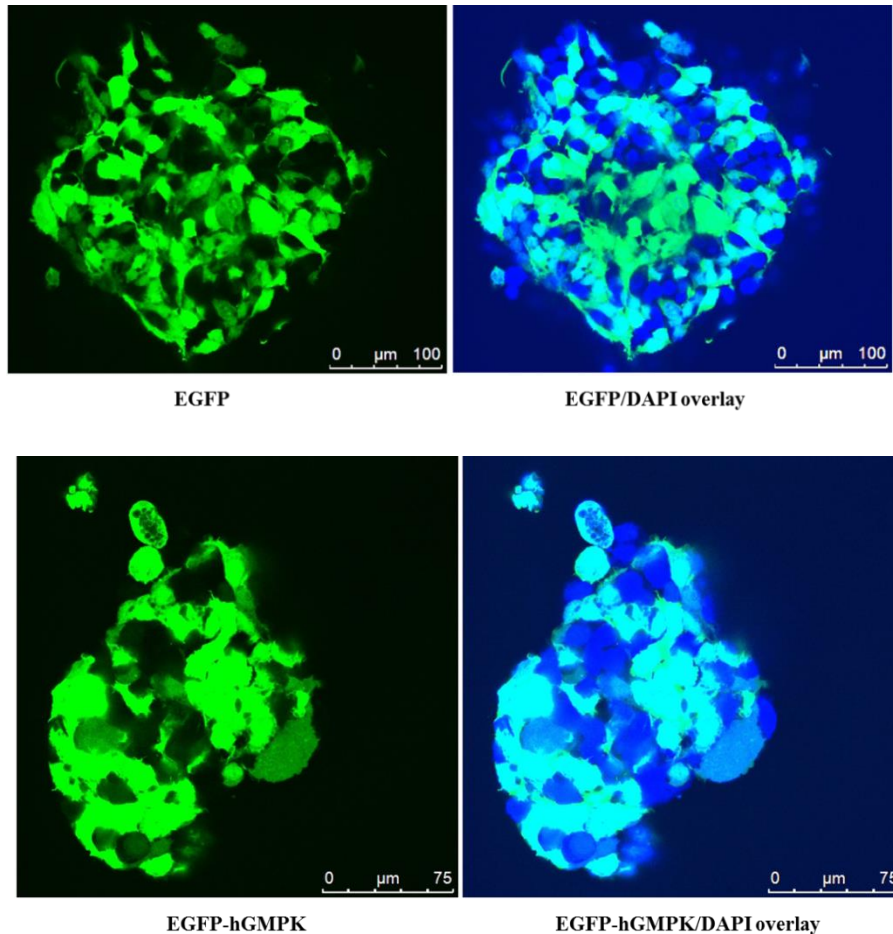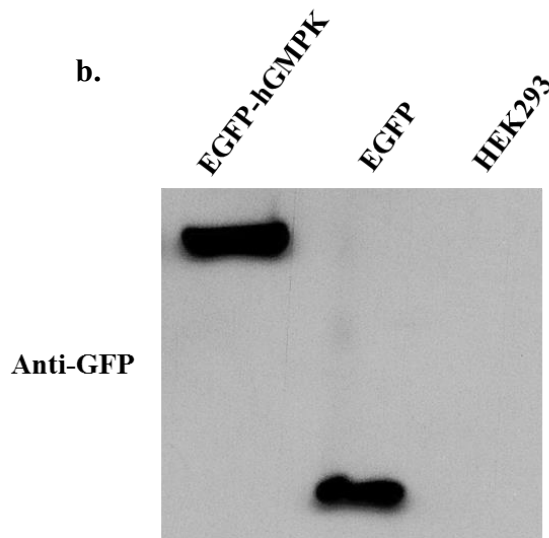
Fig. 3.12. Confocal images and expression levels of EGFP–hGMPK and EGFP-transfected HEK293 cell lines. (a) Confocal images were taken with Leica TCS SP5 with 40x objective. The nuclei were stained with DAPI. (b) Immunoblot of HEK293 cell lysates harboring EGFP-hGMPK and EGFP. Equal amounts (20 µg) of samples were loaded on 12.5 % SDS-PAGE. Proteins were transferred to a nitrocellulose membrane, and immunoblotting was performed using anti-GFP IgG antibody.

3.1.4.2 Sensitivity of EGFP-hGMPK-expressing stable cell line to 6-thioguanine

The sensitivity of stable cell lines expressing EGFP-hGMPK and EGFP was tested against varied concentrations of 6-thioguanine (6-TG) by using the MTT assay as described before (section 2.2.12.2). All cell lines were treated with 6-TG in the range of 0.001-1000 µM in 24 well plates incubated at 37 ^{o}C for 48 h. Each concentration of 6-TG was tested in triplicate. The dose-response curves were plotted for EGFP-hGMPK and EGFP by Gnuplot 5.0 software using a nonlinear regression (Fig. 3.13a & b). The EC$_{50}$ value of EGFP-hGMPK was ~6 times lower than the control EGFP cells (Fig. 3.13c). Expression of hGMPK sensitized the cells even at 10 nM concentration (3rd data point in the plot) of 6-TG whereas the control EGFP-expressing cells were not affected at this concentration.

a.

EGFP

b.

EGFP-hGMPK

c.

HEK293 stable cell line	6-TG EC$_{50}$ ± SD (µM)
EGFP-hGMPK	~10 ± 3
EGFP	~62 ± 7
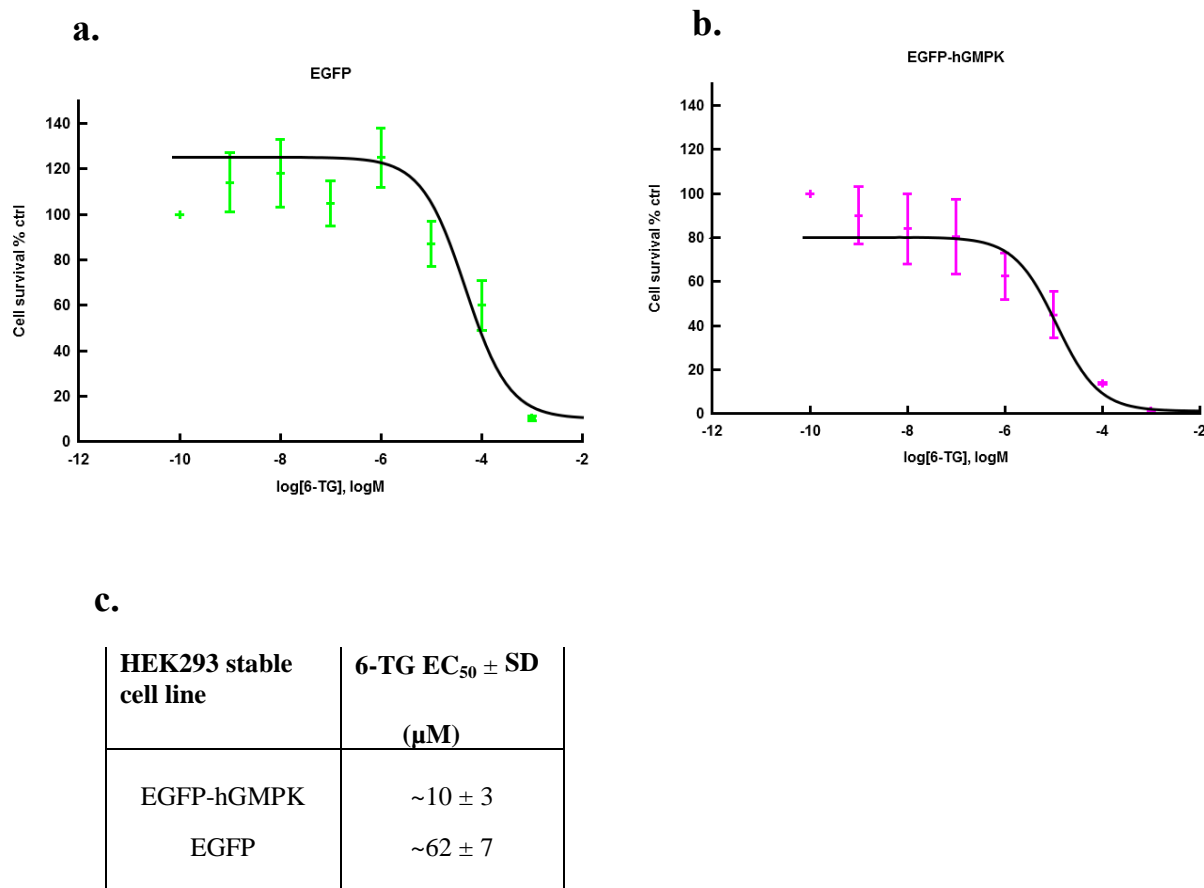
Fig. 3.13. Sensitivity of EGFP-hGMPK and EGFP-expressing HEK293 stable cell lines to 6-thioguanine. (a) & (b), 6-thioguanine (6-TG) dose-response curves of EGFP and EGFP-hGMPK-expressing stable cell lines, respectively. Cell survival was measured by the MTT assay after 48 h incubation with 6-TG. Untreated samples (0 µM 6-TG) were set at 100% cell survival. Each data point (mean ± SD, n=3) is indicated as percentage of the value for control wells with no 6-TG. Sigmoidal curves were fitted to data points of stable cell lines over a 10^6-fold range of 6-TG concentrations (0.001-1000 µM) using Gnuplot 5.0 software (nonlinear regression fitting). (c) Determination of 6-TG EC$_{50}$ for EGFP-hGMPK and EGFP HEK293 stable cell lines. EC$_{50}$ values are shown as mean ± SD.

3.1.5 Electrochemical detection of guanosine monophosphate with a quantum dot-based biosensor modified with hGMPK

A new approach has been developed to detect kinase-catalyzed reactions based on the light-triggered electrochemical sensing of NADH in a three-step coupled-enzyme assay as described in section 2.2.7. We demonstrated the proof of concept of using a photochemical sensor for the

electrochemical detection of a nucleotide kinase-catalyzed reaction. This novel technique senses the change in photocurrent in a GMP concentration-dependent way by a CdS/ZnS quantum dot (QD)-based electrode [87, 88].

First, we tested the validity of our approach for the control reaction between pyruvate (the product of pyruvate kinase in step 2) and NADH catalyzed by lactate dehydrogenase (LDH). The sensor electrode was composed of CdS/ZnS QDs which were layered on a gold electrode via stilbenedithiol (StDT). Varied concentrations of pyruvate substrate were used in the reaction buffer (100 mM HEPES pH 7.5, 100 mM KCl and 20 mM $MgCl_2$) in a 1 ml electrochemical cell container. Electron hole pairs were generated upon illumination of QDs. The electron transfer takes place in between the CdS/ZnS QDs electrode and the $NADH/NAD^+$ redox pair in solution. The current "I" was recorded before and after the reaction was completed at a constant bias voltage, U = +50 mV. It was found that the CdS/ZnS-modified gold electrode was acting as a transducer for the analysis of pyruvate which depletes NADH by oxidizing it into NAD^+ in the LDH-catalyzed reaction. The change in pyruvate concentration was directly proportional to the change in photocurrent under pulsed illumination at a constant biased potential of +50 mV against Ag/AgCl (Fig. 3.14a & b).

a.

b.
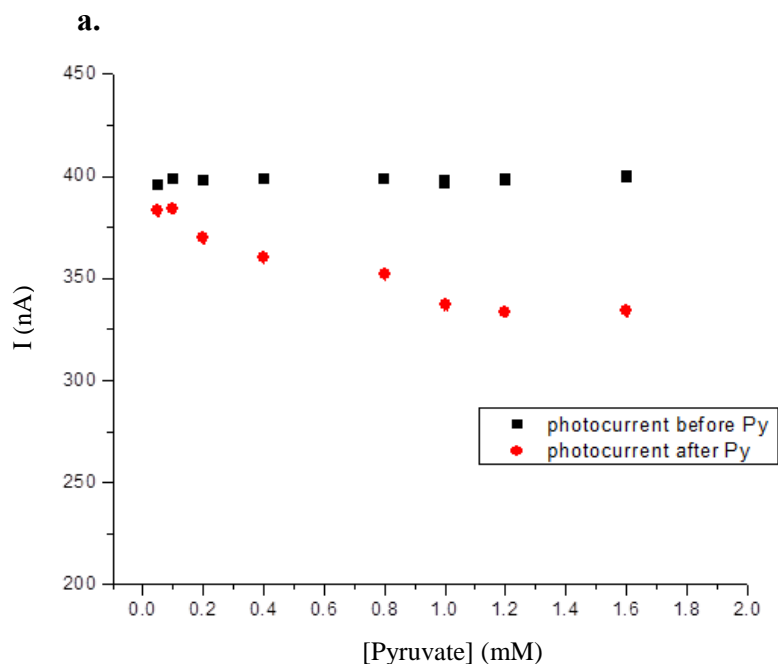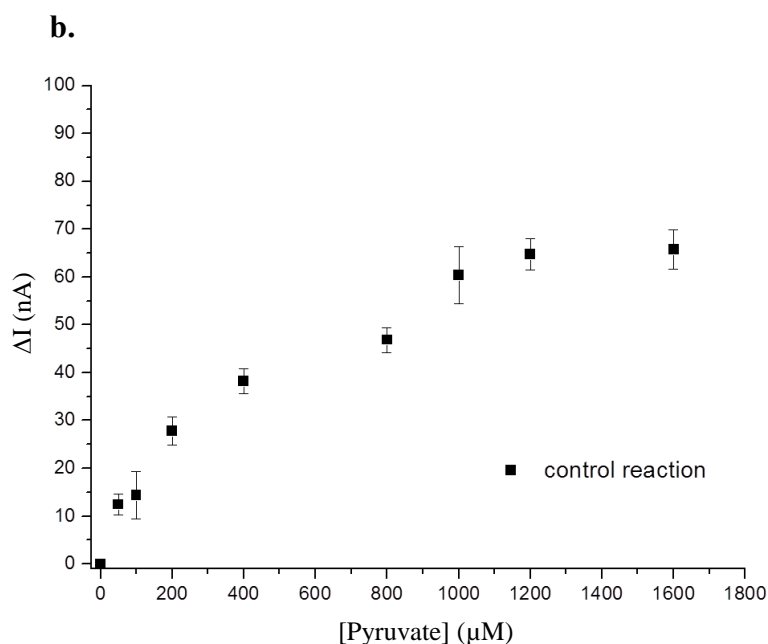
Fig. 3.14. **Dose-response curve for the detection of pyruvate in the control reaction.** (a) Photocurrent detected before and after adding pyruvate (Py) to the reaction mixture. (b) The control reaction as shown below was catalyzed by lactate dehydrogenase (LDH) in a buffer (100 mM HEPES pH 7.5, 100 mM KCl and 20 mM $MgCl_2$) containing 1.2 mM NADH at a constant bias potential of +50 mV and 25 $^{\circ}$C. The concentration of pyruvate was varied from 50 μM to 1600 μM. Before adding pyruvate to the reaction mixture, the current was recorded. Upon addition of pyruvate, the reaction was allowed to complete for 5-10 min. The current was detected again and the change in current "ΔI" in nA was plotted against the respective pyruvate concentrations "c" in μM using the OriginLab software (version 8.1). Data points are represented as means ± standard deviation of triplicate measurements.

- $\text{Pyruvate} + \text{NADH} + \text{H}^+ \underset{}{\overset{\text{LDH}}{\longleftrightarrow}} \text{Lactate} + \text{NAD}^+$
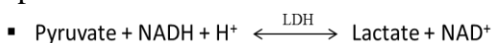
A linear relationship was observed between pyruvate concentration and the change in photocurrent. We coupled this control reaction to two other steps in the coupled assay as depicted in Materials and Methods section 2.2.7. In this case, the concentration of GMP (substrate of hGMPK) was altered in the range of 50 μM-1600 μM at constant concentrations of hGMPK and all other constituents of the assay. Each time, 5-10 min were provided for the reaction to complete. The change in photocurrent (ΔI) was measured upon illumination of CdS/ZnS QDs before and after adding the GMP substrate to the reaction mixture. It was plotted against the respective GMP concentrations using OriginLab 8.1 software (Fig. 3.15a & b). The data showed

that GMP can be detected from a very low concentration like 50 μM to as high as 1200 μM under the specified experimental conditions.

a.

b.
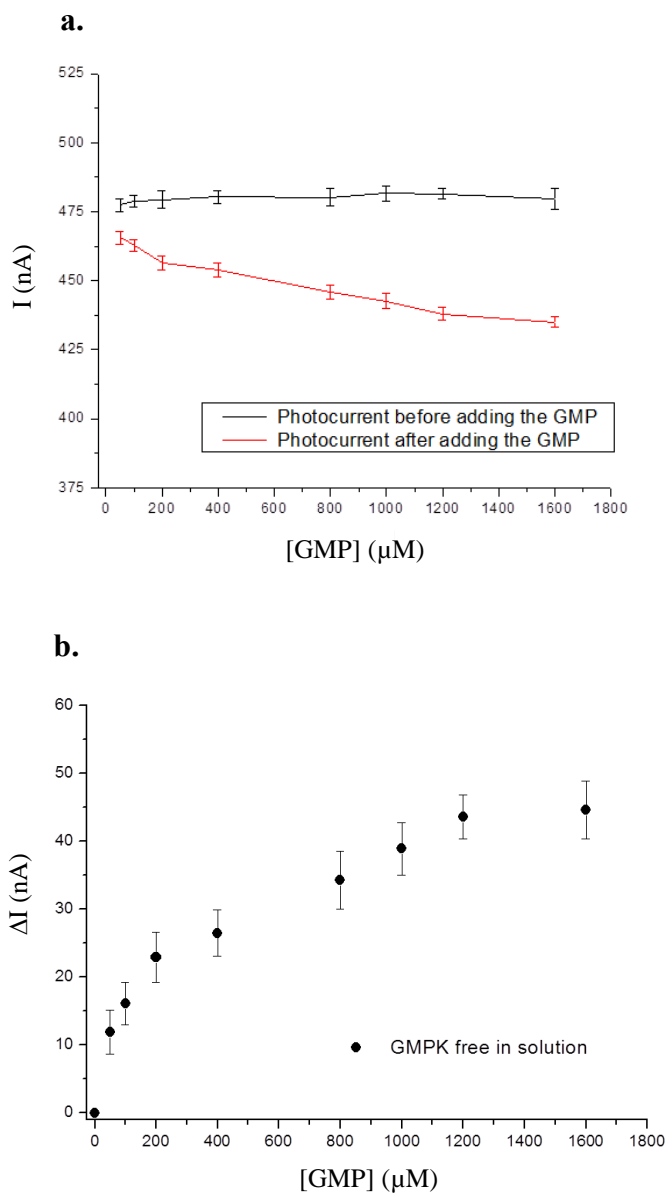
Fig. 3.15. Dose-response curve for the detection of GMP in the electrochemical detection assay (see section 2.2.7). (a) Photocurrent detected before and after adding GMP to the reaction mixture. (b) The first reaction in the three-step coupled assay was catalyzed by hGMPK in the buffer (100 mM HEPES pH 7.5, 100 mM KCl and 20 mM MgCl$_2$) at a constant bias potential of +50 mV and 25 ^{o}C. The current was recorded before adding the GMP substrate (50 μM-1600 μM) to the reaction mixture. The reaction was allowed to complete for 5-10 min and the current was detected again. The change in current "ΔI" in nA was plotted against the respective GMP concentration in μM using the OriginLab 8.1 software. Data points are represented as means ± standard deviation of triplicate measurements.

124

To shape our electrochemical detection system into a photosensor format, we immobilized the enzymes on the CdS/ZnS QDs electrode using layer-by-layer assembly of the polyelectrolytes, PSS (poly (styrene sulfonate)) and PAH (poly (allylamine hydrochloride)), as shown in Fig. 3.16 [114, 115].
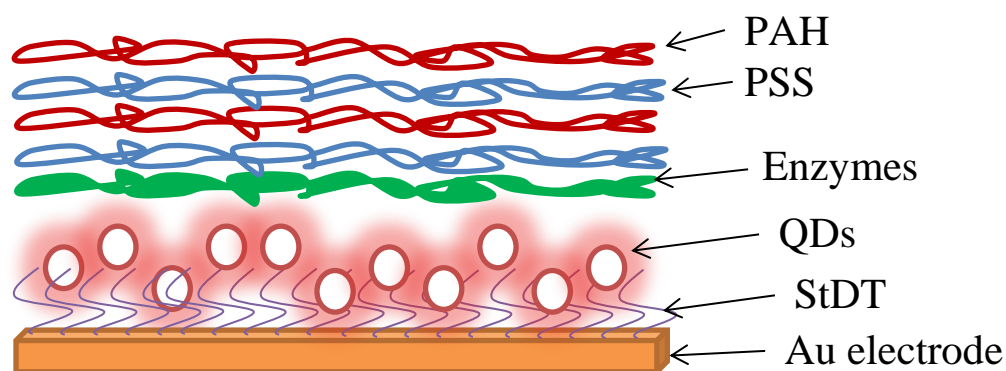
Fig 3.16. Human GMPK immobilized on QDs/StDT/Au using polyelectrolyte bilayers. CdS/ZnS QDs were attached to the Au (gold) electrode surface via StDT (stilbenedithiol). The hGMPK and two helper enzymes (PK and LDH) were immobilized on the electrode surface by two bilayers of PSS (poly (styrene sulfonate)) and PAH (poly (allylamine hydrochloride)) polyelectrolytes using the layer-by-layer (lbl) assembly technique.

Human GMPK and two helper enzymes (PK and LDH) were deposited on the QDs electrode by two consecutive bilayers of PSS/PAH, with PAH forming the outermost layer. The overall structure of the assembled electrode can be expressed as $(PSS/PAH)_2$/hGMPK;PK;LDH/QDs/ StDT/Au/Si. This geometry was used for the detection of the GMP-dependent oxidation current as shown in the dose-response curve in Fig. 3.17.

a.

b.
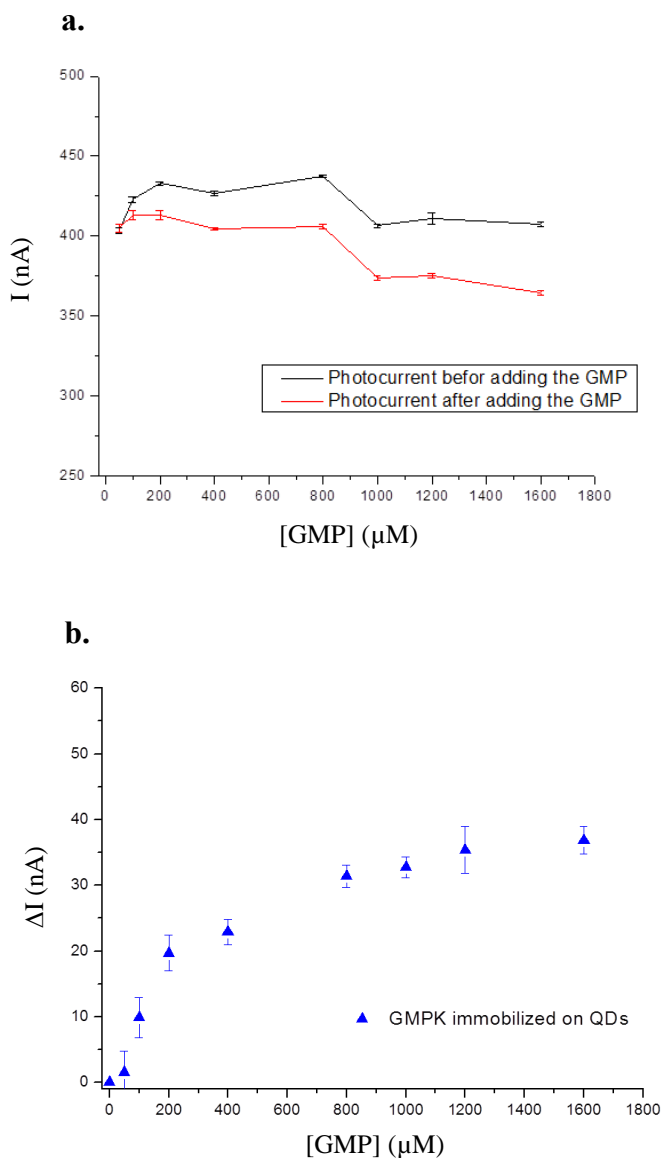
Fig 3.17. Dose-response curve for the detection of GMP by human GMPK immobilized on QDs/StDT/Au electrode surface. (a) Photocurrent detected before and after adding GMP to the reaction mixture. (b) The hGMPK enzyme was immobilized on the gold electrode surface at the bottom of the electrochemical cell using PSS and PAH polyelectrolytes bilayers (the specific geometry is shown in Fig. 3.16). The first reaction in the three-step electrochemical detection assay (see section 2.2.7) was catalyzed by hGMPK in the buffer (100 mM HEPES pH 7.5, 100 mM KCl, 20 mM MgCl$_2$), and detected at a constant bias potential of +50 mV and 25 $^\circ$C. The current was recorded before adding GMP to the reaction mixture. Various amounts of GMP (50 μM-1600 μM) were added to the reaction cell. The reaction was allowed to complete for 5-10 min, and the current was detected again. The change in current "ΔI" in nA was plotted against the GMP concentration in μM using the OriginLab 8.1. Data points are represented as means $\pm$ standard deviation of triplicate measurements.

