

Dissertation ETH No. 13456

**Structure elucidation of viral and human  
Thymidine kinases and Enolase  
from *Alternaria alternata***

A dissertation submitted to the  
**Swiss Federal Institute of Technology Zurich**  
for the degree of  
Doctor of Natural Sciences

presented by

**Andrea E. Prota**

Pharmacist (Eidg. dipl. Apotheker)  
ETH Zürich

born January 12th, 1969  
citizen of Spreitenbach, Switzerland  
&  
Villafranca in Lunigiana, Italy

accepted on the recommendation of

Prof. Dr. G. Folkers, examiner  
Dr. L. Scapozza, co-examiner  
Dr. O. Zerbe, co-examiner

**1999**

Per Carla,

e per mamma, Massimo e Luca

e in memoria a papà, nonna Elide e nonno Vittorio

Lass dir von den Spiegeleien  
Unsrer Physiker erzählen,  
Die am Phänomen sich freuen,  
Mehr sich mit Gedanken quälen.

Spiegel hüben, Spiegel drüben,  
Doppelstellung auserlesen;  
Und dazwischen ruht im Trüben  
Als Kristall das Erdewesen.

Dieses zeigt, wenn jene blicken,  
Allerschönste Farbenspiele;  
Dämmerlicht, das beide schicken,  
Offenbart sich dem Gefühle.

Schwarz wie Kreuze wirst du sehen,  
Pfaueaugen kann man finden;  
Tag und Abendlicht vergehen,  
Bis zusammen beide schwinden.

Und der Name wird ein Zeichen,  
Tief ist der Kristall durchdrungen:  
Aug' in Auge sieht dergleichen  
Wundersame Spiegelungen.

Lass den Makrokosmos gelten,  
Seine spenstischen Gestalten!  
Da die lieben kleinen Welten  
Wirklich Herrliches enthalten.

*(Goethe, Entoptische Farben)*

# TABLE OF CONTENTS

TABLE OF CONTENTS.....	I
ABBREVIATIONS.....	V
SUMMARY.....	1
ZUSAMMENFASSUNG.....	5
SOMMARIO.....	9
CHAPTER 1 .....	11
<b>Introduction</b>	
1.1 Thymidine Kinases .....	12
1.1.1 The pyrimidine salvage pathway.....	12
1.1.2 The short TKs (Type II).....	14
The human cytosolic TK (TK 1, fetal TK) .....	14
Mitochondrial TK (TK 2) .....	16
Cell line TK.....	16
1.1.3 TK and cancer .....	17
1.1.4 The long TKs (Type I).....	19
The structure of the HSV1 TK .....	19
1.1.5 TK and Therapies .....	24
Antiviral therapy .....	24
Gene therapy .....	24
1.2 Enolase .....	27
1.2.1 The sequence homology .....	27
1.3 Crystallography .....	29
1.3.1 A historical background of macromolecular Crystallization.....	31
1.3.1 Definition & principles of protein crystallization.....	33
1.3.2 Crystallization techniques .....	37
Vapor diffusion .....	37
Free interface diffusion.....	38
Batch Crystallization.....	39
Dialysis.....	39
1.3.3 Preparation of Crystals for the X-ray experiment.....	40
1.3.4 From data collection to solving the structure .....	41
1.3.5 Dynamic light scattering