The data revealed that the concentration of GMP can be detected within the range of 50 µM to 800 µM by using this biosensor. Above 800 µM, the photocurrent response was saturated. The deposition of hGMPK in the above-mentioned setup successfully showed activity which is an important precondition for the sensor. However, the maximum current detected (37 nA) was comparatively lower than that of free enzyme in solution (45 nA). It may be due to the electrostatic interactions of the nucleotide substrates with the negatively (PSS) and positively (PAH) charged polyelectrolytes. The three dose-response curves mentioned above are compared in Fig. 3.18.
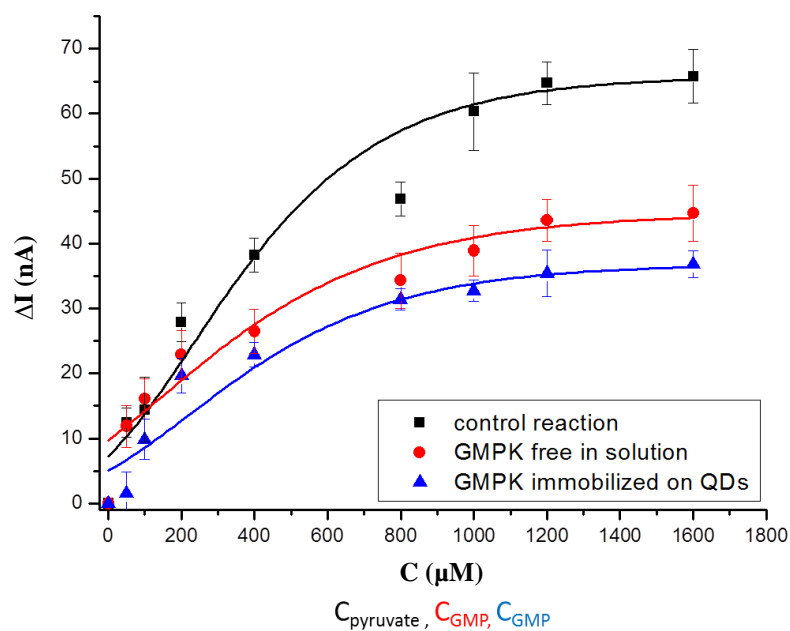
Fig. 3.18. Comparison of the dose-response curves for the detection of GMP in the three-step electrochemical detection assay and pyruvate in the one-step control reaction. The change in current ΔI (nA) was plotted against the respective concentrations "C" of the substrate in µM: pyruvate as "$C_{pyruvate}$" and GMP as "C_{GMP}" using the OriginLab 8.1 software. The control reaction is catalyzed by pyruvate kinase (PK).

3.1.6 Human GMPK-catalyzed reactions in polyelectrolyte containers of different shapes and sizes

For the synthesis of calcium carbonate ($CaCO_3$) microparticles of different shapes and sizes, the precipitation reaction was optimized by varying the concentrations of calcium chloride ($CaCl_2$) and sodium carbonate (Na_2CO_3) solutions and stirring at 500 rpm for 30 min. To trigger the nucleation for anisotropic $CaCO_3$ particles, initial salt concentrations and the ethylene glycol contents in the reaction medium were altered. Exterior and interior surface morphology, crystallinity, porosity and loading capacity of resulting $CaCO_3$ particles were characterized. This work was done in collaboration with Dr. Andre Skirtach and Dr. Alexey Yashchenok from the Department of Interfaces, Max-Planck Institute of Colloids and Interfaces, Golm/Potsdam.

As shown in Fig. 3.19, the initial salt concentration ratio plays a major role in the shaping process of $CaCO_3$ particles. At a very high initial salt concentrations mixed in equal molar ratio ($CaCl_2$: Na_2CO_3, 1 M: 1 M), the $CaCO_3$ nuclei tend to form rhomboidal structures (Fig. 3.19a & d). The axis ratio of the fabricated structures varies between 1.68-1.75. The shape of the $CaCO_3$ particles shifted to elongated rhomboidal (or distorted ellipsoidal) when the amount of calcium ions are 10 times lower than the amount of carbonate ions (0.1 M: 1 M), (Fig. 3.19b & e). The axis ratio of these particles changes between 2.70-2.93. Spherical particles were obtained when the amount of carbonate ions was 10 times lower than the amount of calcium ions (1 M: 0.1 M), (Fig. 3.19c & f).

Due to the rapid nucleation process, secondary nucleation was also observed which occured on the side surfaces resulting in irregular star-like, flower-like structures (Fig. 3.19a, b, g & h). However, we reduced the amount of these irregular structures from 35% to 15% by changing the solubility of salts via alteration of the ethylene glycol content (66.6-80%).

In order to produce anisotropic layer-by-layer (LbL) capsules, particles were loaded with hGMPK enzyme. The enzyme-loaded particles were coated with consecutive LbL of positively charged PAH and negatively charged PSS (Fig. 3.20). The amount of protein loaded per particle was calculated by using Bradford dye-binding assay [147]. Actually, the protein loaded into the microparticles was determined indirectly by measuring the protein concentration and volume of the supernatant left after each protein loading step. By subtracting the amount of protein lost in the supernatant and during each washing step, the total protein entrapped in the particles was

calculated. The particles were counted in a hemocytometer to estimate the average amount of protein loaded per particle. In future work, more direct methods could be utilized like mass-sensing with resonating microchannels due to its ability to measure single nanoparticle mass with sub-femtogram resolution [146]. The catalytic activity of the encapsulated enzyme was determined after coating with one and two bi-layers of polyelectrolytes using the standard NADH-dependent spectroscopic assay (see section 2.2.6). The specific activities for the hGMPK-loaded microparticles are summarized in Table 3.4. The highest loading capacity was observed for star particles, and the lowest was for cube particles. Nevertheless, after LbL coating, the activity of ellipsoidal microparticles was highly retained whereas that of other microparticles was decreased by increasing the number of coating layers up to four successive layers.
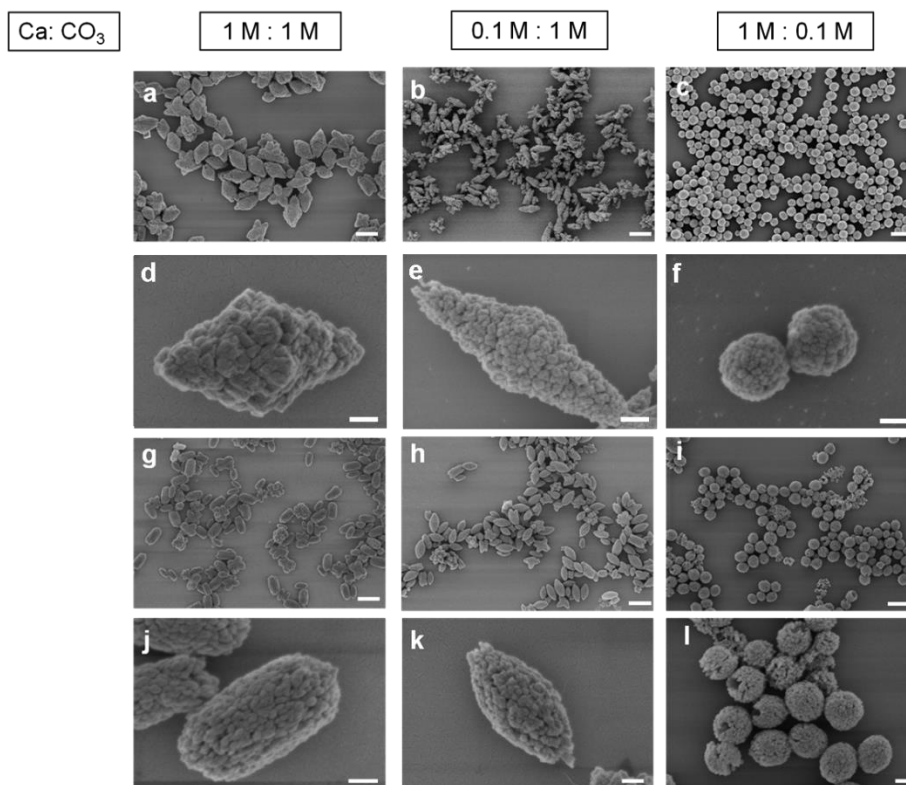
Fig. 3.19. Scanning electron microscopy (SEM) images of calcium carbonate particles. The calcium carbonate particles were synthesized by mixing $CaCl_2$ and Na_2CO_3 salts at 500 rpm for 30 min in the presence of 66.6 % ethylene glycol (a-f), in 80 % ethylene glycol (g-l) medium. The bars on images (a-c) and (g-I) are 2 µm, (d-f) and (j-l) are 300 nm, respectively. Ca: CO_3 at the upper left side of the figure represents the concentration ratios between $CaCl_2$ and Na_2CO_3 salts, respectively.

Table 3.4. Specific activities of encapsulated human GMPK in polyelectrolyte containers of different shapes and sizes

Polyelectrol-yte layers	0.5 µm spherical particles (U/mg)	3-5 µm spherical particles (U/mg)	Ellipsoidal particles (U/mg)	Star particles (U/mg)	Cube particles (U/mg)
No layer, microparticles with adsorbed hGMPK	4.55	14.47	14.1	18.76	0.58
1st layer coated (PAH)	1.69	10.18	13.8	16.1	0.44
2nd layer coated (PSS)	0.54	4.82	12.4	8.0	0.29
4 layers coated (PAH→PSS→ PAH→PSS)	0.27	3.22	11.62	5.1	0.16

The confocal and wide field images of rhomboidal, ellipsoidal and spherical calcium carbonate microparticles are shown in Fig. 3.20 a-e.
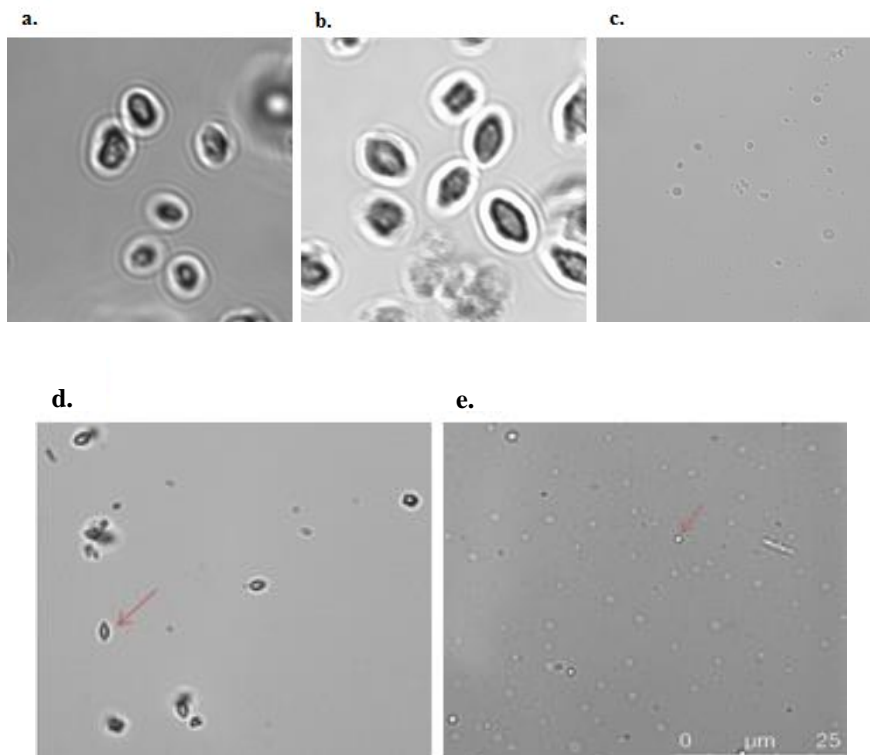
Fig. 3.20. Confocal and wide-field images of calcium carbonate microparticles. a, b & c are the confocal images of rhomboidal, ellipsoidal and spherical microparticles coated with polyelectrolytes (PAH and PSS) layers, respectively. d & e are the wide-field images of ellipsoidal and spherical microcapsules.

3.1.7 A novel spectrophotometric and fluorometric enzyme-coupled assay for human GMPK

Human guanylate kinase (hGMPK), like other nucleoside/nucleotide kinases, is an important intracellular enzyme playing a critical role in the activation of several nucleoside analog prodrugs. The conventional assay used for activity measurements of nucleoside and nucleotide kinases is based on the absorbance of NADH at 340 nm [84, 99]. Nevertheless, there are certain disadvantages to this assay. One problem is that nucleoside/nucleotide analogs such as 6-thioguanosine and 6-thioGMP absorb light at 340 nm making it aberrant to use the NADH-dependent assay for monitoring phosphorylation of these nucleotides (Fig. 3.21). Besides that, it

was highly useful to devise a more sensitive assay which could be used both in absorbance and fluorescence modes for the kinetic analysis of nucleoside/nucleotide kinases. For that purpose, we developed a four-step enzyme-coupled assay which can be used in absorbance mode at 570 nm and in fluorescence mode at λ_{ex}= 568 nm and λ_{em}= 584 nm (see Materials and Methods section 2.2.8).
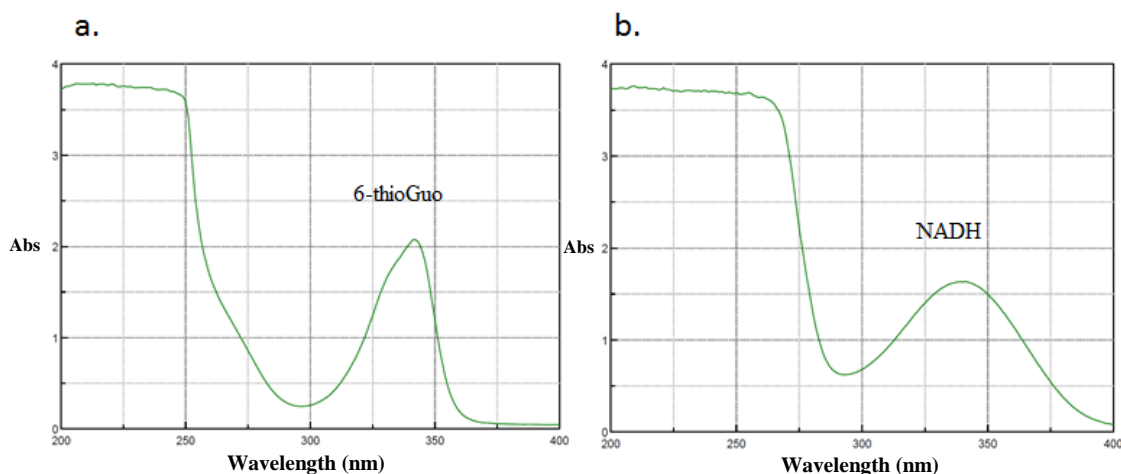
Fig. 3.21. Absorbance maxima of 6-thioguanosine and NADH. Both 6-thioguanosine (6-thioGuo) and NADH have λ_{max} at 340 nm.

The assay was optimized by using the natural substrate of hGMPK, GMP. In the first step of the assay, GMP is phosphorylated to GDP by hGMPK using ATP as a phosphoryl group donor. In the second step, GDP and ADP are converted to GTP and ATP by pyruvate kinase using phosphoenolpyruvate as a phosphate donor resulting in the formation of pyruvate. The pyruvate is reacted with inorganic phosphate and oxygen in the third step catalyzed by pyruvate oxidase producing hydrogen peroxide (H_2O_2), carbon dioxide, and acetyl phosphate. In the fourth step, H_2O_2 is used by horseradish peroxidase (HRP) to oxidize nonfluorescent Amplex Red to resorufin which has excellent absorbance as well as fluorescence properties (see section 2.2.8). The pyruvate oxidase (lpPOX) used in this study was from *Lactobacillus plantarum*, and the lpPOX containing plasmid was kindly gifted by Prof. Kai Tittmann from the department of Bioanalytics, Georg-August-Universität Göttingen. I amplified the ORF (1809 bp) of the enzyme with the desired NdeI and BamHI restriction sites for cloning into the pJC20HisN expression vector. The protein was purified by affinity chromatography using Protino Ni-IDA resin

(Macherey-Nagel) and by gel filtration chromatography using a Superdex 200 column (GE Healthcare) (Fig. 3.22).
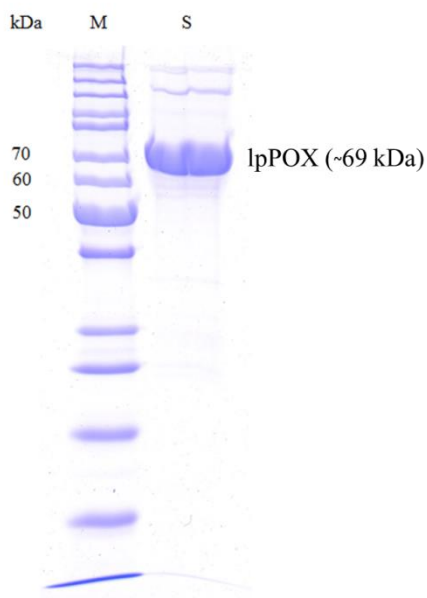
Fig. 3.22. SDS-PAGE of pyruvate oxidase. Pyruvate oxidase (lpPOX) from *Lactbacillus plantarum* was purified by affinity chromatography using Protino Ni-IDA resin (Macherey-Nagel), followed by gel filtration chromatography using a Superdex 200 column (GE Healthcare). Purity of the homogeneous lpPOX was tested by SDS-PAGE (12%). Lanes: M, marker proteins (PageRuler Unstained Protein Ladder, Thermo Scientific), and S is lpPOX with N-terminal decahistidine tag (~69 kDa).

All four reactions in the coupled assay were optimized in a stepwise manner starting from reaction 4, then combined reaction 3 and 4, afterwards 2, 3 and 4, and finally all four reactions (data not shown). The final reaction mixture for the standard assay in the absorbance mode contained 2 mM ATP, 5 nM hGMPK, 2 mM PEP, 1 unit of pyruvate kinase (PK), 1.2 µM of pyruvate oxidase (lpPOX), 0.2 mM thiamine pyrophosphate (TPP), 10 µM flavin adenine dinucleotide (FAD), 25 µM Amplex Red (resorufin, $\varepsilon = 5.4 \times 10^4$ M^{-1}cm^{-1} at 570 nm) and 0.1 µM HRP. The buffer-H (2X) used consisted of 100 mM HEPES pH 7.2, 20 mM MgCl$_2$, 100 mM KCl and 50 mM potassium phosphate. The concentration of GMP was varied in a final reaction

volume of 200 µl, and the absorbance was measured at 570 nm, 25 $^{\circ}$C. As the reactions proceed, Amplex Red in the final step is converted into resorufin which gives high absorbance at 570 nm.

Before determining the steady-state kinetic parameters, it was necessary to test the linearity of the reaction velocity as a function of the hGMPK concentration as shown in Fig. 3.23b. It indicated that hGMPK can be used in the concentration range of 0.5 nM to 18 nM. Using 5 nM hGMPK, the concentration of GMP was varied up to 20 K_m. Turnover rates (k_{obs}) were calculated, and the values were fit to the Michaelis-Menten equation (see equation 1 in section 2.2.6) (Fig. 3.23a) using Gnuplot 5.0 software. The K_m calculated was 30 ± 5 µM, and k_{cat} was 80 ± 4 s^{-1}. These values are in good agreement with parameters determined by the standard NADH-dependent spectroscopic assay (see section 2.2.6). The assay protocol as mentioned above can also be used in the fluorescence mode at excitation wavelength λ_{ex} = 568 nm and emission wavelength λ_{em} = 584 nm [116].
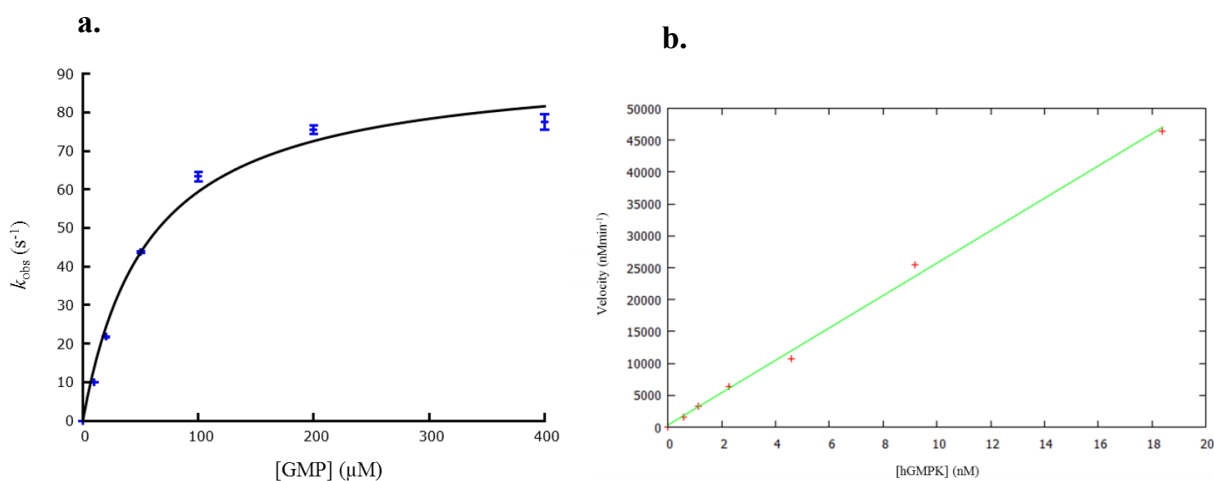
Fig. 3.23. Validation of the spectrophotometric and fluorometric coupled-assay in absorbance mode. (a) Michaelis-Menten plot of hGMPK. Steady-state turnover rates (k_{obs}) were measured for different concentrations of GMP using the newly developed spectrophotometric and fluorometric enzyme-coupled assay based on the use of Amplex Red. The Michaelis-Menten plot was generated by using Gnuplot 5.0 software. All measurements were performed at 25 $^{\circ}$C in 50 mM HEPES pH 7.2, 10 mM MgCl$_2$, 50 mM KCl and 25 mM potassium phosphate using 5 nM hGMPK. Error bars indicate ± standard deviation of triplicate measurements. (b) Different concentrations of hGMPK plotted versus the reaction rate in nM min^{-1}.

3.2 Biochemical characterization of human mitochondrial thymidine kinase

3.2.1 Cloning, expression, and purification of recombinant hTK2

Five isoforms of human mitochondrial thymidine kinase (hTK2) are reported in the UniProt database among which the long isoform (UniProt O00142-1) is 265 amino acids (aa) in length [138]. For higher expression in *E.coli*, we used a synthetic and codon-optimized DNA of the long isoform of hTK2. Based on the comparison of hTK2 with the closely related nucleoside kinases from mouse (81% identity), rat (80% identity), and *Drosophila melanogaster* deoxyribonucleoside kinase (46% identity), and its IUPred (intrinsically unstructured proteins prediction tool) and secondary structure prediction analyses as indicated in Fig. 3.24, I generated a series of six truncated versions of hTK2 by PCR. These truncated forms were designated as hTK2-Δ44N (N-terminal 44 amino acids were truncated, 222 aa), hTK2-Δ44N/Δ8C (N-terminal 44 aa and C-terminal 8 aa were truncated, 214 aa), hTK2-Δ44N/Δ25C (N-terminal 44 aa and C-terminal 25 aa were truncated, 197 aa), hTK2-Δ50N (N-terminal 50 aa were truncated, 216 aa), hTK2-Δ50N/Δ8C (N-terminal 50 aa and C-terminal 8 aa were truncated, 208 aa) and hTK2-Δ50N/Δ25C (N-terminal 50 aa and C-terminal 25 aa were truncated, 191 aa). The main purpose was to remove the N-terminal mitochondrial translocation signal peptide and to determine the effect of truncations on the solubility, stability, and activity of the enzyme [138, 140, 209, 210]. The truncation may also increase the chances of crystallization [143, 175, 177]. Nevertheless, in our hand, hTK2 (full-length and truncated forms) was insoluble and aggregated in inclusion bodies when produced as a fusion with only histidine-tag or with N-terminal His$_6$-SUMO tag (SMT3, 101 aa) in BL21(DE3)pLysS cells.

To get hTK2 protein in native soluble form, I followed a systematic protein expression and purification strategy. First, I tried to dissolve the inclusion bodies in detergents like Triton X-100, urea, CHAPS, sarkosyl, and L-arginine. However, refolding of hTK2 by dialysis against a native buffer did not recover thymidine kinase activity. Then, I ligated and expressed hTK2 in a variety of vectors with different N-terminal tags like His$_6$, His$_6$-SUMO, GST, His$_6$-MBP, and His$_{14}$-MBP-SUMObr. In addition, a number of *E.coli* expression strains were tested including BL21(DE3), BL21(DE3)pLysS, C41(DE3), Origami B(DE3), and Rosetta-gami B(DE3). Moreover, expression conditions were varied, such as using different incubation temperature in the range of 16-37 $^{\circ}$C, IPTG concentrations from 0.1 to 1 mM, induction time in the range of 2-

24 h, and various liquid media including LB, TB, and lactose-containing auto-inducing media were used. After all these trials, finally, I succeeded in optimizing a protocol for obtaining hTK2 in soluble and active form. This protocol is based on expressing hTK2 with N-terminal His$_{14}$-MBP-SUMObr tag (~60 kDa) in BL21(DE3)pLysS cells in LB medium for 16 h induction in the presence of 0.4 mM IPTG at 18 ^{o}C [82, 83]. It was observed that chaperones (GroEL/ES) were consistently copurified with the hTK2 fusion proteins [139]. For removal of these chaperones, I used one extra washing step in the presence of 5 mM ATP. After purification by affinity chromatography using Protino Ni-IDA resin (Macherey-Nagel), the His$_{14}$-MBP-SUMObr tag was cleaved by SUMO-protease. The hTK2 was further purified by anion exchange chromatography using DEAE-Sepharose FF (GE Healthcare Life Sciences). The optimized protocol was applied to all six truncated forms of hTK2 as mentioned above. The final yield obtained was 4-8 mg/l culture. Purity and solubility of hTK2 were tested by 12 % SDS-PAGE as shown in Fig. 3.25 & 3.26 (only two truncated forms are shown). The activity was determined by the NADH-dependent spectroscopic assay (see section 2.2.6). It was found that the double-truncated hTK2-Δ44N/Δ8C had higher catalytic efficiency (k_{cat}/K_m = 9.3 x 10^3 M^{-1}s^{-1}) than the single-truncated hTK2-Δ44N (k_{cat}/K_m = 3.7 x 10^3 M^{-1}s^{-1}). Nevertheless, no activity was detected for the rest of four truncated forms i.e., hTK2-Δ44N/Δ25C, hTK2-Δ50N, hTK2-Δ50N/Δ8C, and hTK2-Δ50N/Δ25C. The two truncated forms of hTK2 were screened for crystallization. After several screening trials, we obtained crystals for hTK2-Δ44N/Δ8C, but unfortunately the quality of the crystals was poor, and they diffracted badly. Besides that, we observed that the crystal formation was not reproducible, and also sometimes the protein tended to aggregate during purification. We hypothesized that hTK2 requires eukaryotic chaperones for proper folding and perhaps undergoes posttranslational modifications; therefore we opted for producing hTK2 in eukaryotic cells such as insect cells [144].

In conclusion, we have developed an optimized protocol to produce hTK2 in soluble and active form in *E. coli*. In addition, we generated a shorter and more active form of hTK2. The crystallization experiments need to be further optimized.

IUPred
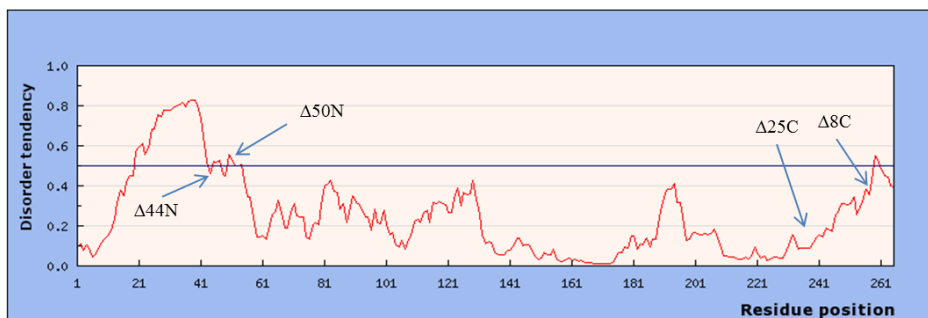
The long isoform hTK2 (265 aa) was truncated into the following forms:
1. hTK2-Δ44N   (222 aa)
2. hTK2-Δ44N/Δ8C (214 aa)
3. hTK2-Δ44N/Δ25C (197 aa)
4. hTK2-Δ50N   (216 aa)
5. hTK2-Δ50N/Δ8C (208 aa)
6. hTK2-Δ50N/Δ25C (191 aa)

Secondary structure prediction tools used:
SOPM (Geourjon and Deléage, 1994)
SOPMA (Geourjon and Deléage, 1995)
HNN (Guermeur, 1997)
MLRC (Guermeur *et al.*, 1998)
DPM (Deléage and Roux, 1987)
DSC (King and Sternberg, 1996)

Fig. 3.24. Truncations of hTK2. Six truncated forms of hTK2 were generated based on the IUPred and secondary structure prediction analysis.
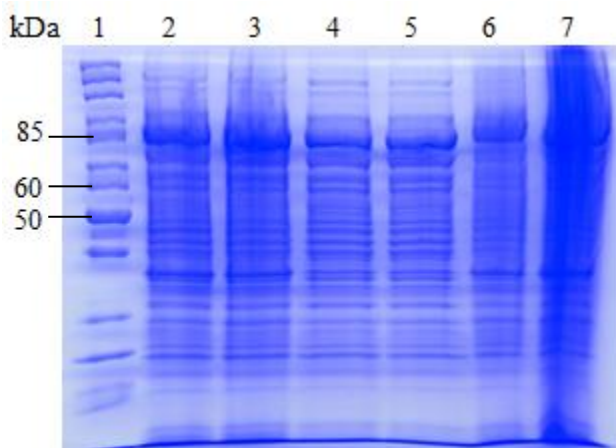
Fig. 3.25. SDS-PAGE of the recombinant hTK2. Two truncated forms of hTK2 were expressed in BL21(DE3)pLysS and purified by affinity chromatography using Protino Ni-IDA resin (Macherey-Nagel). Purity was tested by 12% SDS-PAGE. Lanes: 1, Molecular weight markers (in kDa); 2, cell lysate of fusion hTK2-Δ44N (~86 kDa); 3, cell lysate of fusion hTK2-Δ44N/Δ8C (~85 kDa); 4, supernatant of fusion hTK2-Δ44N ; 5, supernatant of fusion hTK2-Δ44N/Δ8C; 6, cell debris of fusion hTK2-Δ44N; 7, cell debris of fusion hTK2-Δ44N/Δ8C.
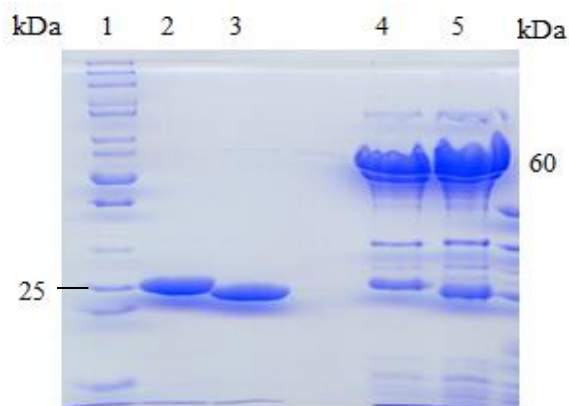
Fig. 3.26. SDS-PAGE of the purified hTK2. After purification by affinity chromatography, the N-terminal His$_{14}$-MBP-SUMObr tag was cleaved by SUMO-protease and removed from the hTK2 samples. Lanes: 1, molecular weight markers; 2, hTK2-Δ44N (26 kDa) after cleavage of the N-terminal tag; 3, hTK2-Δ44N/Δ8C (25 kDa) after SUMO-protease cleavage of the N-terminal tag; 4, 5, His$_{14}$-MBP-SUMObr (~60 kDa) tag removed from the sample after cleavage.

3.2.2 Intracellular localization of hTK2

Most of the mitochondrial proteins which are translated in the cytosol have an N-terminal targeting signal sequence which directs the translocation of the protein into mitochondria. Although biochemical and cell fractionation experiments have shown that TK2 is localized in mitochondria [141], several published cDNA sequences of TK2 do not have N-terminal mitochondrial leader sequences [138, 142]. For mouse mitochondrial thymidine kinase (mTK2), *in vitro* translation and translocation experiments demonstrated that the N-terminal signal sequence directed the import of mTK2 into mitochondria. However, hTK2 as a fusion with GFP (TK2-GFP) when expressed in CHO (Chinese hamster ovary) cells, failed to be transported into mitochondria unlike mitochondrial deoxyguanosine kinase (dGK-GFP) [154]. The mitochondrial import signal was not identified in the predicted primary structure of hTK2. We asked the same question for hTK2, which has a putative 38-amino acid N-terminal signal sequence (MITOPROT, ExPASy server), to test its cellular localization. For that purpose, three constructs were generated, one full-length hTK2 (265 aa) cloned into pEGFP-N1 with C-terminal EGFP, and two N-terminally truncated forms hTK2-Δ44N and hTK2-Δ44N/Δ8C ligated into EGFP-C1 with N-terminal EGFP for comparison. They were transfected into HEK293 cells, and the expression of fusion constructs was observed by confocal microscopy as shown in Fig. 3.27a-c. The full-length

hTK2 with the N-terminal mitochondrial signal sequence was translocated into mitochondria. In contrast, the two truncated forms were uniformly distributed in the cytoplasm.

a.

b.   **c.**
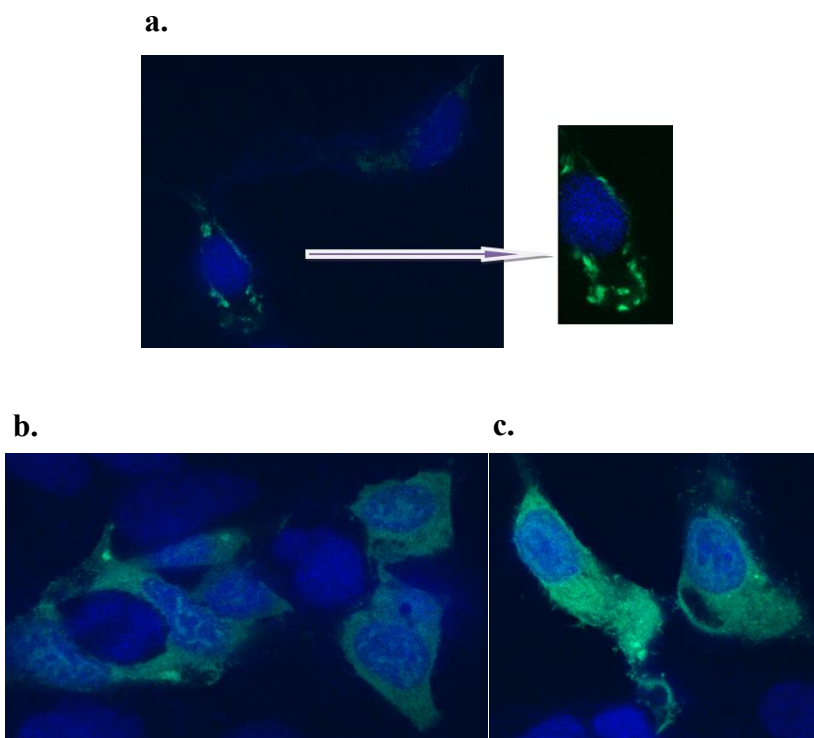
Fig. 3.27. Subcellular localization of hTK2. Confocal images were taken with Leica TCS SP5 with 40x objective. The nuclei were stained with DAPI. (a) Mitochondrial translocation of full-length hTK2 by transfecting HEK293 cells with the fusion construct hTK2-EGFP [pEGFP-N1]. (b) Cytosolic localization of a single truncated hTK2 by transfecting HEK293 cells with EGFP-hTK2-Δ44N[pEGFP-C1]. (c) Cytosolic localization of a double-truncated hTK2 by transfecting HEK293 cells with EGFP-hTK2-Δ44N/Δ8C[pEGFP-C1].

3.2.3 Screening of hTK2 mutants for improved activity towards nucleoside analogs

Seven libraries of hTK2 mutants were generated by error-prone PCR and were screened for their enhanced sensitivity towards anticancer and antiviral nucleoside analogs including gemcitabine, AZT (3´-azido-thymidine), Ara-C (arabinofuranosyl cytidine), BVDU (bromovinyldeoxyuridine), GCV (ganciclovir), cladribine, and 5-fluorouracil. All mutants were ligated into the pGEX-RB vector and were transformed into the TK-deficient *E.coli* strain KY895 [91, 92]. In the first screening step, all transformed colonies of mutants were plated on a TK-selection plate. Only

those mutants harboring thymidine kinase activity could grow on this plate. About 200 colonies were picked and used for streaking on M9 plates containing the nucleoside analogs as mentioned above with varied concentrations. The colonies with improved metabolic phosphorylation of nucleoside analogs would be more sensitive and could not grow. The LD_{100} (the lowest concentration of a nucleoside analog that causes 100% lethality of the KY895 strain transformed with a hTK2 mutant in pGEX-RB vector) was calculated for all 200 colonies, and it was found that mutants M5 and M17 showed 25 times enhanced sensitivity (LD_{100} of 5 μM) towards gemcitabine as compared to wild-type hTK2-Δ44N/Δ8C (LD_{100} of 125 μM). Similarly, fourteen mutants had 10 times higher sensitivity towards gemcitabine. Moreover, five mutants displayed three times lower LD_{100} values for AZT. Nevertheless, the amino acid distribution of these mutations needs to be determined by sequencing their ORF (open reading frames). In addition, these mutants will have to be produced as recombinant proteins and their activities on the corresponding nucleoside analogs need to be determined by the NADH-dependent spectroscopic assay [99].

3.3 Biochemical characterization of *E.coli* guanosine-inosine kinase

3.3.1 Cloning, expression, and purification of recombinant *E.coli* guanosine-inosine kinase

The *E.coli* guanosine-inosine kinase (ecGSK) uses both guanosine and inosine as substrates [148, 149]. As the three-dimensional structure of ecGSK is not known, we aimed to recombinantly produce the enzyme and to crystallize it. For this purpose, the 1302 bp (434 aa) ORF (open reading frame) of ecGSK was amplified by PCR using the *E.coli* XL1-Blue genomic DNA as a template. It was cloned into our standard pET-14bSUMOΔThr plasmid and expressed in BL21(DE3)pLysS. The enzyme was initially purified by affinity chromatography using Protino Ni-IDA resin, and then by anion exchange chromatography using DEAE Sepharose FF column and by gel filtration chromatography using Superdex 200 column. The purity of the enzyme was tested by 12% SDS-PAGE (Fig. 3.28). A good yield of ecGSK was obtained, about 20 mg/l culture. The activity was determined by the conventional NADH-dependent spectroscopic assay (see section 2.2.6) using guanosine as a substrate, resulting in K_m of 55 μM and k_{cat} of ~4 s^{-1}.
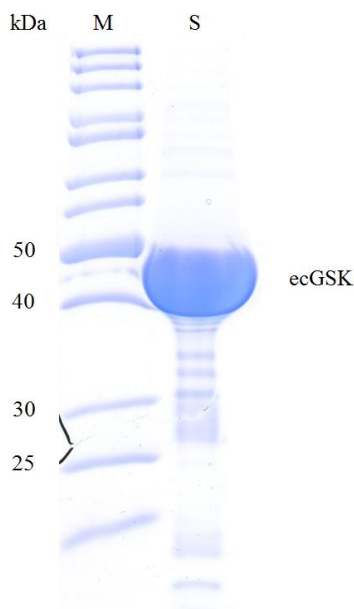
Fig. 3.28. SDS-PAGE of the purified ecGSK. The *E.coli* guanosine-inosine kinase (ecGSK) was purified by affinity chromatography using Protino Ni-IDA resin, followed by anion exchange chromatography using DEAE Sepharose FF (1 ml) and finally by gel filtration chromatography using Superdex 200 column. The purity was tested by 12 % SDS-PAGE. Lanes: M, marker proteins; S, ecGSK (48.45 kDa).

3.3.2 Structural characterization of ecGSK

The ecGSK with final concentration of 35 mg/ml (0.72 mM) containing 7 mM AMP-PNP and guanosine was screened for crystallization by using a 4-channel liquid handling robot (Genomic Solutions). Crystallization plates were stored in the monitoring system (Formulatrix) with automated imaging of the drops. Crystals were obtained after 5 days incubation as shown in Fig. 3.29. This work was done in collaboration with Dr. Vladimir Pena and Dr. Tales Rocha de Moura from the Macromolecular Crystallography research group at the MPI-bpc, Goettingen. Unfortunately, the crystals diffracted badly, and the collected data was not good enough for three-dimensional structure analysis. In order to improve crystals quality, we used several approaches including the limited proteolysis method. The ecGSK was treated with proteases like chymotrypsin, Gluc-C, subtilisin and thermolysin that cut the protein normally at exposed regions such as loops and other highly flexible and disordered regions (Fig. 3.30).

Fig. 3.29. Crystals of ecGSK.
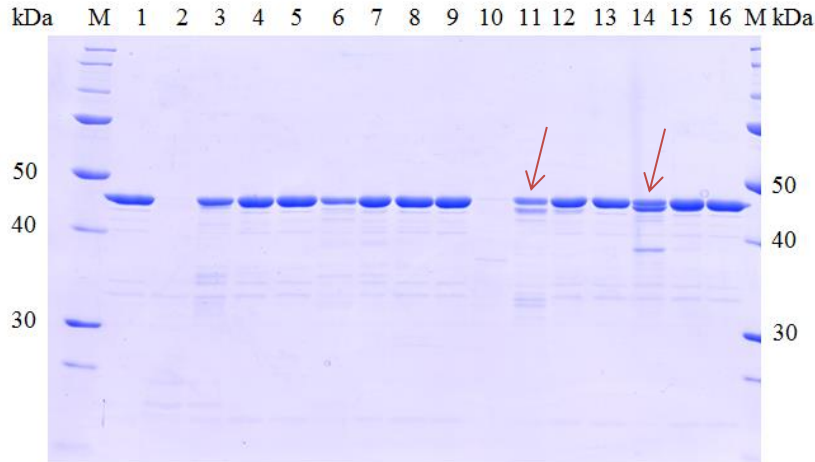
Fig. 3.30. Limited proteolysis of ecGSK. For limited proteolysis, 15 µl of ecGSK (0.7 mg/ml, 48.45 kDa) was taken in 16 microtubes and 4 µl of chymotrypsin, Gluc-C, subtilisin and thermolysin in 1/10, 1/100, and 1/1000 dilutions were added to the respective tubes and incubated at 20 °C for 30 min. About 4 µl of Laemmli sample buffer (6X) was added to each tube and heated at 95 °C for 5 min. The samples were loaded on 12% SDS-PAGE. Lanes: M, marker proteins ladder; 1, as a control and not treated with the protease; 2-4, treated with chymotrypsin; 5, as a control; 6-8, treated with Gluc-C; 9, as a control; 10-12, treated with subtilisin; 13, not treated with the protease; 14-16, treated with thermolysin.

As indicated in wells 11 and 14 in Fig. 3.30, a small fragment of ~2 kDa was removed from the ecGSK full-length protein. Taking this into consideration when analyzing the secondary structure by IUPred (Prediction of intrinsically unstructured proteins) (Appendix Fig. 5), it was found that there is a disordered structure at the C-terminus and a predicted random coil at the N-terminus. Based on these observations, I made three truncated constructs designated as ecGSK-Δ30N (30 amino acids truncated at the N-terminus), ecGSK-Δ21C (21 amino acids truncated at the C-terminus), and ecGSK-Δ30N/Δ21C (30 amino acids at the N-terminus and 21 amino acids at the C-terminus truncated) by PCR amplification. They were recombinantly produced in BL21(DE3)pLysS and purified to homogeneity by affinity chromatography followed by gel filtration chromatography using Superdex 200 column. The three truncated forms of ecGSK were tested for their enzymatic activities, but unfortunately they were found to be catalytically inactive although the recombinant proteins were soluble. Consequently we did not proceed with the crystallization trials of the truncated constructs. In the future, the crystallization of wild-type ecGSK will be optimized by testing different crystallization conditions.

4. Discussion

4.1 Biochemical characterization of human guanylate kinase

Guanylate kinase (GMPK) is an essential enzyme in purine nucleotide synthesis by catalyzing the phosphoryl group transfer from ATP to (d)GMP to form (d)GDP for RNA and DNA metabolism [11]. It has a critical role in the recycling of cGMP by converting GMP into GDP which is the first step in the cGMP cycle (cGMP→GMP→GDP→GTP→cGMP), and thereby regulates the supply of guanine nucleotides to signal transduction pathways [14, 15, 36, 38]. Besides its physiological role, the enzyme is required for the intracellular activation of several antileukemic and antiviral drugs including 6-thioguanine, azathioprine, 6-mercaptopurine, 9-β-D-arabinofuranosylguanine (araG), acyclovir and ganciclovir [11, 17-19, 31, 34, 41]. In addition, guanylate kinases from pathogenic bacteria such as *Staphylococcus aureus* and *Mycobacterium tuberculosis* are potential chemotherapeutic targets [32, 33]. Membrane associated guanylate kinase (MAGUK) homologues are involved in complex multicellular functions like cell junction formation and mitotic spindle orientation, and it was reported that the guanylate kinase domain of MAGUKs, which lack kinase activity, has evolved from the guanylate kinase enzyme [40]. In spite of all these crucial roles, human guanylate kinase (hGMPK) is the least studied guanylate kinase in its family of enzymes with no structural information available. Therefore, I centered my work on the structural and functional analysis of this medically most relevant enzyme.

4.1.1 Kinetic characterization of wild-type hGMPK and site specific mutants

The hGMPK enzyme was produced as a recombinant protein with SUMO and GST tags, and was purified to high purity in enzymatically active form. This successful production is contrary to the published data reported by Brady et al [29], and Ardiani et al [16, 34] who obtained the hGMPK as inactive protein upon expression in *E.coli.* In our case, the fusion tags and the optimized expression and purification conditions might have played an important role in the stability and proper folding of hGMPK when produced in BL21(DE3)pLysS cells [98,118-120]. The hGMPK was found to be a monomer like *Mycobacterium tuberculosis,* yeast and mammalian GMPKs; in contrast, *E.coli* and *Staphylococcus aureus* GMPKs are multimeric [10, 11, 33, 35, 58, 60]. Kinetic parameters determined for wild-type hGMPK were K_m of 25 μM and k_{cat} of 79 s^{-1} when GMP was used as substrate. These kinetic values are similar to those

previously reported for the human erythrocyte enzyme (K_m of 15-24 µM) [121], but differ from the kinetics of the closely related mouse guanylate kinase (mGMPK) with K_m of 59.02 µM and k_{cat} of ~21 s^{-1} [16]. Our data suggest that hGMPK has about 9 times higher catalytic efficiency (k_{cat}/K_m of 316 x 10^4 M^{-1}s^{-1}) than mGMPK (k_{cat}/K_m of 35 x 10^4 M^{-1}s^{-1}), but is slightly less efficient than yeast GMPK which has a K_m of 91 µM and k_{cat} of 394 s^{-1} as reported by Li et al [35]. As the crystal structure of mGMPK (PDB 1LVG) was published in 2002 [11], and since this enzyme shares high sequence identity (88%) with hGMPK, we could directly apply the structural information of mGMPK to hGMPK in the form of a homology model that we used for designing a series of point mutations at multiple sites to analyze their roles in catalysis and domain movements [11]. Besides two carboxylic residues, S37 (which corresponds to S35 in yeast GMPK) plays an important role in substrate discrimination between GMP and AMP by making a single hydrogen bond with the carbonyl oxygen of guanine at position 6 [11, 53].

I introduced four point mutations at residue S37 (S37A, S37C, S37P, and S37Y) by site-directed mutagenesis to understand its role in binding GMP at the active site of hGMPK. Mutation of S37 to alanine, having a nonpolar methyl group with no possibility of forming a hydrogen bond, reduced its catalytic efficiency (2.4 x 10^4 M^{-1}s^{-1}) by a factor of 132 as compared to that of wild-type hGMPK (316 x 10^4 M^{-1}s^{-1}). Ardiani et al [16] suggested that this mutation makes the GMPK molecule resistant to the antileukemic drug metabolite 6-thioGMP. Similar results were obtained for the S37C mutant with k_{cat}/K_m of 3.0 x 10^4 M^{-1}s^{-1}. In this case, the sulfhydryl group of cysteine may favor the sulfur at position 6 of 6-thioGMP instead of oxygen in case of GMP; however, this hypothesis needs to be tested by structural studies. I had chosen the S37P mutation based on multiple sequence alignment of hGMPK with other GMPKs and with guanylate kinase domains (GK^{dom}) of MAGUKs which are catalytically inactive scaffolding proteins that organize protein complexes at cell or synaptic junctions [76]. Like S37 in GMPKs, proline is highly conserved at this position in MAGUKs. The S37P mutation reduced the catalytic efficiency by a factor of 1,053 i.e., k_{cat}/K_m of 0.3 x 10^4 M^{-1}s^{-1}. Johnston et al [40] had demonstrated that S35 in yeast GMPK (which corresponds to S37 in hGMPK) when mutated to proline transformed the guanylate kinase enzyme (GK^{enz}) into a non-enzymatic phosphoprotein binding domain (GK^{dom}) of MAGUKs with the gain of spindle-orienting activity. They further showed by X-ray crystallography, NMR and fluorescence quenching experiments that the loss of enzymatic activity was due to a change in the dynamics of the guanylate kinase. Actually, S37 resides in the

dynamic hinge region that connects the NMP-binding region (region that binds GMP) and the CORE region. Its mutation to proline, which is the least flexible amino acid in terms of sterically allowed conformations, prevented (or highly reduced) the GMP-induced closing movement of the NMP-binding domain of yeast GMPK and thus hindered an essential step in catalysis. They also found that the proline mutant could still bind ATP and GMP nucleotides like we observed in the form of highly reduced activity [77].

Like proline at position 37, there is another conserved proline residue at position 42 in MAGUKs which corresponds to N42 in hGMPK. We were curious to find out whether proline at this position would also drastically change the enzyme kinetics. Unexpectedly, our data showed that the N42P mutation increased the K_m by a factor of only two, and the k_{cat} was rather slightly increased. N42 is neither reported to interact with substrates nor is it located in the hinge region. It belongs to the NMP-binding region next to R41 that however interacts with the phosphate of GMP, and consequently the effect on catalysis was small. The most drastic effect was observed when we substituted tyrosine for S37 because it made hGMPK completely inactive. This was contrary to the results published by Ardiani et al [16], where this mutation resulted in ~7-fold decrease in the catalytic efficiency of mGMPK. The complete loss of enzymatic activity may be due to the fact that tyrosine has a bulky phenol side chain that may not be favorable owing to adjacent T83 and E72 residues. The T83 and E72 residues form bidentate interactions with the carbonyl at position 6, and with two hydrogen bond donors at positions N1 and the amine group at position 2 of guanine, respectively [11]. As the S37Y mutant did not aggregate and was a soluble protein, implying that it was properly folded, the mutation may have only disturbed the active conformation of hGMPK. High resolution NMR or X-ray crystallography structures of S37 variants in the presence of GMP or its analogs will provide detailed explanation for the effects of these substitutions on hGMPK kinetics. Comparing the four S37 mutants, we can rank the guanylate kinase activity as: wt-hGMPK > S37C > S37A > S37P. T83 in hGMPK corresponds to S80 in yeast GMPK, and when we introduced the T83S mutation, it showed indeed an improved k_{cat} like yeast GMPK as reported by Li et al [35]. Nevertheless, the overall catalytic efficiency (132 x 10^4 $M^{-1}s^{-1}$) was still lower than that of wt-hGMPK. The T83A mutant had decreased activity because the methyl group of alanine might not form a bidentate interaction like threonine as Sekulic et al had reported [11]. The Y81 is conserved among all GMPKs and in MAGUKs (e.g. hDlg). Both in mGMPK as reported by Sekulic et al [11] and in

yeast GMPK as shown by Stehle and Schulz [53], Y81 interacts with the α-phosphate of GMP via a single hydrogen bond. When we substituted Y81 with phenylalanine to disrupt this interaction, the K_m was increased by a factor of 12 and k_{cat} was reduced fourfold. Zhang et al [61] had demonstrated by NMR two-dimensional spectra analysis that the phenylalanine mutant of yeast GMPK was properly folded, and that its conformation was highly similar to that of wild-type. They attributed the changes in kinetic parameters to the removal of the hydrogen bond between the hydroxyl group of tyrosine and the α-phosphate of GMP which has a role in stabilizing binary and ternary complexes.

In conclusion, we have shown that a single hydrogen bond between S37 and the carbonyl oxygen of guanine in GMP is critical for GMP binding and for catalysis. Besides its catalytic role, S37 is required for the dynamics of the hinge region that facilitates the movement of the NMP-binding region upon binding of substrates. Its substitution with proline or tyrosine drastically affects hGMPK activity. Similarly, the bidentate interaction of T83 with the carbonyl oxygen of guanine in GMP is required for catalysis. Y81 interacts with the phosphate of GMP and has a role in binary complex stabilization.

4.1.2 Substrate-induced conformational changes in hGMPK studied by small angle X-ray scattering

NMP kinases (ATP:NMP phosphoryltransferases) undergo large domain movements upon binding of both substrates [5], as integral part of their catalysis [122]. These changes can be determined by comparing the enzyme structures in different conformational states [9]. We determined hGMPK in at least four forms designated as open form (hGMPK$_{apo}$), two partially closed forms (hGMPK$_{PC}$), and a completely closed form (hGMPK$_{closed}$) by small angle X-ray scattering (SAXS). Up until now, no GMPK structure has been analyzed in all of these four conformational states. Rg (radius of gyration) values suggested that the hGMPK$_{apo}$ molecule was bigger (~21 Å) in size than the hGMPK$_{closed}$ (~19 Å) and the two partially closed forms (19.7 Å & 20.2 Å). It was observed that the product formation in the presence of ATP and GMP caused a small increase in the hGMPK$_{RO}$ (reopened form) size as compared to the closed forms. Based on Guinier and pair distance distribution function P(r) analysis, we concluded that the closed forms of hGMPK were more compact than all other forms in solution:

$$hGMPK_{closed} > hGMPK_{PC} \geq hGMPK_{RO} > hGMPKapo$$

We found that the three-dimensional surface reconstruction of hGMPK$_{apo}$ was a globular domain with two projections originating from opposite sides of the molecule, termed large P1 and small P2 (Fig. 3.4a & 3.6a). These two projections moved with respect to the center of hGMPK molecule and became less conspicuous in the closed and partially reopened forms upon binding of nucleotides. The superposition of closed and partially reopened forms of hGMPK onto its open form indicated substrate-induced conformational changes due to movements of P1 and P2 domains (Fig. 3.4b-d). Blaszczyk et al [10] observed significant movement of the NMP-binding domain and LID domain when they compared the crystal structures of yeast GMPK in unliganded and GMP-bound forms. Similarly, Johnston et al [40] demonstrated by NMR and fluorescence quenching experiments that yeast GMPK undergoes GMP-induced conformational changes. Sekulic et al [11] determined the crystal structure of mouse GMPK (mGMPK) in closed form (GMP and ADP bound), and overlaid it on the known structures of yeast GMPK in open (no bound substrate) and partially closed forms (GMP bound). They observed substrate-induced domain movements. Choi and Zocchi [13] used an allosteric spring probe to demonstrate GMP-induced changes in GMPK from *Mycobacterium tuberculosis*. The overall fold of hGMPK is very similar to other members of the NMP-kinases family, in particular to mouse and yeast GMPKs [10, 11, 35]. As there is no high-resolution structure available for hGMPK, and crystallization of this enzyme failed [11], we decided to determine its low resolution SAXS structures. Moreover, an *in silico* homology model was constructed for hGMPK, based on the highly identical (88%) mGMPK (PDB 1LVG) structure used as a template, to highlight its structural features. It was predicted that hGMPK has three dynamic structural regions termed NMP-binding region (NMP-BR), CORE, and LID regions interconnected by four hinges.

Our next step was to identify P1 and P2 projections (lobes) that we observed in the SAXS models of hGMPK. For that purpose, we superposed the hGMPK SAXS structures on the crystal structure of mGMPK$_{closed}$. Interestingly, it was found that the P1 and P2 regions in the hGMPK aligned with the NMP-BR and LID regions of mGMPK, respectively (Fig. 3.6a-d). Similar results were obtained when we compared the SAXS structures of hGMPK with its homology models. Based on these comparative analyses, we concluded that the apo-form has much more open conformation than all other partially closed, completely closed and reopened forms [10, 11]. Binding of GMP to hGMPK causes a significant movement of the NMP-BR towards the LID region with comparatively smaller movement of the LID region in the same direction. Overall,

GMP binding brings the two regions closer to each other making the enzyme molecule relatively compact in the partially closed conformation [10]. The simultaneous binding of two nucleotides, such as GMP and an ATP-analog, brings the two regions even closer to each other, making the molecule adopt its most compact, fully closed conformation [11]. The open form is highly flexible whereas the closed form is the most rigid one [10, 11]. Nevertheless, besides the good fit, the closed forms of human and mouse GMPKs differ at certain positions which could partly be due to differences in their amino acid sequences, but also due to small structural variations between in-solution and crystal structures.

In the course of catalysis, kinases undergo numerous rounds of open-to-closed conformational transitions [41], and we observed a mixed-state population of enzyme/substrate/product complexes designated as reopened form (hGMPK$_{RO}$) when hGMPK was provided with its two natural substrates, ATP and GMP. The size of hGMPK$_{RO}$ is slightly bigger than the closed forms. This leads to the outward movement of P1 (NMP-BR) in the hGMPK structure as supported by the P(r) profile. For detailed analysis, time-resolved measurements would be performed in future studies.

In summary, in this work, conformational analysis of hGMPK in its open, closed, partially closed and reopened forms have been presented. Comparing the SAXS structures of hGMPK in these different states revealed large conformational changes that occur during catalysis. The open-to-closed conformational transition of the hGMPK molecule induced by binding of ligands supports the model of induced fit mechanism. The best fitting of the two closed forms from human and mouse GMPKs indicates the highly conserved three-dimensional fold in GMPKs.

4.1.3 Structural characterization of hGMPK by NMR

For optimizing the higher yield of isotope-labeled (^{15}N, ^{15}N/^{13}C) hGMPK, the labeled enzyme was recombinantly produced in *E.coli* which is a convenient, rapid and economical host organism for producing isotopically labeled proteins [104]. Therefore, there is a huge demand to develop a variety of strategies for achieving high-level protein expression in *E.coli*. One of the basic approaches which we used was the pET expression system (Novagen) and the *E.coli* strain BL21(DE3)pLysS which contains the pLysS plasmid for tight regulation of the T7 promoter. The pET-14bSUMOΔThr vector, used for cloning of hGMPK, had the SUMO (Small Ubiquitin-like Modifier) tag for enhanced expression, solubility and stability of hGMPK [98, 118]. The

hGMPK production was high (10-15 mg/l culture) when expressed in rich media such as TB (terrific broth) and autoinducing media (see section 2.1.9.1). Nevertheless, when expression was carried out in M9 minimal growth media for ^{13}C and ^{15}N labeling, a very low yield was observed (~1 mg/l culture). Structure elucidation by NMR requires a large quantity (0.5-1.5 mM in 260-550 µl volume) and high quality (>95% purity) of isotopically (^{15}N, ^{15}N/^{13}C) labeled recombinant protein, the production of which is costly and time-consuming. Thus, it was necessary to set up an optimized and simple protocol for higher production of single- and double-labeled hGMPK. I extracted information from the available protocols particularly those reported by Cai et al [105], Marley et al [103], and Sivashanmugam et al [102], and made several modifications to the existing protocols. For instance, the commercially available "basal vitamins eagle media" were omitted from the minimal growth media recipe as reported by Marley et al [103], and different expression conditions were tested, like varied induction time and temperature, while monitoring their effects on hGMPK yield. The final optimized protocol was mainly based on (1) double-colony selection, and (2) high cell-density method. The double-colony selection procedure was used to optimize high-level protein production [102, 103]. The method was based on selecting several colonies from a freshly streaked agar plate containing the appropriate antiobiotics. They were expressed in minimal growth media, and the expression level was tested on SDS-PAGE. The colony which gave the highest protein production was selected and re-streaked on an agar plate for the second round of selection. The high-level expressing colony was then used for hGMPK expression by a modified high cell density method. In this method, the hGMPK-containing cells were grown in a rich medium like standard LB-broth up to OD_{600} of ~0.7 (mid log phase). The medium was then switched to fourfold smaller volume of the labeled minimal growth medium for the expression under optimum conditions. The cells were harvested and used for protein purification by affinity chromatography using Protino Ni-IDA resin; the protein was further purified by gel filtration chromatography using Superdex 75 column. Our protocol, overall improved the single- and double-labeled hGMPK production by a factor of 8, and yielded about 8 mg/l culture of highly pure isotope-labeled enzyme.

As preliminary NMR experiments, the 1 mM ^{15}N-labeled hGMPK$_{apo}$ (open form) was titrated against increasing concentrations of GMP substrate to generate the enzyme species designated as ^{15}N-labeled hGMPK$_{GMP}$ (partially closed form). The ^{15}N-^{1}H HSQC spectra obtained indicated significant changes in the positions of amino acid residues upon binding of GMP. These results

were complementary to our findings by SAXS analysis showing that hGMPK undergoes major conformational changes upon binding of its substrates. About 177 residues out of total 197 amino acids of hGMPK have been identified from the analysis of the NMR 3D spectra. This is a collaborative work which is still in progress to solve the three-dimensional structure of hGMPK. It will enable us to understand the catalytic mechanism of hGMPK and to rationally design its catalytically more efficient mutants against the antileukemic 6-thioGMP. The engineered mutants may be tested for suicide gene therapy of acute lymphoblastic leukemic cells either alone or as a fusion partner with *E.coli* guanosine-inosine kinase (ecGSK) which phosphorylates 6-thioguanosine to 6-thioGMP.

To conclude, we optimized the production of isotopically labeled hGMPK for NMR structural elucidation. The ^{15}N-^{1}H HSQC spectra analysis indicated the GMP-induced conformational changes in hGMPK which were complementary to our findings by SAXS studies.

4.1.4 Enhanced cytotoxicity of the antileukemic drug 6-thioguanine by expressing hGMPK in HEK293 cells

The purine analog 6-thioguanine (6-TG) has been extensively used for the treatment of acute lymphoblastic leukemia [19]. It is readily transported into cells [123], and must be enzymatically converted to the nucleoside monophosphate form for further phosphorylations by nucleotide kinases [106]. We speculated that hGMPK, which catalyzes the second phosphorylation step in the conversion of 6-thioguanine to 6-thioGTP/6-thiodGTP, may be the bottleneck enzyme in the metabolic activation of this drug. Once incorporated into RNA and DNA, 6-thio-nucleotides exert their cytotoxic effects by triggering the post-replicative DNA mismatch repair system, mitochondrial DNA dysfunction, and reactive oxygen species formation [17-19].

Previously, Ardiani et al [16] reported that substitutions at S37 in mGMPK conferred genetic complementation to a conditional GMPK-deficient *E.coli* strain whereas the wild-type mGMPK-expressing cells were sensitive to the clinically used 6-TG prodrug. These authors suggested that drug resistance to 6-TG may be due to point mutations in GMPK. Our aim was to investigate whether hGMPK would enhance the cytotoxicity of 6-TG when tested in mammalian HEK293 cells. This will provide a basis for using the engineered hGMPK as a potential suicide gene for efficiently killing of cancer cells.

Yuan et al [31] reported that long-term use of 6-TG like many other chemotherapeutic agents is associated with several side effects including carcinogenicity. DNA containing 6-thioguanine once exposed to UVA light oxidizes 6-thioguanine to guanine-S^6-sulfonic acid (^{SO3H}G). The ^{SO3H}G blocks DNA polymerases and elicits ambiguous coding properties making the DNA highly mutagenic which is associated with increased skin cancer risks [129]. We hypothesized that this problem occurs because endogenous hGMPK inefficiently phosphorylates 6-thioGMP, and consequently high drug doses are required for prolonged periods of time to treat acute lymphoblastic leukemia. It has been demonstrated that the catalytic efficiency of GMPK for 6-thioGMP is about 8,000 times lower than its catalytic efficiency for GMP as reported by Ardiani et al [16]. Thus, overexpression of hGMPK might overcome the bottleneck of metabolic phosphorylation by enhancing the activation of 6-TG at even low dose levels. For that purpose, I made two fusion constructs of hGMPK, one with N-terminal EGFP and the other with C-terminal EGFP for fluorescence visualization in cells. Both fusion constructs along with a control EGFP were used to transfect HEK293 stable-cell lines. It was observed that the expression of hGMPK with C-terminal EGFP was very low as compared to the N-terminal EGFP construct as examined microscopically and by immunoblot analysis. We concluded that EGFP fused to the C-terminus may have impaired the proper folding of hGMPK [125-128]. In contrast, the N-terminal EGFP did not adversely affect the activity of hGMPK as I verified by the NADH-dependent enzyme-coupled assay using the bacterially produced EGFP-hGMPK fusion protein. Therefore, the EGFP-hGMPK transfected HEK293 cells were tested for their sensitivity to varied concentrations of 6-TG (0.001-1000 μM) incubated for 48 h at 37 ^{o}C. Western blot analysis revealed similar expression levels for EGFP and EGFP-hGMPK. Cell viability of stable cell lines was determined by using the MTT assay [124]. The EC$_{50}$ values derived from the dose-response curves indicated that EGFP-hGMPK-expressing stable cell lines were more sensitive (about six times lower EC$_{50}$) to 6-TG than the EGFP-expressing stable cell lines serving as control. In conclusion, the overexpression of hGMPK increased the sensitivity of HEK293 cells to 6-TG due to enhanced activation of the prodrug.

4.1.5 Electrochemical detection of guanosine monophosphate with a quantum dot-based biosensor modified with hGMPK

The aim of this study was to develop a novel approach for the detection of nucleotide kinase-catalyzed reactions based on the electrochemical sensing of NADH in a multi-step coupled-enzyme assay, and in addition, to demonstrate the proof of concept of a biosensor for hGMPK immobilized on CdS/ZnS quantum-dots (QDs) modified gold electrode.

Schubert et al [88] reported that using the CdSe/ZnS QDs electrode, NADH addition to a buffer changes the photocurrent under illumination at a constant electrode potential. This indicates that electrons are transferred from NADH to the excited quantum dots under illumination. The NADH oxidation cannot simply be enhanced by the applied bias potential, rather it was determined by the catalytic properties of the semiconductor nanoparticles carrying the enzyme [88]. Thus, such a QDs-modified electrode system can be used for the electrochemical detection of NADH at a constant bias potential. It will provide a basis to construct biosensors for those enzymes which catalyze NADH/NAD$^+$-dependent redox reactions directly or indirectly in a coupled assay. Using this principle, we decided to apply the NADH-dependent enzyme-coupled assay to the CdS/ZnS QDs-modified gold electrode system for monitoring the hGMPK-catalyzed reaction (see section 2.2.7) [99]. In the first step of the assay, hGMPK catalyzes the phosphorylation of GMP to GDP using ATP as a phosphoryl group donor [11, 29]. The formation of NDPs (ADP and GDP) is coupled to two additional reactions catalyzed by pyruvate kinase (PK) and lactate dehydrogenase (LDH), respectively. It was predetermined that the amount of NADH oxidized in the enzymatic reactions linearly depends on the concentration of GMP. This change in NADH concentration was sensed by the CdS/ZnS QDs-modified gold electrode under illumination for 5-10 sec. The electrode provided a stable light-switchable layer.

First, a control experiment, which is basically the last reaction of our coupled assay as mentioned above, was performed. It is a redox reaction between pyruvate and NADH catalyzed by LDH to form NAD$^+$ and lactate. All measurements were performed in 100 mM HEPES buffer pH 7.5 containing 100 mM KCl and 20 mM MgCl$_2$ at a constant bias potential of +50 mV at room temperature. The concentration of pyruvate was varied from 50 μM to 1600 μM keeping other constituents of the reaction mixture constant i.e., 1.2 mM NADH and 15 units of LDH. The current was recorded in the presence and absence of pyruvate upon illumination of the CdS/ZnS QDs electrode. A linear relationship was found between the change in photocurrent and pyruvate

concentration up to 1000 μM pyruvate, above which saturation started. Similar results were reported by Schubert et al [88] for the detection of NADH produced by the reaction between glucose and NAD^+ catalyzed by glucose dehydrogenase. In that case, the formation of NADH was sensed at a constant bias potential of +50 mV (versus Ag/AgCl, 1 M KCl) using CdSe/ZnS QDs-modified gold electrode [88], and it was demonstrated that the change in current was dependent on the concentration of glucose. Khalid et al [87] determined a similar dose-response curve for the detection of 4AP (4-aminophenol) at constant bias potential of +200 mV using a CdS QDs electrode.

Under our optimized assay conditions, the addition of GMP substrate to the reaction buffer in the absence of hGMPK showed no signal change in the potential range from -100 to +200 mV (versus Ag/AgCl, 3 M KCl) (data not shown). This means that unlike NADH, GMP cannot be oxidized at the QDs-modified gold electrode. However, in the presence of 18 nM hGMPK and all other constituents of the assay (as stated above), photocurrent changes were detected in the GMP-concentration dependent way. GMP was detected in the concentration range from 50 μM to 1200 μM, and saturation started above this value. Mora et al [130] developed an analytical method for the determination of AMP (adenosine 5′-monophosphate) coupled to 6-thioguanosine and a glassy carbon electrode with gold nanoparticles. In comparison, our approach is different because it is an enzyme-based detection of nucleotides which could be applied to any type of kinases in general, and can be used for studying their kinetics.

To demonstrate a proof of sensor, we immobilized the assay enzymes (hGMPK, PK and LDH) on the CdS/ZnS QDs gold electrode using layer-by-layer assembly of polyelectrolytes, PSS (poly (styrene sulfonate)) and PAH (poly (allylamine hydrochloride)) [87, 89, 114, 115]. The photocurrent measurements were carried out at varied concentrations of GMP. The response curve obtained for GMP-to-GDP conversion was similar to that when hGMPK was free in solution, i.e. not immobilized on the QDs surface. The maximum current (37 nA) detected in case of immobilized hGMPK was slightly lower than that for free hGMPK (45 nA). These little variations may be due to electrostatic interactions of the nucleotide substrates with the negatively (PSS) and positively (PAH) charged polyelectrolytes. The aim of this study was to demonstrate the proof of concept of a biosensor for detecting hGMPK activity; building of a practical sensor will further need the optimization of reaction conditions, most importantly extending the half-life

of enzymes in use. Potentially, this assay may be applied to pathologically important kinases like creatine kinase in serum which is of considerable diagnostic significance [319].

In summary, a new technique for the detection of nucleotide kinase-catalyzed reactions in a coupled assay has been developed. This system demonstrates the functioning of a light-controlled bioelectrochemical sensor for hGMPK.

4.1.6 A novel spectrophotometric and fluorometric enzyme-coupled assay for hGMPK

The conventional NADH-dependent spectroscopic assay has been used for the activity measurements of nucleoside and nucleotide kinases [84, 99]. In this assay, absorbance of NADH ($\varepsilon = 6.22$ mM^{-1}cm^{-1} at 340 nm) in the reaction mixture is monitored at 340 nm, and it gradually drops down to zero upon oxidation of NADH to NAD$^+$. Nevertheless, using this assay we encountered a serious problem when we were trying to use 6-thioGMP and 6-thioguanosine nucleotides in the assay conditions. Like NADH, both 6-thioGMP and 6-thioguanosine have absorption maxima at ~340 nm [18, 134]. To overcome this wavelength problem, we decided to develop an assay that can be used in absorbance mode at a wavelength without background absorbance, but also may be applicable in the fluorescence mode.

Thus, we set up a four-step Amplex Red-based coupled assay in which the first reaction is catalyzed by the nucleotide kinase of interest producing nucleoside diphosphates (NDPs). The NDPs are coupled to phosphoenolpyruvate in the next step catalyzed by pyruvate kinase to form nucleoside triphosphates (NTPs) and pyruvate. In the third reaction, pyruvate is coupled to inorganic phosphate catalyzed by recombinant pyruvate oxidase (from *Lactobacillus plantarum,* ~69 kDa) producing hydrogen peroxide (H_2O_2) and acetyl phosphate. In the last step, horseradish peroxidase uses H_2O_2 to oxidize the nonfluorescent Amplex Red to resorufin which has excellent absorbance as well as fluorescence properties with $\varepsilon = 5.4 \times 10^4$ M^{-1}cm^{-1} at 570 nm [116, 135]. We optimized the assay in absorbance mode at 570 nm. About 5 nM of hGMPK was used with varied concentrations of GMP (0-500 µM). Steady-state turnover rates (k_{obs}) were calculated, and the Michaelis-Menten plot was constructed by Gnuplot 5.0 software. Kinetic parameters were deduced from the plot with K_m of 30 ± 5 µM and k_{cat} of 80 ± 4 s^{-1}. To validate our assay, the kinetic parameters for hGMPK were also determined by the conventional NADH-dependent spectroscopic assay. Kinetic values obtained from both assays were in very good agreement. As

resorufin is a highly fluorescent compound, the assay can be used in fluorescence mode as well [116].

In summary, we established the Amplex Red-based absorption spectrophotometric and fluorometric enzyme-coupled assay as an alternative to the conventional NADH-dependent spectroscopic assay. It overcomes the overlapping wavelength problem associated with strong absorption of 6-thioguanine nucleotides at 340 nm, and it has the advantage to be applicable both in absorbance and fluorescence modes.

4.1.7 hGMPK-catalyzed reactions in polyelectrolyte containers of different shapes and sizes

We synthesized calcium carbonate ($CaCO_3$) microparticles of different shapes and sizes, and loaded them with hGMPK to test their loading capacity and enzymatic activities.

$CaCO_3$ microparticles were synthesized by mixing $CaCl_2$ and Na_2CO_3 solutions in the presence of ethylene glycol [131]. It was found that by increasing the ethylene glycol (EG) content from 66.6% up to 80%, an enhanced density was provided in the reaction mixture, which reduced the stability of precipitated $CaCO_3$ particles. Under these conditions, ellipsoidal particles with varying axis ratios were produced. Some ellipsoidal microparticles have round edges and resemble rod-shaped bacteria such as *Bacillus*; others have sharp ends and bear resemblance to *Euglena*. Our findings revealed that the solubility of salts is one of the key parameters for the shaping process of anisotropic $CaCO_3$ particles. It was found that high concentrations and equal ratios of carbonate and calcium ions causes anisotropic rhomboidal geometry, whereas high carbonate and lower calcium ions resulted in anisotropic elliptical particles. However, low concentration of carbonate and high concentration of calcium ions produced isotropic spherical particles.

Microparticles were loaded with hGMPK by incubation in 1 mg/ml stock solution on a shaker for 15 min. After one washing step with water, microparticles were counted in a hemocytometer, and the protein amount was determined by the Bradford dye-binding assay as reported by De Temmerman et al [131]. The particles were coated with consecutive layers of positively charged PAH and negatively charged PSS, and the enzymatic activity was determined by the NADH-dependent spectroscopic assay [99, 133]. Comparison of the specific activities of the loaded microparticles revealed that ellipsoidal particles had highest protein-loading capacities as

compared to microparticles of all other shapes used in this study. We are on the way to further characterize these particles of different shapes for their loading capacity with *Drosophila melanogaster* deoxyribonucleoside kinase (Dm-dNK) and hGMPK fused to EGFP. Moreover, we will use biocompatible polyelectrolytes like poly L-arginine and dextran sulfate for biological applications [132].

In conclusion, ellipsoidal particles with loaded hGMPK demonstrated higher specific activities, after coating with polyelectrolytes, as compared to microparticles of all other shapes including spherical, rhomboidal, star and cube.

4.2 Biochemical characterization of human mitochondrial thymidine kinase

Human mitochondrial thymidine kinase (hTK2), which phosphorylates dThd, dCyd and dUrd to their corresponding monophosphates, is constitutively expressed in all tissues and is localized in the mitochondrial matrix providing dNTPs (deoxyribonucleoside triphosphates) for the replication and maintenance of mitochondrial DNA [151, 152, 157, 159, 199, 222]. Like mouse TK2, hTK2 has an N-terminal mitochondrial targeting signal for its translocation into mitochondria [138, 140, 150, 151]. Thymidine kinase activity has been detected in mitochondria [155-159]. Nevertheless, cytosolic forms of hTK2 were also reported [152, 153]. The cytosolic forms of hTK2 will affect the dNTPs pool in cytoplasm. Thus, hTK2 may play an important role not only in the mitochondrial but also in the nuclear DNA replication and repair system [139]. Wang and Eriksson [140] demonstrated by *in vitro* translation and translocation experiments that a 40 residues long N-terminal signal sequence directs the transport of mouse TK2 into mitochondria. In contrast, Johansson et al [154], when expressing TK2-GFP in CHO (Chinese hamster ovary) cell lines, observed that TK2-GFP fusion was not imported into mitochondria but remained in the cytoplasm. However, unlike TK2, the dGK-GFP (dGK, deoxyguanosine kinase) fusion protein was transported into mitochondria [154]. Besides that, these authors did not specify the mitochondrial import signal in hTK2 [154]. In this entire scenario, we were curious to perform similar experiments with the long isoform (265 amino acids) [138] of hTK2 as a fusion with GFP to determine its subcellular localization in HEK293 cells. There exist several isoforms for hTK2 that may have originated from alternative splicing of a single gene [138, 139]. We selected the long isoform containing the putative 38-amino acid N-terminal mitochondrial targeting signal as predicted by the MITOPROT tool (ExPASy server). The long isoform

(UniProt O00142-1), as previously reported by Wang et al [138], has 41 N-terminal amino acids different from the 10 N-terminal residues of the short isoform (UniProt O00142-2) but sharing the remaining 224 amino acids. When we overexpressed the long isoform in HEK293 cells, it indeed showed mitochondrial localization as a TK2-EGFP protein when examined by confocal microscopy. The pattern of mitochondrial appearance as a network of ovoidal structures was similar to the one observed for dGK-GFP [154]. Unlike the full-length hTK2, the N-terminal truncated forms were uniformly distributed in the cytoplasm.

When hTK2 was produced recombinantly in BL21(DE3)pLysS cells as a fusion with an N-terminal histidine tag, it extensively aggregated as inclusion bodies. Similar results were previously reported by Barroso et al [139] when they expressed hTK2 as a fusion with N-terminal MBP (maltose binding protein) tag. Aggregation in inclusion bodies can be reversible and the protein may be recovered in native folded form [315, 316]. Therefore, I used several detergents including Triton X-100, urea, CHAPS, sarkosyl, SDS, and L-arginine to solubilize and subsequently refold the recombinant hTK2 by dialyzing against a native buffer [160-164, 170, 171]. Nevertheless, all these efforts to regain the activity of denatured hTK2 were ineffective. Similarly, I used different *E.coli* expression strains, most importantly BL21(DE3)pLysS and C41(DE3) that are commonly recommended for the expression of proteins either toxic to cells or membrane-bound [165]. In addition, expression of hTK2 from a variety of vectors containing different tags and/or fusion partners was tested. It was accompanied by varying experimental conditions such as reducing the rate of protein synthesis by lowering the growth temperature or lowering inducer (IPTG) concentration. Moreover, the growth medium was changed by the addition of 1% glucose to repress protein induction, or ethanol was added to upregulate the synthesis of heat shock proteins [166-168]. Finally, I solved the aggregation problem by expressing hTK2 as a fusion with His_{14}-MBP-SUMObr tag (~60 kDa) in BL21(DE3)pLysS cells using LB (Luria-Bertani) medium while inducing with low IPTG concentration (~0.4 mM) at low temperature (18 ^{o}C) [82, 83]. But, we observed that along with soluble hTK2, *E.coli* chaperonins (e.g., GroEL) of molecular weight ~60 kDa were also coexpressed and copurified when tested by SDS-PAGE, and identified by LC-MS/MS [173]. These heat shock proteins support the proper folding of proteins and are induced under stress conditions [169, 172]. They may prevent self-aggregation of incompletely folded proteins [139].

We managed to get rid of GroEL after washing the sample with ATP/Mg^{2+} during the affinity purification step when hTK2 fusion protein was bound to the Protino Ni-IDA resin [174].

As hTK2 is the only human deoxyribonucleoside kinase (dNK) in its family of dNKs whose three-dimensional structure is missing, we decided to determine its crystal structure. For that purpose, I truncated the full-length hTK2 (265 aa) to make 6 constructs of variable lengths by deleting residues from both its N- and C-termini based on IUPred and secondary structure predictions [176, 178]. Graslund et al [175] reported that the multi-construct approach is a successful strategy to produce soluble recombinant proteins in *E. coli* and to enhance the chances of success for obtaining well-diffracting protein crystals. Similarly, Savitsky et al [177] demonstrated that more than 80% of the crystal structures which he obtained were derived from truncated proteins. For hTK2, two out of six truncated forms were catalytically active whereas the other four were inactive. The double-truncated hTK2-Δ44N/Δ8C (N-terminal 44 aa and C-terminal 8 aa truncated) was a dimeric protein and had nearly 3-fold higher catalytic efficiency than the single-truncated hTK2-Δ44N (N-terminal 44 aa truncated). Only the double-truncated hTK2-Δ44N/Δ8C was successfully crystallized but unfortunately the crystals badly diffracted. As described above, hTK2 is prone to aggregation when recombinantly produced in *E.coli* which points to lack of proper folding that may give rise to poor sample quality. Based on these observations, we hypothesized that hTK2 requires eukaryotic chaperones and potentially post-translational modifications for proper folding and structural stability. Thus, its expression in eukaryotic cells such as insect cells might yield samples amenable to crystallization. Sun et al [179] demonstrated that providing oxidizing or reducing agents to TK2 caused its modifications that led to large variations of the activity and stability *in vitro*. Similarly, Hazra et al [180] reported that post-translational phosphorylation of S74 in human deoxycytidine kinase modulates its kinetic properties.

Directed evolution is a powerful tool to generate variants of desired properties and to study the relationship between sequence, structure and function of a protein [91, 181]. Therefore, we employed the method of error-prone PCR based on serial dilution steps of the hTK2 template DNA. It enabled us to control mutagenesis efficiency while simultaneously avoiding the PCR saturation problem [91]. Our aim was to use this technique to generate hTK2 variants for better activity against antiviral and anticancer nucleoside analogs [183]. The double-truncated hTK2-Δ44N/Δ8C construct, which was catalytically more active than other forms, was selected as a

template for random mutagenesis and for establishing a screening procedure based on TK complementation of the TK-deficient *E.coli* strain KY895 [92-94, 182]. Gemcitabine (dFdC, 2´,2´-difluorodeoxycytidine) which is an anticancer deoxyribonucleoside analogue, and the antiviral AZT (Zidovudine, 3´-azido-2´,3´-dideoxythymidine) are poor substrates of hTK2 [184]. Our low-frequency mutagenesis and subsequent screening method revealed two mutants (M5 and M17) that increased the sensitivity of transformed KY895 to gemcitabine by 25 times and fourteen mutants by 10-fold. In addition, five mutants enhanced the sensitivity of KY895 to AZT by a factor of 3. However, the positions of these mutations in the hTK2 primary structure and their kinetic characterization need to be determined. Moreover, these mutants could potentially be tested as suicide genes in cancer cell lines.

In summary, we demonstrated in cell culture experiments that the full-length hTK2 was translocated into mitochondria guided by its N-terminal mitochondrial targeting signal whereas the N-terminally truncated forms were localized in the cytoplasm. As hTK2 forms inclusion bodies upon expression in *E.coli*, therefore to solve this aggregation problem, we optimized an expression and purification protocol for the recombinant production of hTK2 in catalytically active and soluble form. The protocol is mainly based on the expression of hTK2 with an N-terminal His_{14}-MBP-SUMObr tag under optimum conditions. We also optimized conditions for the crystallization of hTK2. Through directed evolution using error-prone PCR followed by a two step screening procedure, we generated hTK2 mutants with several fold higher efficiency for phosphorylating nucleoside analog prodrugs like antiviral AZT and anticancer gemcitabine.

4.3 Biochemical characterization of *E.coli* guanosine-inosine kinase

The *E.coli* guanosine-inosine kinase (ecGSK, gsk, EC 2.7.1.73) has both guanosine and inosine kinase activities [185, 186]. The ecGSK gene was mapped on the *E.coli* chromosome by Jochimsen et al [185], and the guanosine and inosine kinase activities were identified. Harlow et al [149] and Mori et al [148] reported the cloning and characterization of the purified gene product. Its 434 amino acid primary structure is completely different from human deoxyguanosine kinase [187, 188]. We used the nucleotide sequence reported by Mori et al [148], and amplified it from the genomic DNA of *E.coli* XL1-Blue cells [95]. The enzyme was recombinantly produced in appreciable amount, and was purified to homogeneity by chromatographic techniques including affinity chromatography using Protino Ni-IDA resin,

anion exchange chromatography using DEAE Sepharose FF column, and by gel filtration chromatography using Superdex 200 column. We found that ecGSK, besides the physiological substrates Guo and Ino, also phosphorylates 6-thioguanosine which is the nucleoside form of the clinically used antileukemic drug 6-thioguanine (6-TG). This property is unique in the sense that commonly used nucleoside kinases (e.g., Dm-dNK and dCK) do not phosphorylate this nucleoside analog [184, 189, 190]. Therefore, our finding may open new possibilities of using ecGSK as a suicide gene in combination with 6-thioguanosine which is not accepted by endogenous human nucleoside kinases as a substrate.

The three-dimensional structure of ecGSK is unknown, so its structure elucidation will provide a more detailed understanding of its function. We tried its crystallization by the vapor diffusion method [191, 192]. It crystallized in 100 mM Bis-tris buffer at pH 6.2, but unfortunately the diffraction data obtained was not good enough for solving its structure. In order to enhance the chances of getting well diffracting crystals, I truncated the disordered terminal regions from ecGSK, as predicted by IUPred, secondary structure prediction analysis and limited proteolysis. I recombinantly produced its N-and C-terminal truncated versions but unexpectedly they were catalytically inactive [175, 177]. This indicated that both termini of the ecGSK are either part of the active site or required for stabilization of the enzyme´s active conformation. Overall, the initial crystallization trials were satisfactory, but further efforts are needed to optimize crystallization conditions for getting better quality crystals for solving the ecGSK structure [191].

In conclusion, we recombinantly produced ecGSK which manifested a unique property of phosphorylating the nucleoside form of the antileukemic drug 6-thioguanine. In order to solve its three-dimensional structure, we optimized conditions for its crystallization.

Appendix

Human guanylate kinase (hGMPK, GUK1, GMK, ATP:GMP phosphotransferase, EC 2.7.4.8), UniProt identifier Q16774-1, 197 amino acids, and 591 bp ORF

```
  1
ATG TCG GGC CCC AGG CCT GTG GTG CTG AGC GGG CCT TCG GGA GCT GGG AAG AGC ACC
 M   S   G   P   R   P   V   V   L   S   G   P   S   G   A   G   K   S   T
CTG CTG AAG AGG CTG CTC CAG GAG CAC AGC GGC ATC TTT GGC TTC AGC GTG TCC CAT
 L   L   K   R   L   L   Q   E   H   S   G   I   F   G   F   S   V   S   H
ACC ACG AGG AAC CCG AGG CCC GGC GAG GAG AAC GGC AAA GAT TAC TAC TTT GTA ACC
 T   T   R   N   P   R   P   G   E   E   N   G   K   D   Y   Y   F   V   T
AGG GAG GTG ATG CAG CGT GAC ATA GCA GCC GGC GAC TTC ATC GAG CAT GCC GAG TTC
 R   E   V   M   Q   R   D   I   A   A   G   D   F   I   E   H   A   E   F
TCG GGG AAC CTG TAT GGC ACG AGC AAG GTG GCG GTG CAG GCC GTG CAG GCC ATG AAC
 S   G   N   L   Y   G   T   S   K   V   A   V   Q   A   V   Q   A   M   N
CGC ATC TGT GTG CTG GAC GTG GAC CTG CAG GGT GTG CGG AAC ATC AAG GCC ACC GAT
 R   I   C   V   L   D   V   D   L   Q   G   V   R   N   I   K   A   T   D
CTG CGG CCC ATC TAC ATC TCT GTG CAG CCG CCT TCA CTG CAC GTG CTG GAG CAG CGG
 L   R   P   I   Y   I   S   V   Q   P   P   S   L   H   V   L   E   Q   R
CTG CGG CAG CGC AAC ACT GAA ACC GAG GAG AGC CTG GTG AAG CGG CTG GCT GCT GCC
 L   R   Q   R   N   T   E   T   E   E   S   L   V   K   R   L   A   A   A
CAG GCC GAC ATG GAG AGC AGC AAG GAG CCC GGC CTG TTT GAT GTG GTC ATC ATT AAC
 Q   A   D   M   E   S   S   K   E   P   G   L   F   D   V   V   I   I   N
GAC AGC CTG GAC CAG GCC TAC GCA GAG CTG AAG GAG GCG CTC TCT GAG GAA ATC AAG
 D   S   L   D   Q   A   Y   A   E   L   K   E   A   L   S   E   E   I   K
AAA GCT CAA AGG ACC GGC GCC TAA
 K   A   Q   R   T   G   A   *
                197
```
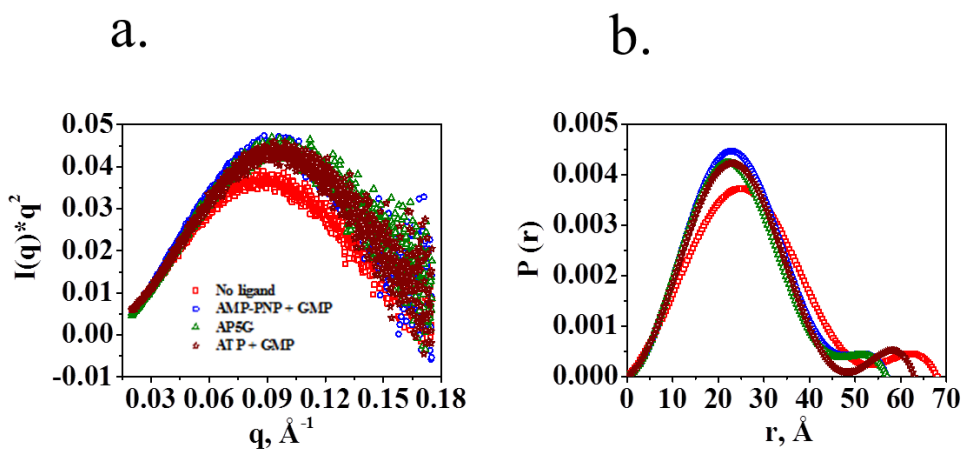

Fig. 1. Human GMPK conformations upon binding of different nucleotides. (a) Kratky profiles, and (b) computed P(r) curves.
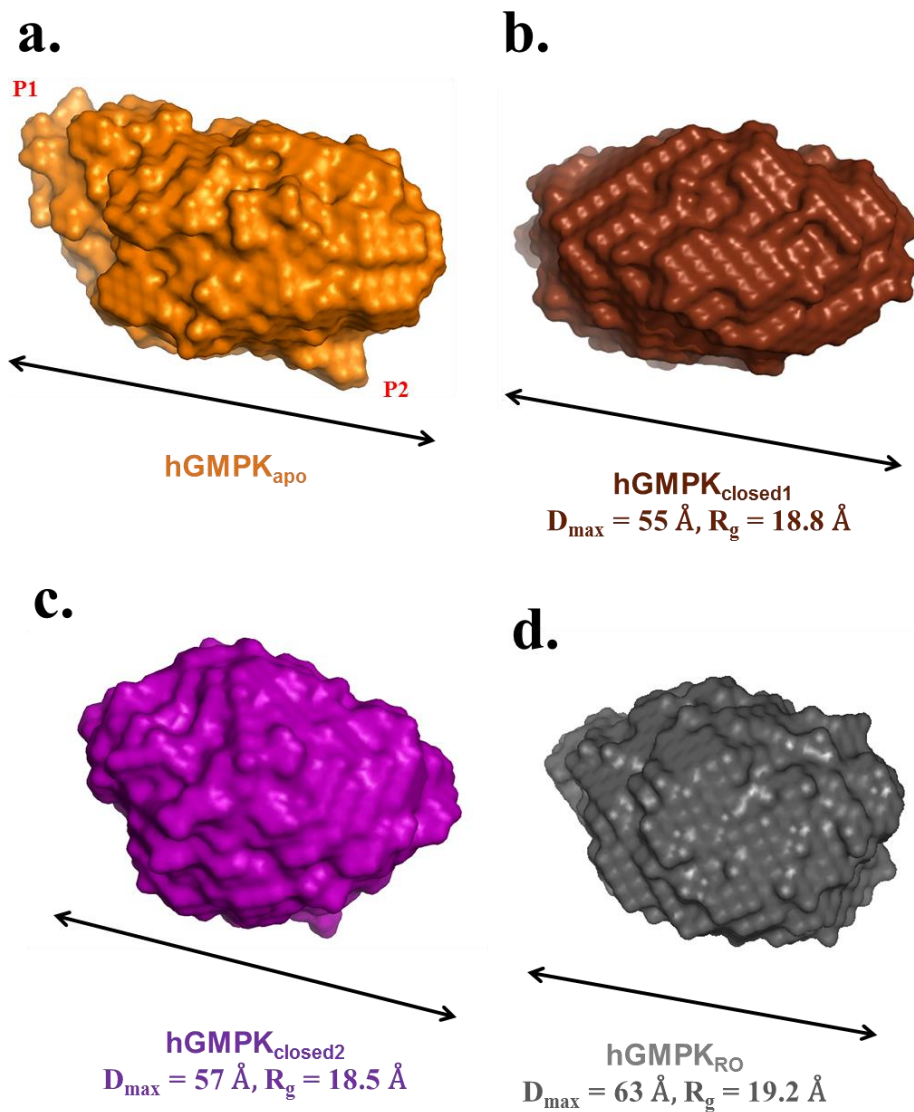
Fig. 2. Three-dimensional surface reconstructions of human GMPK in different conformational forms. (a) hGMPK in the unliganded form is represented as hGMPK$_{apo}$, (b) with bound GMP and AMP-PNP as hGMPK$_{GMP+AMP-PNP}$ or hGMPK$_{closed1}$, (c) with bound Ap5G as hGMPK$_{Ap5G}$ or hGMPK$_{closed2}$, and (d) in the presence of GMP and ATP as hGMPK$_{ATP+GMP}$ or hGMPK$_{RO}$ (reopened form).

Human mitochondrial thymidine kinase (hTK2, TK2, ATP:thymidine 5´-phosphotransferase, EC 2.7.1.21), UniProt identifier O00142-1, synthetic and codon-optimized gene for expression in *E.coli*, 265 amino acids in length,795 bp ORF

```
  1
atg ctg ctg tgg cct ctg cgt ggt tgg gct gct cgt gca ctg cgt tgc ttt ggt cct
 M   L   L   W   P   L   R   G   W   A   A   R   A   L   R   C   F   G   P
ggt tct cgt ggt agt ccg gct tct ggt cct ggt cct cgt cgt gtt cag cgt cgt gcc
 G   S   R   G   S   P   A   S   G   P   G   P   R   R   V   Q   R   R   A
tgg cct cct gat aaa gaa caa gag aaa gaa aaa aaa agc gtg atc tgt gtg gaa ggc
 W   P   P   D   K   E   Q   E   K   E   K   K   S   V   I   C   V   E   G
aat att gct agc ggc aaa acg aca tgt ctg gaa ttc ttc tcg aat gcc acc gat gtt
 N   I   A   S   G   K   T   T   C   L   E   F   F   S   N   A   T   D   V
gaa gtt ctg acc gaa ccg gtg agc aaa tgg cgt aat gtc cgt ggg cat aat cct ctg
 E   V   L   T   E   P   V   S   K   W   R   N   V   R   G   H   N   P   L
ggt ctg atg tat cat gat gct tcc cgc tgg gga ctg aca ctg cag act tat gta cag
 G   L   M   Y   H   D   A   S   R   W   G   L   T   L   Q   T   Y   V   Q
ctg acc atg ctg gat cgt cat aca cgt cct cag gtt tct agc gtt cgc ctg atg gaa
 L   T   M   L   D   R   H   T   R   P   Q   V   S   S   V   R   L   M   E
cgt tct atc cat agc gcc cgt tat atc ttc gtg gaa aat ctg tat cgt tct ggc aaa
 R   S   I   H   S   A   R   Y   I   F   V   E   N   L   Y   R   S   G   K
atg cct gaa gtt gac tat gtt gtg ctg tcc gaa tgg ttt gat tgg att ctg cgc aac
 M   P   E   V   D   Y   V   V   L   S   E   W   F   D   W   I   L   R   N
atg gat gtt agc gtg gac ctg att gtg tat ctg cgc acc aat ccg gaa aca tgc tat
 M   D   V   S   V   D   L   I   V   Y   L   R   T   N   P   E   T   C   Y
caa cgt ctg aaa aaa cgc tgc cgt gaa gag gaa aaa gtg atc cca ctg gag tat ctg
 Q   R   L   K   K   R   C   R   E   E   E   K   V   I   P   L   E   Y   L
gaa gcg att cat cat ctg cat gag gag tgg ctg att aaa ggt tcc ctg ttc cca atg
 E   A   I   H   H   L   H   E   E   W   L   I   K   G   S   L   F   P   M
gct gcc ccg gta ctg gta atc gaa gcg gac cat cac atg gag cgt atg ctg gaa ctg
 A   A   P   V   L   V   I   E   A   D   H   H   M   E   R   M   L   E   L
ttt gag caa aac cgt gac cgt att ctg acc ccg gaa aac cgt aaa cac tgc ccg taa
 F   E   Q   N   R   D   R   I   L   T   P   E   N   R   K   H   C   P   *
                                                                     265
```
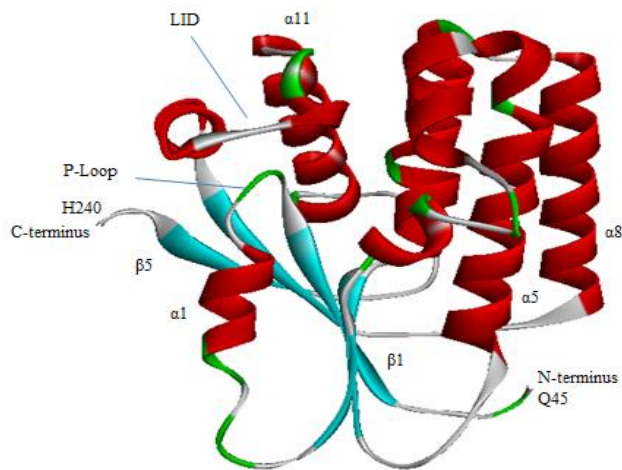

Fig. 3. Homology model of hTK2-Δ44N/Δ8C (by SWISS Model)

E.coli guanosine-inosine kinase (ecGSK, gsk, ATP:guanosine 5´-phosphotransferase, EC 2.7.1.73), UniProt P0AEW6-1, 434 amino acids and 1302 bp in length

```
 1
atg aaa ttt ccc ggt aaa cgt aaa tcc aaa cat tac ttc ccc gta aac gca cgc gat
 M   K   F   P   G   K   R   K   S   K   H   Y   F   P   V   N   A   R   D
ccg ctg ctt cag caa ttc cag cca gaa aac gaa acc agc gct gcc tgg gta gtg ggt
 P   L   L   Q   Q   F   Q   P   E   N   E   T   S   A   A   W   V   V   G
atc gat caa acg ctg gtc gat att gaa gcg aaa gtg gat gat gaa ttt att gag cgt
 I   D   Q   T   L   V   D   I   E   A   K   V   D   D   E   F   I   E   R
tat gga tta agc gcc ggg cat tca ctg gtg att gag gat gat gta gcc gaa gcg ctt
 Y   G   L   S   A   G   H   S   L   V   I   E   D   D   V   A   E   A   L
tat cag gaa cta aaa cag aaa aac ctg att acc cat cag ttt gcg ggt ggc acc att
 Y   Q   E   L   K   Q   K   N   L   I   T   H   Q   F   A   G   G   T   I
ggt aac acc atg cac aac tac tcg gtg ctc gcg gac gac cgt tcg gtg ctg ctg ggc
 G   N   T   M   H   N   Y   S   V   L   A   D   D   R   S   V   L   L   G
gtc atg tgc agc aat att gaa att ggc agt tat gcc tat cgt tac ctg tgt aac act
 V   M   C   S   N   I   E   I   G   S   Y   A   Y   R   Y   L   C   N   T
tcc agc cgt acc gat ctt aac tat cta caa ggc gtg gat ggc ccg att ggt cgt tgc
 S   S   R   T   D   L   N   Y   L   Q   G   V   D   G   P   I   G   R   C
ttt acg ctg att ggc gag tcc ggg gaa cgt acc ttt gct atc agt cca ggc cac atg
 F   T   L   I   G   E   S   G   E   R   T   F   A   I   S   P   G   H   M
aac cag ctg cgg gct gaa agc att ccg gaa gat gtg att gcc gga gcc tcg gca ctg
 N   Q   L   R   A   E   S   I   P   E   D   V   I   A   G   A   S   A   L
gtt ctc acc tca tat ctg gtg cgt tgc aag ccg ggt gaa ccc atg ccg gaa gca acc
 V   L   T   S   Y   L   V   R   C   K   P   G   E   P   M   P   E   A   T
atg aaa gcc att gag tac gcg aag aaa tat aac gta ccg gtg gtg ctg acg ctg ggc
 M   K   A   I   E   Y   A   K   K   Y   N   V   P   V   V   L   T   L   G
acc aag ttt gtc att gcc gag aat ccg cag tgg tgg cag caa ttc ctc aaa gat cac
 T   K   F   V   I   A   E   N   P   Q   W   W   Q   Q   F   L   K   D   H
gtc tct atc ctt gcg atg aac gaa gat gaa gcc gaa gcg ttg acc gga gaa agc gat
 V   S   I   L   A   M   N   E   D   E   A   E   A   L   T   G   E   S   D
ccg ttg ttg gca tct gac aag gcg ctg gac tgg gta gat ctg gtg ctg tgc acc gcc
 P   L   L   A   S   D   K   A   L   D   W   V   D   L   V   L   C   T   A
ggg cca atc ggc ttg tat atg gcg ggc ttt acc gaa gac gaa gcg aaa cgt aaa acc
 G   P   I   G   L   Y   M   A   G   F   T   E   D   E   A   K   R   K   T
cag cat ccg ctg ctg ccg ggc gct ata gcg gaa ttc aac cag tat gag ttt agc cgc
 Q   H   P   L   L   P   G   A   I   A   E   F   N   Q   Y   E   F   S   R
gcc atg cgc cac aag gat tgc cag aat ccg ctg cgt gta tat tcg cac att gcg ccg
 A   M   R   H   K   D   C   Q   N   P   L   R   V   Y   S   H   I   A   P
tac atg ggc ggg ccg gaa aaa atc atg aac act aat gga gcg ggg gat ggc gca ttg
 Y   M   G   G   P   E   K   I   M   N   T   N   G   A   G   D   G   A   L
gca gcg ttg ctg cat gac att acc gcc aac agc tac cat cgt agc aac gta cca aac
 A   A   L   L   H   D   I   T   A   N   S   Y   H   R   S   N   V   P   N
tcc agc aaa cat aaa ttc acc tgg tta act tat tca tcg tta gcg cag gtg tgt aaa
 S   S   K   H   K   F   T   W   L   T   Y   S   S   L   A   Q   V   C   K
tat gct aac cgt gtg agc tat cag gta ctg aac cag cat tca cct cgt tta acg cgc
 Y   A   N   R   V   S   Y   Q   V   L   N   Q   H   S   P   R   L   T   R
ggc ttg ccg gag cgt gaa gac agc ctg gaa gag tct tac tgg gat cgt taa
 G   L   P   E   R   E   D   S   L   E   E   S   Y   W   D   R   *
                                                          434
```
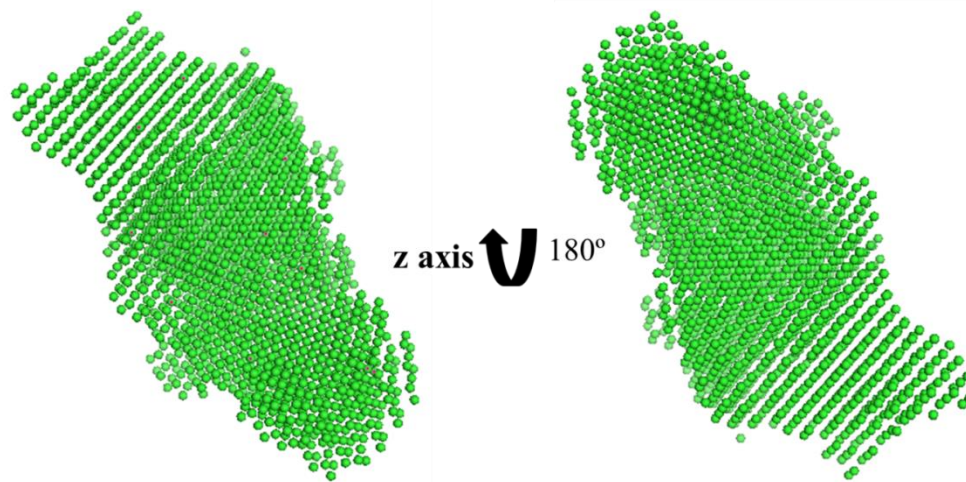
Fig. 4. Molecular surface reconstruction of ecGSK determined by SAXS

IUPred
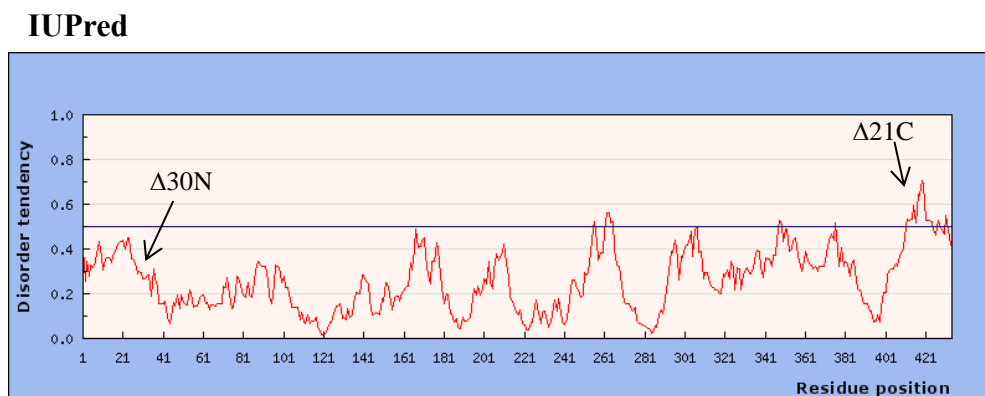
Truncated forms of ecGSK (434 aa):
1. ecGSK-Δ30N   (405 aa)
2. ecGSK-Δ21C   (413 aa)
3. ecGSK-Δ30N/Δ21C   (384 aa)

Secondary structure prediction tools:
SOPM (Geourjon and Deléage, 1994)
SOPMA (Geourjon and Deléage, 1995)
HNN (Guermeur, 1997)
MLRC (Guermeur *et al.*, 1998)
DPM (Deléage and Roux, 1987)

Fig. 5. Truncations of ecGSK. Based on the IUPred, limited proteolysis and secondary structure prediction analysis, three truncated forms of ecGSK (434 aa) were generated and recombinantly produced in *E.coli*.

Bibliography

[1]   Koshland, D.E. (1958). Application of a Theory of Enzyme Specificity to Protein Synthesis. Proceedings of the National Academy of Sciences 44, 98-104.

[2]   Gerstein, M., Lesk, A.M. and Chothia, C. (1994). Structural mechanisms for domain movements in proteins. Biochemistry 33, 6739-49.

[3]   Koshland, D.E. (1995). The Key–Lock Theory and the Induced Fit Theory. Angewandte Chemie International Edition in English 33, 2375-2378.

[4]   Jencks, W.P. (1975). Binding energy, specificity, and enzymic catalysis: the circe effect. Adv Enzymol Relat Areas Mol Biol 43, 219-410.

[5]   Muller-Dieckmann, H.J. and Schulz, G.E. (1995). Substrate specificity and assembly of the catalytic center derived from two structures of ligated uridylate kinase. J Mol Biol 246, 522-30.

[6]   Schulz, G.E., Muller, C.W. and Diederichs, K. (1990). Induced-fit movements in adenylate kinases. J Mol Biol 213, 627-30.

[7]   Gerstein, M., Schulz, G. and Chothia, C. (1993). Domain closure in adenylate kinase. Joints on either side of two helices close like neighboring fingers. J Mol Biol 229, 494-501.

[8]   Yan, H. and Tsai, M.D. (1999). Nucleoside monophosphate kinases: structure, mechanism, and substrate specificity. Adv Enzymol Relat Areas Mol Biol 73, 103-34.

[9]   Vonrhein, C., Schlauderer, G.J. and Schulz, G.E. (1995). Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. Structure 3, 483-90.

[10]   Blaszczyk, J., Li, Y., Yan, H. and Ji, X. (2001). Crystal structure of unligated guanylate kinase from yeast reveals GMP-induced conformational changes. J Mol Biol 307, 247-57.

[11]   Sekulic, N., Shuvalova, L., Spangenberg, O., Konrad, M. and Lavie, A. (2002). Structural characterization of the closed conformation of mouse guanylate kinase. J Biol Chem 277, 30236-43.

[12]   Beck, B.J., Huelsmeyer, M., Paul, S. and Downs, D.M. (2003). A mutation in the essential gene gmk (encoding guanlyate kinase) generates a requirement for adenine at low temperature in Salmonella enterica. J Bacteriol 185, 6732-5.

[13]   Choi, B. and Zocchi, G. (2007). Guanylate kinase, induced fit, and the allosteric spring probe. Biophys J 92, 1651-8.

[14]   Hall, S.W. and Kuhn, H. (1986). Purification and properties of guanylate kinase from bovine retinas and rod outer segments. Eur J Biochem 161, 551-6.

[15]   Konrad, M. (1992). Cloning and expression of the essential gene for guanylate kinase from yeast. J Biol Chem 267, 25652-5.

[16]   Ardiani, A., Goyke, A. and Black, M.E. (2009). Mutations at serine 37 in mouse guanylate kinase confer resistance to 6-thioguanine. Protein Eng Des Sel 22, 225-32.

[17]   Karran, P. (2006). Thiopurines, DNA damage, DNA repair and therapy-related cancer. Br Med Bull 79-80, 153-70.

[18]   Karran, P. and Attard, N. (2008). Thiopurines in current medical practice: molecular mechanisms and contributions to therapy-related cancer. Nat Rev Cancer 8, 24-36.

[19]   Zhang, F., Fu, L. and Wang, Y. (2013). 6-thioguanine induces mitochondrial dysfunction and oxidative DNA damage in acute lymphoblastic leukemia cells. Molecular & Cellular Proteomics 12, 3803-11.

[20]   Jain, R. et al. (2013). X-ray scattering experiments with high-flux X-ray source coupled rapid mixing microchannel device and their potential for high-flux neutron scattering investigations. The European Physical Journal E 36, 1-9.

[21]   Petri, M., Frey, S., Menzel, A., Gorlich, D. and Techert, S. (2012). Structural characterization of nanoscale meshworks within a nucleoporin FG hydrogel. Biomacromolecules 13, 1882-9.

[22]   Petri, M., Menzel, A., Bunk, O., Busse, G. and Techert, S. (2011). Concentration effects on the dynamics of liquid crystalline self-assembly: time-resolved X-ray scattering studies. J Phys Chem A 115, 2176-83.

[23]   Bowler, M.W. (2013). Conformational dynamics in phosphoglycerate kinase, an open and shut case? FEBS Letters 587, 1878-1883.

[24]   Klenow, H. and Lichtler, E. (1957). On the enzymic formation and the isolation of polyphosphates of adenine deoxyribosides. Biochim. Biophys. Acta 23, 6-12.

[25]   BERGER, A., SCHILTZ, E. and SCHULZ, G.E. (1989). Guanylate kinase from *Saccharomyces cerevisiae*. Isolation and characterization, crystallization and preliminary X-ray analysis, amino acid sequence and comparison with adenylate kinases. Eur. J. Biochem. 184, 433-443.

[26]   Konrad, M. (1992). Cloning and expression of the essential gene for guanylate kinase from yeast. J. Biol. Chem. 267, 25652-25655.

[27]   Gentry, D., Bengra, C., Ikehara, K. and Cashel, M. (1993). Guanylate kinase of *Escherichia coli* K-12. J. Biol. Chem. 268, 14316-14321.

[28]   Gaidarov, I.O., Suslov, O.N. and Abdulaev, N.G. (1993). Enzymes of the cyclic GMP metabolism in bovine retina. I. Cloning and expression of the gene for guanylate kinase. FEBS Letters 335, 81-84.

[29]   Brady, W.A., Kokoris, M.S., Fitzgibbon, M. and Black, M.E. (1996). Cloning, characterization, and modeling of mouse and human guanylate kinases. J. Biol. Chem. 271, 16734-16740.

[30]   Woods, D.F. and Bryant, P.J. (1991). The Discs-Large Tumor Suppressor Gene of Drosophila Encodes a Guanylate Kinase Homolog Localized at Septate Junctions. (1991). Cell 66, 451-464.

[31]   Yuan, B., O'Connor, T.R. and Wang, Y. (2010). 6-Thioguanine and S^6-Methylthioguanine Are Mutagenic in Human Cells. ACS Chemical Biology 11, 1021-1027.

[32]   Omari, K.E., Dhaliwal, B., Lockyer, M., Charles, I., Hawkins, A.R. and Stammers, D.K. (2006). Structure of *Staphylococcus aureus* guanylate monophosphate kinase. Acta Cryst. F62, 949-953.

[33]   Hible, G., Christova, P., Renault, L., Seclaman, E., Thompson, A., Girard, E., Munier-Lehmann, H. and Cherfils, J. (2006). Unique GMP-binding site in *Mycobacterium tuberculosis* guanosine monophosphate kinase. Proteins 62, 489–500.

[34]   Ardiani, A., Sanchez-Bonilla, M. and Black, ME. (2010). Fusion enzymes containing HSV-1 thymidine kinase mutants and guanylate kinase enhance prodrug sensitivity in vitro and in vivo. Cancer Gene Therapy 17, 86–96.

[35]   Li, Y., Zhang, Y. and Yan, H. (1996). Kinetic and Thermodynamic Characterizations of Yeast Guanylate Kinase. J. Biol. Chem. 271, 28038-28044.

[36]   Brady, W.A., Kokoris, M.S., Fitzgibbon, M. and Black, M.E. (1996). Cloning, characterization, and modeling of mouse and human guanylate kinases. J. Biol. Chem. 271, 16734-16740.

[37] Kumar, V., Spangenberg, O. and Konrad, M. (2000). Cloning of the guanylate kinase homologues AGK-1 and AGK-2 from *Arabidopsis thaliana* and characterization of AGK-1. Eur. J. Biochem. 267, 606-615.

[38] Fitzgibbon, J., Katsanis, N., Wells, D., Delhanty, J., Vallins, W. and Hunt, D.M. (1996). Human guanylate kinase (GUK1): cDNA sequence, expression and chromosomal localization. FEBS Letters 385, 185-188.

[39] Oeschger, M.P. and Bessman, M.J. (1966). Purification and Properties of Guanylate Kinase from *Escherichia coli.* J. Biol. Chem. 241, 5452-5460.

[40] Johnston, C.A., Whitney, D.S., Volkman, B.F., Doe, C.Q. and Prehoda, K.E. (2011). Conversion of the enzyme guanylate kinase into a mitotic-spindle orienting protein by a single mutation that inhibits GMP-induced closing. PNAS 108, 973–978.

[41] Johansson, M., Amiri, M. and Karlsson, A. (2005). Phosphorylation of 9-β-D-arabinofuranosylguanine monophosphate by Drosophila melanogaster guanylate kinase. Biochemical Pharmacology 70, 987-992.

[42] Van Rompay, A.R., Johansson, M. and Karlsson, A. (2000). Phosphorylation of nucleosides and nucleoside analogs by mammalian nucleoside monophosphate kinases. Pharmacology & Therapeutics 87, 189–198.

[43] Johansson, N.G. and Eriksson, S. (1996). Structure-activity relationships for phosphorylation of nucleoside analogs to monophosphates by nucleoside kinases. Acta Biochimica Polonica 43, 143-160.

[44] Reichard, P. (1988). Interactions between deoxyribonucleotide and DNA synthesis. Ann. Rev. Biochem. 57, 349-74.

[45] Van Rompay, A.R., Johansson, M. and Karlsson, A. (2003). Substrate specificity and phosphorylation of antiviral and anticancer nucleoside analogs by human deoxyribonucleoside kinases and ribonucleoside kinases. Pharmacology & Therapeutics 100, 119-139.

[46] Deville-Bonne, D., Amri, C.E., Meyer, P., Chen, Y., Agrofoglio, L.A. and Janin, J. (2010). Human and viral nucleoside/nucleotide kinases involved in antiviral drug activation: Structural and catalytic properties. Antiviral Research 86, 101-120.

[47] Chen, Y.L., Lin, D.W. and Chang, Z.F. (2008). Identification of a putative human mitochondrial thymidine monophosphate kinase associated with monocytic/macrophage terminal differentiation. (2008). Genes to Cells 13, 679-689.

[48] Topalis, D., Kumamoto, H., Amaya Velasco, M. F., Dugue, L., Haouz, A., Alexandre, J. A., Gallois-Montbrun, S., Alzari, P. M., Pochet, S., Agrofoglio, L. A. & Deville-Bonne, D. (2007). Nucleotide binding to human UMP-CMP kinase using fluorescent derivatives -a screening based on affinity for the UMP-CMP binding site. The FEBS journal. 274, 3704-14.

[49] Pai, E. F., Sachsenheimer, W., Schirmer, R. H. & Schulz, G. E. (1977). Substrate positions and induced-fit in crystalline adenylate kinase. Journal of molecular biology. 114, 37-45.

[50] Alexandre, J. A., Roy, B., Topalis, D., Pochet, S., Perigaud, C. & Deville-Bonne, D. (2007). Enantioselectivity of human AMP, dTMP and UMP-CMP kinases. Nucleic acids research. 35, 4895-904.

[51] Schulz, G. E., Muller, C. W. & Diederichs, K. (1990). Induced-fit movements in adenylate kinases. Journal of molecular biology. 213, 627-30.

[52] Sachsenheimer, W. & Schulz, G. E. (1977). Two conformations of crystalline adenylate kinase. Journal of molecular biology. 114, 23-36.

[53]   Stehle, T. and Schulz, G.E. (1992). Refined structure of the complex between guanylate kinase and its substrate GMP at 2·0 Å resolution. J. Mol. Biol. 224, 1127-1141.

[54]   Hible, G., Daalova, P., Gilles, A.M. and Cherfils, J. (2006). Crystal structures of GMP kinase in complex with ganciclovir monophosphate and Ap5G. Biochimie 88, 1157-64.

[55]   Stehle, T. and Schulz, G.E. (1990). Three-dimensional Structure of the Complex of Guanylate Kinase from Yeast with its Substrate GMP. J. Mol. Biol. 211, 249-254.

[56]   Blaszczyk, J., Li, Y., Yan, H. and Ji, X. (2001). Crystal structure of unligated guanylate kinase from yeast reveals GMP-induced conformational changes. J Mol Biol 307, 247-57.

[57]   Zhu, J., Shang, Y., Chen, J. and Zhang, M. (2012). Structure and function of the guanylate kinase-like domain of the MAGUK family scaffold proteins. Frontiers in Biology 7, 379-396.

[58]   Gentry, D., Bengra, C., Ikehara, K. and Cashel, M. (1993). Guanylate kinase of *Escherichia coli* K-12. J. Biol. Chem. 268, 14316-14321.

[59]   Hible, G., Christova, P., Renault, L., Seclaman, E., Thompson, A., Girard, E., Munier-Lehmann, H. and Cherfils, J. (2006). Unique GMP-binding site in Mycobacterium tuberculosis guanosine monophosphate kinase. Proteins 62, 489-500.

[60]   Hible, G., Renault, L., Schaeffer, F., Christova, P., Zoe Radulescu, A., Evrin, C., Gilles, A.M. and Cherfils, J. (2005). Calorimetric and crystallographic analysis of the oligomeric structure of Escherichia coli GMP kinase. J Mol Biol 352, 1044-59.

[61]   Zhang, Y., Li, Y., Wu, Y. and Yan, H. (1997). Structural and Functional Roles of Tyrosine 78 of Yeast Guanylate Kinase. Journal of Biological Chemistry 272, 19343-19350.

[62]   Augen, J. and Wold, F. (1986). How much sequence information is needed for the regulation of amino terminal acetylation of eukaryotic proteins? Trends in Biochem. Sci. 11, 494-497.

[63]   Driessen, H. P. C., de Jong, W. W., Tesser, G. I. & Bloemendeal, H. (1985). The mechanism of N-terminal acetylation of proteins. Crit. Rev. Biochem. 18, 281-325.

[64]   Bachmair, A., Finley, D. & Varshavsky, A. (1986). In vivo half-life of a protein is a function of its amino terminal residue. Science, 234, 179-186.

[65]   Hershko, A., Heller, H., Eytan, E., Kakij, G. and Rose, I.A. (1984). Role of the α-amino group of protein in ubiquitin-mediated protein breakdown. Proc. Natl. Acad. Sci. 81, 7021-7025.

[66]   Sinev, M.A., Sineva, E.V., Ittah, V. and Haas, E. (1996). Towards a mechanism of AMP-substrate inhibition in adenylate kinase from *Escherichia coli*. FEBS Letters 397, 273-276.

[67]   Zhang, Y., Li, Y. and Honggao, Y. (1997). Structural and functional roles of tyrosine-50 of yeast guanylate kinase. Techniques in Protein Chemistry 8, 679-689.

[68]   Schulz, G.E., Müller, C.W. and Diederichs, K. (1990). Induced-fit Movements in Adenylate Kinases. J. Mol. Biol. 213, 627-630.

[69]   Abele, U. and Schulz, G.E. (1995). High-resolution structures of adenylate kinase from yeast ligated with inhibitor Ap5A, showing the pathway of phosphoryl transfer. Protein Sci. 4, 1262-1271.

[70]   Berry, M. B., Meador, B., Bilderback, T., Liang, P., Glaser, M. and Phillips, G. N. Jr. (1994). The closed conformation of a highly flexible protein: the structure of E. coli adenylate kinase with bound AMP and AMPPNP. Proteins 19, 183-98.

[71]   Dreusicke, D. and Schulz, G.E. (1986). The glycine-rich loop of adenylate kinase forms a giant anion hole. FEBS Letters 208, 301-304.

[72]   Stolworthy, T.S. and Black, M.E. (2001). The mouse guanylate kinase double mutant E72Q/D103N is a functional adenylate kinase. Protein Engineering 14, 903-909.

[73]   Agarwal, R.P., Scholar, E.M., Agarwal, K.C. and Parks, R.E.Jr. (1971). Identification and isolation on a large scale of guanylate kinase from human erythrocytes. Effects of monophosphate nucleotides of purine analogs. Biochemical Pharmacology 20, 1341-1354.

[74]   Miller, W.H. and Miller, R.L. (1980). Phosphorylation of Acyclovir (Acycloguanosine) Monophosphate by GMP Kinase. The Journal of Biological Chemistry 255, 7204-7207.

[75]   Jamil, T., Fisher, R.A. and Harris, H. (1975). Studies on the properties and tissue distribution of the isozymes of guanylate kinase in man. Hum Hered. 25, 402-13.

[76]   de Mendoza, A., Suga, H. and Ruiz-Trillo, I. (2010). Evolution of the MAGUK protein gene family in premetazoan lineages. BMC Evol Biol 10, 93.

[77]   Li, Y., Spangenberg, O., Paarmann, I., Konrad, M. and Lavie, A. (2002). Structural basis for nucleotide-dependent regulation of membrane-associated guanylate kinase-like domains. J Biol Chem 277, 4159-65.

[78]   te Velthuis, A.J.W., Admiraal, J.F. and Bagowski, C.P. (2007). Molecular evolution of the MAGUK family in metazoan genomes. BMC Evol Biol 7, 129.

[79]   Anderson, J.M. (1996). Cell signalling: MAGUK magic. Current Biology 6, 382-384.

[80]   Kistner, U., Garner, C.C. and Linial, M. (1995). Nucleotide binding by the synapse associated protein SAP90. FEBS Letters 359, 159-63.

[81]   Spangenberg O. (2001). Guanylatkinase: Von einem aktiven Enzym zu einem inaktiven Multidomänen-Protein. PhD Thesis, Göttingen.

[82]   Frey, S. and Gorlich, D. (2014). A new set of highly efficient, tag-cleaving proteases for purifying recombinant proteins. J Chromatogr A 1337, 95-105.

[83]   Frey, S. and Gorlich, D. (2014). Purification of protein complexes of defined subunit stoichiometry using a set of orthogonal, tag-cleaving proteases. J Chromatogr A 1337, 106-15.

[84]   Hazra, S., Ort, S., Konrad, M. and Lavie, A. (2010). Structural and kinetic characterization of human deoxycytidine kinase variants able to phosphorylate 5-substituted deoxycytidine and thymidine analogues. Biochemistry 49, 6784-90.

[85]   McSorley, T., Ort, S., Monnerjahn C. and Konrad M. (2014). A designed equine herpes thymidine kinase (EHV4 TK) variant improves ganciclovir-induced cell-killing. Biochemical Pharmacology 87, 435-444.

[86]   Karamitros, C.S., Yashchenok, A.M., Mohwald, H., Skirtach, A.G. and Konrad, M. (2013). Preserving catalytic activity and enhancing biochemical stability of the therapeutic enzyme asparaginase by biocompatible multilayered polyelectrolyte microcapsules. Biomacromolecules 14, 4398-406.

[87]   Khalid, W., Gobel, G., Huhn, D., Montenegro, J.M., Rivera-Gil, P., Lisdat, F. and Parak, W.J. (2011). Light triggered detection of aminophenyl phosphate with a quantum dot based enzyme electrode. J Nanobiotechnology 9, 46.

[88]   Schubert, K., Khalid, W., Yue, Z., Parak, W.J. and Lisdat, F. (2010). Quantum-dot-modified electrode in combination with NADH-dependent dehydrogenase reactions for substrate analysis. Langmuir 26, 1395-400.

[89]   Riedel, M., Gobel, G., Abdelmonem, A.M., Parak, W.J. and Lisdat, F. (2013). Photoelectrochemical sensor based on quantum dots and sarcosine oxidase. Chemphyschem 14, 2338-42.

[90]   Leung, D.W., Chen, E. and Goeddel, D.V. (1989). A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. Technique 1: 11–15.

[91]   Braman, Jeff. (2010). In Vitro Mutagenesis Protocols. Methods in Molecular Biology 634, 103-109.

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

[93]   Knecht, W et al. (2007). Drosophila deoxyribonucleoside kinase mutants with enhanced ability to phosphorylate purine analogs. Gene Therapy 14, 1278-1286.

[94]   Knecht, W., Petersen, B.M. and Piskur, J. (2000). Identification of Residues Involved in the Specificity and Regulation of the Highly Efficient Multisubstrate Deoxyribonucleoside Kinase from *Drosophila melanogaster*. J. Mol. Biol. 301, 827-837.

[95]   Bullock, W.O., Fernandez, J.M. and Short, J.M. (1987). XL1-Blue: A highly efficient plasmid transforming recA *Escherichia coli* strain with ß-galactosidase selection. Biotechniques 5, 376-378.

[96]   Studier, F.W., Rosenberg, A.H., Dunn, J.J. and Dubendorff, J.W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. Methods Enzymol 185, 60-89.

[97]   Miroux, B. and Walker, J.E. (1996). Over-production of proteins in *Escherichia coli*: Mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol 260, 289-298.

[98]   Panavas, T., Sanders, C. and Butt T.R. (2009). SUMO Fusion Technology for Enhanced Protein Production in Prokaryotic and Eukaryotic Expression Systems. Methods in Molecular Biology, SUMO Protocols 497, 303-317.

[99]   Agarwal, K. C., Miech, R. P. and Parks, R. E., Jr. (1978). Guanylate kinases from human erythrocytes, hog brain, and rat liver. Methods Enzymol. 51, 483-490.

[100]   Blondin, C., Serina, L., Wiesmüller, L., Gilles, A.M. and Barzu, O. (1994). Improved Spectrophotometric Assay of Nucleoside Monophosphate Kinase Activity Using the Pyruvate Kinase/Lactate Dehydrogenase Coupling System. Analytical Biochemistry 220, 219-221.

[101]   Zheng, L., Baumann, U., and Reymond JL.(2004). An efficient one-step site-directed and site-saturation mutagenesis protocol. Nucleic Acids Research 32, e115.

[102]   Sivashanmugam, A., Murray, V., Cui, C., Zhang, Y., Wang, J. and Li, Q. (2009). Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. PROTEIN SCIENCE 18, 936-948.

[103]   Marley, J., Lu, M. and Bracken, C. (2001). A method for efficient isotopic labeling of recombinant proteins. Journal of Biomolecular NMR 20**,** 71–75.

[104]   Tolia, N.H. and Tor, L.J. (2006). Strategies for protein coexpression in *Escherichia coli*. Nature Methods 1, 55-64.

[105]   Cai, M., Huang, Y., Sakaguchi, K., Clore, G.M., Gronenborn, A.M. and Craigie, R. (1998). An efficient and cost-effective isotope labeling protocol for proteins expressed in Escherichia coli. J Biomol NMR 11, 97–102.

[106] Elgemeie, G.H. (2003). Thioguanine, Mercaptopurine: Their Analogs and Nucleosides as Antimetabolites. Current Pharmaceutical Design 9, 2627-2642.

[107] Swann, P.F et al. (1996). Role of Postreplicative DNA Mismatch Repair in the Cytotoxic Action of Thioguanine. Science 273, 1109-1111.

[108] Stolworthy, T.S., Krabbenhoft, E. and Black, M.E. (2003). A novel Escherichia coli strain allows functional analysis of guanylate kinase drug resistance and sensitivity. Anal Biochem 322, 40-7.

[109] Yoshida, S., Yamada, M., Masaki, S. & Saneyoshi, M. (1979). Utilization of 2′-deoxy-6-thioguanosine 5′-triphosphate in DNA synthesis *in vitro* by DNA polymerase α from calf thymus. *Cancer Res.* 39, 3955–3958.

[110] Ling, Y. H., Nelson, J. A., Cheng, Y. C., Anderson, R. S. & Beattie, K. L. (1991). 2′-Deoxy-6-thioguanosine 5′-triphosphate as a substrate for purified human DNA polymerases and calf thymus terminal deoxynucleotidyltransferase *in vitro*. *Mol. Pharmacol.* 40, 508–514.

[111] McSorley, T., Ort, S., Monnerjahn, C. and Konrad, M. (2014). A designed equine herpes thymidine kinase (EHV4 TK) variant improves ganciclovir-induced cell-killing. Biochem Pharmacol 87, 435-44.

[112] Tanne, J., Schafer, D., Khalid, W., Parak, W.J. and Lisdat, F. (2011). Light-controlled bioelectrochemical sensor based on CdSe/ZnS quantum dots. Anal Chem 83, 7778-85.

[113] Khalid, W. et al. (2011). Immobilization of quantum dots via conjugated self-assembled monolayers and their application as a light-controlled sensor for the detection of hydrogen peroxide. ACS Nano 5, 9870-6.

[114] Decher, G. and Schlenoff, J. (2002). Multilayer Thin Films: Sequential Assembly of Nanocomposite Materials. Wiley VCH.

[115] Lisdat, F., Dronov, R., Mohwald, H., Scheller, F.W. and Kurth, D.G. (2009). Self-assembly of electro-active protein architectures on electrodes for the construction of biomimetic signal chains. Chemical Communications, 274-283.

[116] Karamitros, C.S., Lim, J. and Konrad, M. (2014). An Amplex Red-based fluorometric and spectrophotometric assay for L-asparaginase using its natural substrate. Anal Biochem 445, 20-23.

[117] Kandeel, M. and Kitade, Y. (2011). Binding dynamics and energetic insight into the molecular forces driving nucleotide binding by guanylate kinase. J Mol Recognit 24, 322-32.

[118] Malakhov, M. P., Mattern, M. R., Mala-khova, O. A., Drinker, M., Weeks, S. D. and Butt, T. R. (2004). SUMO fusions and SUMO-specific protease for efficient expression and purification of proteins. J. Struct. Funct. Genomics 5, 75–86.

[119] Butt, T. R., Edavettal, S. C., Hall, J. P. and Mattern, M. R. (2005). SUMO fusion technology for difficult-to-express proteins. Protein Expr. Purif. 43, 1–9.

[120] Zuo, X., Mattern, M. R., Tan, R., Li, S., Hall, J., Sterner, D. E., Shoo, J., Tran, H., Lim, P., Sarafianos, S. G., Kazi, L., Navas-Martin, S., Weiss, S. R. and Butt, T. R. (2005) Enhanced expression and purification of membrane proteins by SUMO fusion in *Escherichia coli*. J. Struct. Funct. Genomics 6, 103–111.

[121] Boehme, R.E. (1984). Phosphorylation of the Antiviral Precursor 9-(1,3-Dihydroxy-2-propoxymethyl)guanine Monophosphate by Guanylate Kinase Isozymes. The Journal of Biological Chemistry 259, 12346-12349.

[122]   Gutteridge, A. and Thornton, J. (2004). Conformational change in substrate binding, catalysis and product release: an open and shut case? FEBS Letters 567, 67-73.

[123]   Fotoohi, A. K., Lindquist, M., Peterson, C. & Albertoni, F. (2006). Involvement of the concentrative nucleoside transporter 3 and equilibrative nucleoside transporter 2 in the resistance of T-lymphoblastic cell lines to thiopurines. *Biochem. Biophys. Res. Commun.* 343, 208–215.

[124]   Sargent, J.M. and Taylor, C.G. (1989). Appraisal of the MTT assay as a rapid test of chemosensitivity in acute myeloid leukaemia. Br. J. Cancer 60, 206-210.

[125]   Brieger, A., Plotz, G., Hinrichsen, I., Passmann, S., Adam, R. and Zeuzem, S. (2012). C-terminal fluorescent labeling impairs functionality of DNA mismatch repair proteins. PLoS One 7, e31863.

[126]   Han, W., Rhee, J.S., Maximov, A., Lin, W., Hammer, R.E., et al. (2005) C-terminal ECFP fusion impairs synaptotagmin 1 function: crowding out synaptotagmin 1. J Biol Chem 280, 5089–5100.

[127]   Weinheimer, I., Boonrod, K., Moser, M., Zwiebel, M., Fullgrabe, M., et al. (2010) Analysis of an autoproteolytic activity of rice yellow mottle virus silencing suppressor P1. Biol Chem 391, 271–281.

[128]   Motley, A.M., Berg, N., Taylor, M.J., Sahlender, D.A., Hirst, J., et al. (2006) Functional analysis of AP-2 alpha and mu2 subunits. Mol Biol Cell 17, 5298–5308.

[129]   Yuan, B. and Wang, Y. (2008). Mutagenic and cytotoxic properties of 6-thioguanine, S^6-methylguanine, and guanine-S^6-sulfonic acid. Journal of Biological Chemistry 283, 23665-23670.

[130]   Mora, L., Hernández, P., Vicente, J., Galán, F. and Hernández, L. (2008). Electrochemical Study and Direct Determination of Adenosine-5′-Monophosphate Coupled to 6-Thioguanosine and a Glassy Carbon Modified Electrode with Gold Nanoparticles. Electroanalysis 20, 2084-2089.

[131]   De Temmerman, M.L., Demeester, J., De Vos, F. and De Smedt, S.C. (2011). Encapsulation performance of layer-by-layer microcapsules for proteins. Biomacromolecules 12, 1283-9.

[132]   Karamitros, C.S., Yashchenok, A.M., Mohwald, H., Skirtach, A.G. and Konrad, M. (2013). Preserving catalytic activity and enhancing biochemical stability of the therapeutic enzyme asparaginase by biocompatible multilayered polyelectrolyte microcapsules. Biomacromolecules 14, 4398-406.

[133]   Skirtach, A.G., Yashchenok, A.M. and Mohwald, H. (2011). Encapsulation, release and applications of LbL polyelectrolyte multilayer capsules. Chem Commun (Camb) 47, 12736-46.

[134]   Brem, R. and Karran, P. (2012). Multiple forms of DNA damage caused by UVA photoactivation of DNA 6-thioguanine. Photochem Photobiol 88, 5-13.

[135]   Zhou, M., Diwu, Z., Panchuk-Voloshina, N. and Haugland, R.P. (1997). A stable nonfluorescent derivative of resorufin for the fluorometric determination of trace hydrogen peroxide: applications in detecting the activity of phagocyte NADPH oxidase and other oxidases. Anal. Biochem. 253, 162-168.

[136]   Petro, A.I., Volodkin, D.V. and Sukhorukov, G.B. (2005). Protein-Calcium Carbonate Coprecipitation: A Tool for Protein Encapsulation. Biotechnol. Prog. 21, 918-925.

[137]   Pechenkin, M.A., Möhwald, H. and Volodkin, D.V. (2012). pH- and salt-mediated response of layer-by-layer assembled PSS/PAH microcapsules: fusion and polymer exchange. Soft Matter 8, 8659-8665.

[138] Wang, L., Saada, A. and Eriksson, S. (2003). Kinetic properties of mutant human thymidine kinase 2 suggest a mechanism for mitochondrial DNA depletion myopathy. J Biol Chem 278, 6963-8.

[139] Barroso, J.F., Elholm, M. and Flatmark, T. (2003). Tight Binding of Deoxyribonucleotide Triphosphates to Human Thymidine Kinase 2 Expressed in *Escherichia coli*. Purification and Partial Characterization of Its Dimeric and Tetrameric Forms. Biochemistry 42, 15158-15169.

[140] WANG, L. and ERIKSSON, S. (2000). Cloning and characterization of full-length mouse thymidine kinase 2: the N-terminal sequence directs import of the precursor protein into mitochondria. Biochem. J. 351, 469-476.

[141] Saul Kit, S. and Leung, W.C. (1974). Submitochondrial Localization and Characteristics of Thymidine Kinase Molecular Forms in Parental and Kinase-Deficient HeLa Cells. Biochemical Genetics 11, 231-247.

[142] Wanga, L et al. (1999). Human thymidine kinase 2: molecular cloning and characterisation of the enzyme activity with antiviral and cytostatic nucleoside substrates. FEBS Letters 443, 170-174.

[143] Savitsky, P., Bray, J., Cooper, C.D., Marsden, B.D., Mahajan, P., Burgess-Brown, N.A. and Gileadi, O. (2010). High-throughput production of human proteins for crystallization: the SGC experience. J Struct Biol 172, 3-13.

[144] Hazra, S., Szewczak, A., Ort, S., Konrad, M. and Lavie, A. (2011). Post-translational phosphorylation of serine 74 of human deoxycytidine kinase favors the enzyme adopting the open conformation making it competent for nucleoside binding and release. Biochemistry 50, 2870-80.

[145] Barroso, J.F., Carvalho, R.N. and Flatmark, T. (2005). Kinetic Analysis and Ligand-Induced Conformational Changes in Dimeric and Tetrameric Forms of Human Thymidine Kinase 2. Biochemistry 44, 4886-4896.

[146] Burg, T.P., Godin, M., Knudsen, S.M., Shen, W., Carlson, G., Foster, J.S., Babcock, K. and Manalis, S.R. (2007). Weighing of biomolecules, single cells and single nanoparticles in fluid. Nature 446, 1066-9.

[147] Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248−254.

[148] Mori, H., Iida, A., Teshiba, S. and Fujio, T. (1995). Cloning of a Guanosine-Inosine Kinase Gene of *Escherichia coli* and Characterization of the Purified Gene Product. Journal of Bacteriology 177, 4921–4926.

[149] Harlow, K.W., Nygaard, P. and Hove-Jensen, B. (1995). Cloning and Characterization of the *gsk* Gene Encoding Guanosine Kinase of *Escherichia coli*. Journal of Bacteriology 177, 2236–2240.

[150] Wettin, K., Johansson, M., Zheng, X., Zhu, C., and Karlsson, A. (1999). Cloning of mouse mitochondrial thymidine kinase 2 cDNA. *FEBS Lett. 460*, 103-106.

[151] Eriksson, S., Munch-Petersen, B., Johansson, K. and Eklund, H. (2002). Structure and function of cellular deoxyribonucleoside kinases. Cell. Mol. Life Sci. 59, 1327-1346.

[152] Wang, L., Munch-Petersen, B., Sjoëberg, A.H., Hellman, U., Bergman, T., Joërnvall, H. and Eriksson, S. (1999). Human thymidine kinase 2: molecular cloning and characterisation of the enzyme activity with antiviral and cytostatic nucleoside substrates. FEBS Letters 443, 170-174.

[153] Johansson, M. and Karlsson, A. (1997). Cloning of the cDNA and Chromosome Localization of the Gene for Human Thymidine Kinase 2. The Journal of Biological Chemistry 272, 8454–8458.

[154] Johansson, M., Brismar, S. and Karlsson, A. (1997). Human deoxycytidine kinase is located in the cell nucleus. Proc. Natl. Acad. Sci. 94, 11941-11945.

[155] Kit, S. and Leung, W.C. (1974). Submitochondrial localization and characteristics of thymidine kinase molecular forms in parental and kinase-deficient HeLa cells. Biochemical Genetics 11, 231-247.

[156] Bogenhagen, D. and Clayton, D.A. (1976). Thymidine nucleotide supply for mitochondrial DNA synthesis in mouse L-cells. The Journal of Biological Chemistry 251, 2938-2944.

[157] Eriksson, S., Kierdaszuk, B., Munch-Petersen, B., Öberg, B. and Johansson, N.G. (1991). Comparison of the substrate specificities of human thymidine kinase 1 and 2 and deoxycytidine kinase toward antiviral and cytostatic nucleoside analogs. Biological and Biophysical Research communications 176, 586-592.

[158] Arner, E.S.J., Spasokoukotskaja, T. and Eriksson, S. (1992). Selective assays for thymidine kinase 1 and 2 and deoxycytidine kinase and their activities in extracts from human cells and tissues. Biochemical and Biophysical Research Communications 188, 712-718.

[159] Jansson, O., Bohman, C., Munch-Petersen, B. and Eriksson, S. (1992). Mammalian thymidine kinase 2. Direct photoaffinity labeling with [^{32}P] dTTP of the enzyme from spleen, liver, heart and brain. Eur. J. Biochem. 206, 485-490.

[160] Schlager, B., Straessle, A. and Hafen, E. (2012). Use of anionic denaturing detergents to purify insoluble proteins after overexpression. BMC Biotechnol 12, 95.

[161] Tao, H., Liu, W., Simmons, B.N., Harris, H.K., Cox, T.C. and Massiah, M.A. (2010). Purifying natively folded proteins from inclusion bodies using sarkosyl, Triton X-100, and CHAPS. Biotechniques 48, 61-4.

[162] Zardeneta, G. and Horowitz, P.M. (1994). Protein refolding at high concentrations using detergent/phospholipid mixtures. Analytical Biochemistry 218, 392-398.

[163] De Bernardez Clark, E. (1998). Refolding of recombinant proteins. *Current Opinion Biotechnol. 9,* 157-163.

[164] Lilie, H., Schwarz, E. & Rudolph, R. (1998). Advances in refolding of proteins produced in *E. coli*. *Current Opinion Biotechnol. 9,* 497-501.

[165] Miroux, B. and Walker, J.E. (1996). Over-production of Proteins in *Escherichia coli*: Mutant Hosts that Allow Synthesis of some Membrane Proteins and Globular Proteins at High Levels. J. Mol. Biol. 260, 289–298.

[166] Georgiou, G. and Valax, P. (1996). Expression of correctly folded proteins in *Escherichia coil*. Current Opinion in Biotechnology 7, 190-192.

[167] Baneyx, F. (1999). Recombinant protein expression in *Escherichia coli*. Current Opinion in Biotechnology 10, 411-421.

[168] Steczko, J., Donoho, G. A., Dixon, J. E., Sugimoto, T., and Axelrod, B. (1991). Effect of ethanol and low-temperature culture on expression of soybean lipoxygenase L-1 in *Escherichia coli*. Protein Expression Purif. 2, 221-227.

[169] Yura, T., Nagai, H. and Mori, H. (1993). Regulation of the heatshock response in bacteria. Annu. ReV. Microbiol. 47, 321-350.

[170] Umetsu, M., Tsumoto, K., Nitta, S., Adschiri, T., Ejima, D., Arakawa, T. and Kumagai, I. (2005). Nondenaturing solubilization of β2 microglobulin from inclusion bodies by L-arginine. Biochemical and Biophysical Research Communication 328, 189-197.

[171] Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D. and Arakawa, T. (2003). Solubilization of active green fluorescent protein from insoluble particles by guanidine and arginine. Biochemical and Biophysical Research Communication 312, 1383-1386.

[172] Guisbert, E., Herman, C., Lu, C.Z., et al. (2004). A chaperone network controls the heat shock response in *E.coli*. Genes Dev. 18, 2812-2821.

[173] Stone, K.L. and Williams, K.R. (2009). Reverse-Phase HPLC Separation of Enzymatic Digests of Proteins. The Protein Protocols Handbook 3rd edition, 937-946.

[174] Rohman, M. and Harrison-Lavoie, K.J. (2000). Separation of copurifying GroEL from glutathione-S-transferase fusion proteins. Protein Expr Purif 20, 45-47.

[175] Graslund, S. et al. (2008). The use of systematic N- and C-terminal deletions to promote production and structural studies of recombinant proteins. Protein Expr Purif 58, 210-21.

[176] Sagemark, J., Kraulis, P. and Weigelt, J. (2010). A software tool to accelerate design of protein constructs for recombinant expression. Protein Expr Purif 72, 175-8.

[177] Savitsky, P., Bray, J., Cooper, C.D., Marsden, B.D., Mahajan, P., Burgess-Brown, N.A. and Gileadi, O. (2010). High-throughput production of human proteins for crystallization: the SGC experience. J Struct Biol 172, 3-13.

[178] Dosztanyi, Z., Csizmok, V., Tompa, P. and Simon, I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. Bioinformatics 21, 3433-4.

[179] Sun, R., Eriksson, S. and Wang, L. (2012). Oxidative stress induced S-glutathionylation and proteolytic degradation of mitochondrial thymidine kinase 2. J Biol Chem 287, 24304-12.

[180] Hazra, S., Szewczak, A., Ort, S., Konrad, M. and Lavie, A. (2011). Post-translational phosphorylation of serine 74 of human deoxycytidine kinase favors the enzyme adopting the open conformation making it competent for nucleoside binding and release. Biochemistry 50, 2870-80.

[181] Cirino, P.C., Mayer, K.M. and Umeno, D. (2003). Generating mutant libraries using error-prone PCR. Methods in Molecular Biology 231, 3-9.

[182] Hiraga, S., Igarashi, K. and Yura, T. (1967). A deoxythymidine kinase-deficient mutant of *Escherichia Coli*. I. Isolation and some properties. Biochim. Biophys. Acta 145, 41-51.

[183] Lutz, S., Liu, L. and Liu, Y. (2009). Engineering Kinases to Phosphorylate Nucleoside Analogs for Antiviral and Cancer Therapy. Chimia (Aarau) 63, 737-744.

[184] Rompay, A.R.V., Johansson, M. and Karlsson, A. (2003). Substrate specificity and phosphorylation of antiviral and anticancer nucleoside analogues by human deoxyribonucleoside kinases and ribonucleoside kinases. Pharmacology & Therapeutics 100, 119– 139.

[185] Jochimsen, B., Nygaard, P. and Vestergaard, T. (1975). Location on the chromosome of *Escherichia coli* of genes governing purine metabolism. Molecular and General Genetics 143, 85-91.

[186] Hove-Jensen, B. and Nygaard, P. (1989). Role of guanosine kinase in the utilization for nucleotide synthesis in *Escherichia coli*. Journal of General Microbiology 135, 1263-1273.

[187] Wang, L., Hellman, U. and Eriksson, S. (1996). Cloning and expression of human mitochondrial deoxyguanosine kinase cDNA. FEBS Letters 390, 39-43.

[188] Sjöberg, A.H., Wang, L. and Eriksson, S. (1998). Substrate specificity of human recombinant mitochondrial deoxyguanosine kinase with cytostatic and antiviral purine and pyrimidine analogs. Mol Pharmacol. 53, 270-273.

[189] Herrstrom Sjoberg, A., Wang, L. and Eriksson, S. (2001). Antiviral guanosine analogs as substrates for deoxyguanosine kinase: implications for chemotherapy. Antimicrob Agents Chemother 45, 739-42.

[190] Solaroli, N., Johansson, M., Balzarini, J. and Karlsson, A. (2007). Enhanced toxicity of purine nucleoside analogs in cells expressing Drosophila melanogaster nucleoside kinase mutants. Gene Ther 14, 86-92.

[191] Benvenuti, M. and Mangani, S. (2007). Crystallization of soluble proteins in vapor diffusion for x-ray crystallography. Nat Protoc 2, 1633-51.

[192] Dessau, M.A. and Modis, Y. (2011). Protein crystallization for X-ray crystallography. J Vis Exp. 47, e2285, doi: 10.3791/2285.

[193] Arnér, E. S. J. and Eriksson, S. (1995) Mammalian deoxyribonucleoside kinases. Pharmacol. Ther. 67: 155–186.

[194] Austin, W.R. et al. (2012). Nucleoside salvage pathway kinases regulate hematopoiesis by linking nucleotide metabolism with replication stress. J Exp Med 209, 2215-28.

[195] Reichard, P. (1988). Interactions between deoxyribonucleotide and DNA synthesis. Annu Rev Biochem 57, 349–374.

[196] Greenberg, G. R., & Hilfinger, J. M. (1996). Regulation of synthesis of ribonucleotide reductase and relationship to DNA replication in various systems. Prog Nucleic Acid Res Mol Biol 53, 345– 395.

[197] Burton, T. R., Kashour, T., Wright, J. A., & Amara, F. M. (2003). Cellular signaling pathways affect the function of ribonucleotide reductase mRNA binding proteins: mRNA stabilization, drug resistance, and malignancy (Review). Int J Oncol 22, 21–31.

[198] Fairman, J.W., S.R. Wijerathna, M.F. Ahmad, H. Xu, R. Nakano, S. Jha, J. Prendergast, R.M. Welin, S. Flodin, A. Roos, et al. (2011). Structural basis for allosteric regulation of human ribonucleotide reductase by nucleotide-induced oligomerization. *Nat. Struct. Mol. Biol.* 18, 316– 322.

[199] Munch-Petersen, B., Cloos, L., Tyrsted, G. & Eriksson, S. (1991). Diverging substrate specificity of pure human thymidine kinases 1 and 2 against antiviral dideoxynucleosides. J Biol Chem 266, 9032– 9038.

[200] Ellims, P. H. & Van der Weyden, M. B. (1981). Kinetic mechanism and inhibition of human liver thymidine kinase. Biochim Biophys Acta 660, 238– 242.

[201] Shewach, D. S., Reynolds, K. K. & Hertel, L. (1992). Nucleotide specificity of human deoxycytidine kinase. Mol Pharmacol 42, 518– 524.

[202] White, J. C., & Hines, L. H. (1987). Role of uridine triphosphate in the phosphorylation of 1-h-D-arabinofuranosylcytosine by Ehrlich ascites tumor cells. Cancer Res 47, 1820– 1824.

[203] Turk, B., Awad, R., Usova, E., Bjo¨rk, I. & Eriksson, S. (1999). A presteady-state kinetic analysis of substrate binding to human recombinant deoxycytidine kinase: A model for nucleoside kinase action. Biochemistry 38, 8555– 8561.

[204] Gower Jr., W. R., Carr, M. C., & Ives, D. H. (1979). Deoxyguanosine kinase. Distinct molecular forms in mitochondria and cytosol. J Biol Chem 254, 2180– 2183.

[205] Zhu, C., Johansson, M., Permert, J. & Karlsson, A. (1998). Phosphorylation of anticancer nucleoside analogs by human mitochondrial deoxyguanosine kinase. Biochem Pharmacol 56, 1035– 1040.

[206]   Welin, M., Kosinska, U., Mikkelsen, N., Carnrot, C., Zhu, C., Wang, L., Eriksson, S., Munch-Petersen, B. and Eklund, H. (2004). Structures of thymidine kinase 1 of human and mycoplasmic origin. Proc. Natl. Acad. Sci. U.S.A. 101, 17970–17975.

[207]   Mutahir, Z., Clausen, A.R., Andersson, K.M., Wisen, S.M., Munch-Petersen, B. and Piskur, J. (2013). Thymidine kinase 1 regulatory fine-tuning through tetramer formation. FEBS J 280, 1531-41.

[208]   Sabini, E., Ort, S., Monnerjahn, C., Konrad, M. and Lavie, A. (2003). Structure of human dCK suggests strategies to improve anticancer and antiviral therapy. Nat. Struct. Biol. 10, 513–519.

[209]   Munch-Petersen, B. (2000). Functional Expression of a Multisubstrate Deoxyribonucleoside Kinase from Drosophila melanogaster and Its C-terminal Deletion Mutants. Journal of Biological Chemistry 275, 6673-6679.

[210]   Johansson, K., Ramaswamy, S., Ljungcrantz, C., Knecht, W., Piskur, J., Munch-Petersen, B., Eriksson, S. and Eklund, H. (2001). Structural basis for substrate specificities of cellular deoxyribonucleoside kinases. Nat. Struct. Biol. 8, 616–620.

[211]   Wang, L., Choudhury, D., Chattopadhyaya, J. and Eriksson, S. (1999). Stereoisomeric selectivity of human deoxyribonucleoside kinases. Biochemistry 38, 16993–16999.

[212]   Verri, A., Focher, F., Priori, G., Gosselin, G., Imbach, J.L., Capobianco, M., Garbesi, A. and Spadari, S. (1997). Lack of enantiospecificity of human 2'-deoxycytidine kinase: relevance for the activation of beta-L-deoxycytidine analogs as antineoplastic and antiviral agents. Molecular Pharmacolohy, 32–138.

[213]   Gaubert, G., Gosselin, G., Boudou, V., Imbach, J.L., Eriksson, S. and Maury, G. (1999). Low enantioselectivities of human deoxycytidine kinase and human deoxyguanosine kinase with respect to 2´-deoxyadenosine, 2´-deoxyguanosine and their analogs. Biochimie 81, 1041–1047.

[214]   Gentry, G. A. (1992). Viral thymidine kinases and their relatives. Pharmacol Ther 54, 319– 355.

[215]   Eriksson, S. & Wang, L. (1997). Substrate specificities, expression and primary sequences of deoxynucleoside kinases; implication for chemotherapy. Nucleosides Nucleotides 16, 653–659.

[216]   Wild, K., Bohner, T., Folkers, G. and Schulz, G.E. (1997). The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue. Protein Sci. 6, 2097–2106.

[217]   Edwald, B., Sampath, D. and Plunkett, W. (2008). Nucleoside analogs: molecular mechanisms signaling cell death. Oncogene 27, 6522-6537.

[218]   Galmarini, C. M., Mackey, J. R. & Dumontet, C. (2002). Nucleoside analogues and nucleobases in cancer treatment. Lancet Oncol 3, 415–424.

[219]   Lesko, N., Naess, K., Wibom, R., Solaroli, N., Nennesmo, I., von Dobeln, U., Karlsson, A. and Larsson, N.G. (2010). Two novel mutations in thymidine kinase-2 cause early onset fatal encephalomyopathy and severe mtDNA depletion. Neuromuscul Disord 20, 198-203.

[220]   Liu, S. H., Grove, K. L. and Cheng, Y. C. (1998). Unique metabolism of a novel antiviral L-nucleoside analog, 2´-fluoro-5-methyl-beta-L-arabinofuranosyluracil: a substrate for both thymidine kinase and deoxycytidine kinase. Antimicrob. Agents Chemother. 42, 833–839.

[221]   Maury, G. (2000). The enantioselectivity of enzymes involved in current antiviral therapy using nucleoside analogues: a new strategy? Antivir. Chem. Chemother. 11, 165–189.

[222]   Priego, E.M., Karlsson, A., Gago, F., Camarasa, M.J., Balzarini, J. and Pérez-Pérez, M.J. (2012). Recent advances in thymidine kinase 2 (TK2) inhibitors and new perspectives for potential applications. Curr Pharm Des. 18, 2981-94.

[223]   Ikeda, S., Chakravarty, R. and Ives, D. H. (1986). Multisubstrate analogs for deoxynucleoside kinases. Triphosphate end products and synthetic bisubstrate analogs exhibit identical modes of binding and are useful probes for distinguishing kinetic mechanisms. J. Biol. Chem. 261, 15836-15843.

[224]   Kim, M. Y. and Ives, D. H. (1989). Human deoxycytidine kinase: kinetic mechanism and end product regulation. Biochemistry28, 9043-9047.

[225]   Mikkelsen, N. E., Johansson, K., Karlsson, A., Knecht, W., Andersen, G., Piskur, J., Munch-Petersen, B. and Eklund, H. (2003). Structural Basis for Feedback Inhibition of the Deoxyribonucleoside Salvage Pathway: Studies of the Drosophila Deoxyribonucleoside Kinase. Biochemistry 42, 5706-5712.

[226]   Lee, L. S. and Cheng, Y. (1976). Human deoxythymidine kinase II: substrate specificity and kinetic behavior of the cytoplasmic and mitochondrial isozymes derived from blast cells of acute myelocytic leukemia. Biochemistry 15, 3686-3690.

[227]   Söderlund, J. C. and Arner, E. S. (1994). Mitochondrial versus cytosolic activities of deoxyribonucleoside salvage enzymes. Adv. Exp. Med. Biol. 370, 201-204.

[228]   Elholm, M., Hollas, H., Issalene, C., Barroso, J. F., Berge, R. K. and Flatmark, T. (2001). Transient up-regulation of liver mitochondrial thymidine kinase activity in proliferating mitochondria. IUBMB Life 51, 99-104.

[229]   Saada, A., Shaag, A., Mandel, H., Nevo, Y., Eriksson, S. and Elpeleg, O. (2001). Mutant mitochondrial thymidine kinase in mitochondrial DNA depletion myopathy. Nat Genet 29, 342-4.

[230]   Balzarini, J., Van Daele, I., Negri, A., Solaroli, N., Karlsson, A., Liekens, S., Gago, F. and Van Calenbergh, S. (2009). Human mitochondrial thymidine kinase is selectively inhibited by 3'-thiourea derivatives of β-thymidine: identification of residues crucial for both inhibition and catalytic activity. Mol Pharmacol 75, 1127-36.

[231]   Sun, R., Eriksson, S. and Wang, L. (2010). Identification and characterization of mitochondrial factors modulating thymidine kinase 2 activity. Nucleosides, Nucleotides and Nucleic acids 29, 382-385.

[232]   Radivoyevitch, T., Munch-Petersen, B., Wang, L. and Eriksson, S. (2011). Nucleosides, Nucleotides and Nucleic acids 30, 203-209.

[233]   Wang, L., Sun, R. and Eriksson, S. (2011). The kinetic effects on thymidine kinase 2 by enzyme-bound dTTP may explain the mitochondrial side effects of antiviral thymidine analogs. Antimicrobial Agents and Chemotherapy 55, 2552-2558.

[234]   Mancuso, M., Filosto, M., Bonilla, E., Hirano, M., Shanske, S., Vu, T.H. and DiMauro, S. (2003). Mitochondrial Myopathy of Childhood Associated With Mitochondrial DNA Depletion and a Homozygous Mutation (T77M) in the TK2 Gene. Arch Neurol. 60, 1007-1009.

[235]   Mancuso, M. et al. (2002). Mitochondrial DNA depletion: mutations in thymidine kinase gene with myopathy and SMA. Neurology 59, 1197-1202.

[236]   Saada, A., Ben-Shalom, E., Zyslin, R., Miller, C., Mandel, H. and Elpeleg, O. (2003). Mitochondrial deoxyribonucleoside triphosphate pools in thymidine kinase 2 deficiency. Biochemical and Biophysical Research Communication 310, 963-966.

[237] Oskoui, M. et al. (2006). Clinical spectrum of mitochondrial DNA depletion due to mutations in the thymidine kinase 2 gene. Arch Neurol. 63, 1122-1126.

[238] Alberio, S., Mineri, R., Tiranti, V. and Massimo, Z. (2007). Depletion of mtDNA: Syndromes and genes. Mitochondrion 7, 6-12.

[239] Gotz, A. et al. (2008). Thymidine kinase 2 defects can cause multi-tissue mtDNA depletion syndrome. Brain 131, 2841-50.

[240] Akman, H.O. et al. (2008). Thymidine kinase 2 (H126N) knockin mice show the essential role of balanced deoxynucleotide pools for mitochondrial DNA maintenance. Hum Mol Genet 17, 2433-40.

[241] Wang, L. (2010). Deoxynucleoside Salvage Enzymes and Tissue Specific Mitochondrial DNA Depletion. Nucleosides, Nucleotides and Nucleic acids 29, 370-381.

[242] Bartesaghi, S., Betts-Henderson, J., Cain, K., Dinsdale, D., Zhou, X., Karlsson, A., Salomoni, P. and Nicotera, P. (2010). Loss of thymidine kinase 2 alters neuronal bioenergetics and leads to neurodegeneration. Hum Mol Genet 19, 1669-77.

[243] Lesko, N. et al. (2010). Two novel mutations in thymidine kinase-2 cause early onset fatal encephalomyopathy and severe mtDNA depletion. Neuromuscular Disorders 20, 198-203.

[244] Jochimsen, B., Nygaard, P. and Vestergaard, T. (1975). Location on the Chromosome of Escherichia coli of Genes Governing Purine Metabolism. Molec. gen. Genet. 143, 85-91.

[245] Ipata, P.L., Gualerzi, C., Scolozzi, C., Tozzi, M.G., Trinei, M. and Barsacchi, D. (1995). Occurrence of Inosine Kinase as a Distinct Enzyme in *Spirulina platensis*. Biochemical and Biophysical Research Communications 209, 547-553.

[246] Bachmann, B. J. (1990). Linkage map of *Escherichia coli* K-12, edition 8. Microbiol. Rev. 54,130–197.

[247] Yoshikawa, M., Kato, T. and Takenishi, T. (1969). Studies of phosphorylation. III. Selective phosphosphorylation of unprotected nucleosides. Bull. Chem. Soc. Jpn. 42, 3505–3508.

[248] Mori, H., Iida, A., Fujio, T. and Teshiba, S. (1997). A novel process of inosine 5´-monophosphate production using overexpressed guanosine/inosine kinase. Appl Microbiol Biotechnol 48, 693-698.

[249] Usuda, Y., Kawasaki, H., Shimaoka, M. and Utagawa, T. (1997). Molecular cloning and transcriptional analysis of a guanosine kinase gene of Brevibacterium acetylicum ATCC 953. Journal of bacteriology 179, 6959-6964.

[250] Kawasaki, H., Usuda, Y., Shimaoka, M. and Utagawa, T. (2000). Phosphorylation of Guanosine Using Guanosine-inosine Kinase from *Exiguobacterium acetylicum* Coupled with ATP Regeneration. Biosci. Biotechnol. Biochem. 64, 2259-2261.

[251] Kawasaki, H., Usuda, Y., Shimaoka, M. and Utagawa, T. (2000). Development of an assay for purine nucleoside kinase activity in cell extracts and detection of inosine kinase activity in *Brevibacterium acetylicum* ATCC 953, related species, and *Corynebacterium flaccumfaciens* ATCC 6887. Biosci. Biotechnol. Biochem. 64, 761-766.

[252] Kawasaki, H., Shimaoka, M., Usuda Y. and Utagawa, T. (2000). End-product regulation and kinetic mechanism of guanosine-inosine kinase from *Escherichia coli*. Biosci. Biotechnol. Biochem. 64, 972-979.

[253] Matsui, H., Shimaoka, M., Takenaka, Y. and Kawasaki, H. (2001). Gsk disruption leads to guanosine accumulation in *Escherichia coli*. Biosci. Biotechnol. Biochem. 65, 1230-1235.

[254] Wang, H. and Wang, Y. (2010). LC-MS/MS coupled with stable isotope dilution method for the quantification of 6-thioguanine and S^6-methylthioguanine in genomic DNA of human cancer cells treated with 6-thioguanine. Anal Chem. 82, 5797-5803.

[255] Lehn, J.M. (1993). Supramolecular chemistry. Science 260, 1762-1763.

[256] Cheng, K., Blumen, S.R., MacPherson, M.B., Steinbacher, J.L., Mossman, B.T. and Landry, C.C. (2010). Enhanced uptake of porous silica microparticles by bifunctional surface modification with a targeting antibody and a biocompatible polymer. ACS Appl Mater Interfaces 2, 2489-95.

[257] Dendukuri, D., Tsoi, K., Hatton, T.A. and Doyle, P.S. (2005). Controlled synthesis of nonspherical microparticles using microfluidics. Langmuir 21, 2113-2116.

[258] Hwang, D.K., Dendukuri, D. and Doyle, P.S. (2008). Microfluidic-based synthesis of non-spherical magnetic hydrogel microparticles. Lab Chip 8, 1640-1647.

[259] Dugyala, V.R., Daware, S.V. and Basavaraj, M.G. (2013). Shape anisotropic colloids: synthesis, packing behavior, evaporation driven assembly, and their application in emulsion stabilization. Soft Matter 9, 6711-6725.

[260] Simone, E.A., Dziubla, T.D. and Muzykantov, V.R. (2008). Polymeric carriers: role of geometry in drug delivery. Expert Opin Drug Deliv 5, 1283-1300.

[261] Delcea, M., Madaboosi, N., Yashchenok, A.M., Subedi, P., Volodkin, D.V., De Geest, B.G., Mohwald, H. and Skirtach, A.G. (2011). Anisotropic multicompartment micro- and nano-capsules produced via embedding into biocompatible PLL/HA films. Chem Commun (Camb) 47, 2098-100.

[262] Gratton, S.E., Ropp, P.A., Pohlhaus, P.D., Luft, J.C., Madden, V.J., Napier, M.E. and DeSimone, J.M. (2008). The effect of particle design on cellular internalization pathways. Proc Natl Acad Sci U S A 105, 11613-11618.

[263] Geng, Y., Dalhaimer, P., Cai, S., Tsai, R., Tewari, M., Minko, T. and Discher, D.E. (2007). Shape effects of filaments versus spherical particles in flow and drug delivery. Nat Nanotechnol 2, 249-55.

[264] Champion, J.A., Walker, A. and Mitragotri, S. (2008). Role of particle size in phagocytosis of polymeric microspheres. Pharm Res 25, 1815-21.

[265] Mitragotri, S. (2009). In drug delivery, shape does matter. Pharm Res 26, 232-4.

[266] Champion, J.A. and Mitragotri, S. (2006). Role of target geometry in phagocytosis. Proc Natl Acad Sci U S A 103, 4930-4.

[267] Delcea, M., Yaschenok, A., Videnova, K., Kreft, O., Möhwald, H. and Skirtach, A.G. (2010). Multicompartmental micro- and nanocapsules: hierarchy and applications in biosciences. Macromol. Bioscience 10, 465-474.

[268] Yan, Y., Such, G.K., Johnston, A.P.R., Lomas, H. and Caruso, F. (2011). Toward Therapeutic Delivery with Layer-by-Layer Engineered Particles. ACS Nano 5, 4252-4257.

[269] Skirtach, A.G., Munoz Javier, A., Kreft, O., Kohler, K., Piera Alberola, A., Mohwald, H., Parak, W.J. and Sukhorukov, G.B. (2006). Laser-induced release of encapsulated materials inside living cells. Angew Chem Int Ed Engl 45, 4612-4617.

[270] Rivera-Gil, P., De Koker, S., De Geest, B. G. and Parak, W. J. (2009). Intracellular Processing of Proteins Mediated by Biodegradable Polyelectrolyte Capsules. Nano Lett. 9, 4398-4402.

[271] Volodkin, D.V., Petrov, A.I., Prevot, M. and Sukhorukov, G.B. (2004). Matrix polyelectrolyte microcapsules: new system for macromolecule encapsulation. Langmuir 20, 3398–3406.

[272] Mann, S. (2001). Biomineralization: principles and concepts in bioinorganic materials chemistry. Oxford University Press, Oxford.

[273] Yu, S.H., Cölfen, H. and Antonietti, M. (2003). Polymer-Controlled Morphosynthesis and Mineralization of Metal Carbonate Superstructures. J. Phys. Chem. B 107, 7396-7405.

[274] De Geest, B.G., De Koker, S., Sukhorukov, G.B., Kreft, O., Parak, W.J., Skirtach, A.S., Demeester, J., De Smedt, S.C. and Hennink, W.E. (2009). Polyelectrolyte Microcapsules for Biomedical Applications. Soft Matter 5, 282-291.

[275] Lee, I., Han, H. and Lee, S.-Y. (2010). Growth of aragonite calcium carbonate nanorods in the biomimetic anodic aluminum oxide template. Journal of Crystal Growth 312, 1741-1746.

[276] Zhang, Z., Gao, D., Zhao, H., Xie, C., Guan, G, Wang, D., Yu, S.H. (2006). Biomimetic assembly of polypeptide-stabilized $CaCO_3$ nanoparticles. J Phys Chem B. 110, 8613–8618.

[277] Guo, X.H., Xu, A.W. and Yu, S.H. (2008). Crystallization of Calcium Carbonate Mineral with Hierarchical Structures in DMF Solution under Control of Poly(ethyleneglycol)-b-poly(L-glutamic acid): Effects of Crystallization Temperature and Polymer Concentration. Cryst. Growth & Design 8, 1233–1242.

[278] Gebauer, D., Verch, A., Börner, H.G., and Helmut Cölfen. (2009). Influence of Selected Artificial Peptides on Calcium Carbonate Precipitation - A Quantitative Study. Crystal Growth & Design 9, 2398-2403.

[279] Keene, E.C., Evans, J.S. and Estroff, L.A. (2010). Silk Fibroin Hydrogels Coupled with the n16N−β-Chitin Complex: An in Vitro Organic Matrix for Controlling Calcium Carbonate Mineralization. Crystal Growth & Design 10, 5169-5175.

[280] Gebauer, D., Völkel, A. and Cölfen, H. (2008). Stable prenucleation calcium carbonate clusters. Science 322, 1819-1822.

[281] Parakhonskiy, B.V., Haase, A. and Antolini, R. (2012). Sub-Micrometer Vaterite Containers: Synthesis, Substance Loading, and Release. Angewandte Chemie International Edition 51, 1195-1197.

[282] Cölfen, H. and Qi, L. (2001). A Systematic Examination of the Morphogenesis of Calcium Carbonate in the Presence of a Double-Hydrophilic Block Copolymer. Chem. Eur. J. 7, 106-116.

[283] Qiu, X. P., Leporatti, S., Donath, E. and Möhwald, H. (2001). Studies on the drug release properties of polysaccharide multilayers encapsulated ibuprofen microparticles. Langmuir 17, 5375-5380.

[284] [284] Stoll, C., Kudera, S., Parak, W.J. and Lisdat, F. (2006). Quantum dots on gold: electrodes for photoswitchable cytochrome C electrochemistry. Small 2, 741-3.

[285] Katz, E., Zayats, M., Willner, I. and Lisdat, F. (2006). Controlling the direction of photocurrents by means of CdS nanoparticles and cytochrome c-mediated biocatalytic cascades. Chem Commun (Camb), 1395-7.

[286] Stoll, C., Gehring, C., Schubert, K., Zanella, M., Parak, W.J. and Lisdat, F. (2008). Photoelectrochemical signal chain based on quantum dots on gold--sensitive to superoxide radicals in solution. Biosens Bioelectron 24, 260-5.

[287] Liu, Q., Lu, X., Li, J., Yao, X. and Li, J. (2007). Direct electrochemistry of glucose oxidase and electrochemical biosensing of glucose on quantum dots/carbon nanotubes electrodes. Biosens Bioelectron 22, 3203-9.

[288] Wang, Z., Xu, Q., Wang, H.-Q., Yang, Q., Yu, J.-H. and Zhao, Y.-D. (2009). Hydrogen peroxide biosensor based on direct electron transfer of horseradish peroxidase with vapor deposited quantum dots. Sensors and Actuators B: Chemical 138, 278-282.

[289] Tang, L., Zhu, Y., Yang, X., Sun, J. and Li, C. (2008). Self-assembled CNTs/CdS/dehydrogenase hybrid-based amperometric biosensor triggered by photovoltaic effect. Biosens Bioelectron 24, 319-23.

[290] Henglein, A. (1989). Small-Particle Research: Physicochemical Properties of Extremely Small Colloidal Metal and Semiconductor Particles. Chem Rev 89, 1861–1873.

[291] Alivisatos, A.P. (1996). Semiconductor clusters, nanocrystals, and quantum dots. Science 271, 933–937.

[292] Nirmal, M. and Brus, L. (1999). Luminescence Photophysics in Semiconductor Nanocrystals. Acc Chem Res 32, 407–414.

[293] Chan, W.C. (1998). Quantum Dot Bioconjugates for Ultrasensitive Nonisotopic Detection. Science 281, 2016-2018.

[294] Resch-Genger, U., Grabolle, M., Cavaliere-Jaricot, S., Nitschke, R. and Nann, T. (2008). Quantum dots versus organic dyes as fluorescent labels. Nat Methods 5, 763-75.

[295] Wang, C., Gao, X. and Su, X. (2010). In vitro and in vivo imaging with quantum dots. Anal Bioanal Chem 397, 1397-415.

[296] Kucur, E., Riegler, J., Urban, G.A. and Nann, T. (2003). Determination of quantum confinement in CdSe nanocrystals by cyclic voltammetry. Journal of Chemical Physics 119, 2333-2337.

[297] Ehlert, O., Tiwari, A. and Nann, T. (2006). Quantum confinement of the thermodynamic functions for the formation of electrons and holes in CdSe nanocrystals. Journal of Applied Physics 100, 074314.

[298] Kucur, E., Bucking, W., Arenz, S., Giernoth, R. and Nann, T. (2006). Heterogeneous charge transfer of colloidal nanocrystals in ionic liquids. Chemphyschem 7, 77-81.

[299] Yue, Z. et al. (2010). Evaluation of quantum dots applied as switchable layer in a light-controlled electrochemical sensor. Anal Bioanal Chem 396, 1095-103.

[300] Niemeyer, C.M. (2003). Functional hybrid devices of proteins and inorganic nanoparticles. Angew Chem Int Ed Engl 42, 5796-800.

[301] Miller, R.L., Adamczyk, D.L. and Spector, T. (1977). Reassessment of the interactions of guanylate kinase and 6-thioguanosine 5´-phosphate. Biochemical Pharmacology 26, 1573-1576.

[302] El Kouni, M.H. and Cha, Sungman. (1981). A Simple Radioisotopic Assay for Nucleoside Kinases Employing Alumina for Separation of Nucleosides and Nucleotides. Analytical Biochemistry 111, 67-71.

[303] Maness, P.F. and Orengo, A. (1976). Differences between Pyrimidine Nucleoside Monopbosphate Kinase from Rat Novikoff Ascites Hepatoma and Rat Liver. Cancer Research 36, 2312-2316.

[304] Schaffer, M., Huribert, R.B. and Orengo, A. (1973). Guanosine Anabolism for Biosynthesis of Nucleic Acids in Novikoff Ascites Rat Tumor Cells in Culture. Cancer Research 33, 2265-2272.

[305] Sedewitz, B., Schleifer, K.H. and Götz, F. (1984). Purification and biochemical characterization of pyruvate oxidase from *Lactobacillus plantarum*. Journal of Bacteriology 160, 273-278.

[306]   Sedewitz, B., Schleifer, K.H. and Götz, F. (1984). Physiological role of pyruvate oxidase in the aerobic metabolism of *Lactobacillus plantarum*. Journal of Bacteriology 160, 462-465.

[307]   Risse, B., Stempfer, G., Rudolph, R., Möllering, H. and Jaenicke, R. (1992). Stability and reconstitution of pyruvate oxidase from *Lactobacillus plantarum*: Dissection of the stabilizing effects of coenzyme binding and subunit interaction. Protein Science 1, 1699-1709.

[308]   Muller, Y.A., Schumacher, G., Rudolph, R. and Schulz, G.E. (1994). The refined structure of a stabilized mutant and of wild-type pyruvate oxidase from *Lactobacillus plantarum*. J. Mol. Biol. 237, 315-335.

[309]   Lorquet, F., Goffin, P., Muscariello, L., Baudry, J.B., Ladero, V., Sacco, M., Kleerebezem, M. and Hols, P. (2004). Characterization and functional analysis of the poxB gene, which encodes pyruvate oxidase in *Lactobacillus plantarum*. J Bacteriol 186, 3749-59.

[310]   Elion, G. B. (1989). The purine path to chemotherapy. Science 244, 41-47.

[311]   Pui, C. H. and Jeha, S. (2007). New therapeutic strategies for the treatment of acute lymphoblastic leukaemia. Nat. Rev. Drug Discov. 6, 149-165.

[312]   Gentry, B.G., Gentry, S.N., Jackson, T.L., Zemlicka, J. and Drach, J.C. (2011). Phosphorylation of antiviral and endogenous nucleotides to di- and triphosphates by guanosine monophosphate kinase. Biochemical Pharmacology 81, 43-49.

[313]   Wernimont, A. and Edwards, A. (2009). *In Situ* Proteolysis to Generate Crystals for Structure Determination: An Update. PLoS ONE 4, e5094.

[314]   Delaglio, F. et al. (1995). NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J. Biomol. NMR 6, 277-293.

[315]   Singh, S.M. and Panda, A.K. (2005). Solubilization and refolding of bacterial inclusion body proteins. J Biosci Bioeng 99, 303-10.

[316]   Carrio, M.M. and Villaverde, A. (2001). Protein aggregation as bacterial inclusion bodies is reversible. FEBS Letters 489, 29-33.

[317]   Baret, J.C., Miller, O.J., Taly, V., Ryckelynck, M., El-Harrak, A., Frenz, L., Rick, C., Samuels, M.L., Hutchison,J.B., Agresti, J.J., Link, D.R., Weitz,D.A. and Griffiths, A.D. (2009). Fluorescence-activated droplet sorting (FADS): efficient microfluidic cell sorting based on enzymatic activity. Lab Chip 9, 1850–1858.

[318]   Lim, J., Gruner, P., Konrad, M. and Baret, J.C. (2013). Micro-optical lens array for fluorescence detection in droplet-based microfluidics. Lab Chip 13, 1472–1475.

[319]   Brow, T.D., Kakkar, V.V. and Das, S.K. (1999).The significance of creatine kinase in cardiac patients with acute limb ischaemia. J Cardiovasc Surg (Torino) 40, 637-44.

[320]   Igarashi, K., Hiraga, S. and Yura, T. (1967). A Deoxythymidine Kinase Deficient Mutant of *ESCHERICHIA COLI*. II. Mapping and Transduction Studies with Phage φ80. Genetics 57, 643-654.

CURRICULUM VITAE

NAZIMUDDIN KHAN

Date and Place of Birth: November 30, 1978, Charsadda, Khyber Pakhtunkhwa, Pakistan

EDUCATION

2010-2014   Ph.D. Biology (Biochemistry), Georg-August-Universität Göttingen, Germany

2001-2003   M.Sc. Chemistry-Biochemistry, Gomal University D.I. Khan, Pakistan

1998-2000   B.Sc. Chemistry, Zoology and Botany, University of Peshawar, Pakistan

1995-1997   F.Sc. Pre-medical, Chemistry, Physics and Biology, B.I.S.E., Peshawar Pakistan

1994-1995   S.S.C. Chemistry, Physics and Biology, B.I.S.E., Peshawar Pakistan

RESEARCH ACTIVITIES

2010-2014   Ph.D. research student, Enzyme Biochemistry Group, Max Planck Institute for Biophysical Chemistry, Am Fassberg11, 37077 Goettingen, Germany

2005-2009   Junior Research Fellow, International Center for Chemical and Biological Sciences, University of Karachi, Karachi 75270, Pakistan

HONORS AND STUDENTSHIP

2010-2014   DAAD Ph.D. Scholarship, Germany

2005-2009   Junior Research Fellow, ICCBS, University of Karachi, Pakistan

2005-2009   HEC Indigenous Fellowship, Pakistan

2008   GRE Biochemistry, Cell and Molecular Biology, ETS, USA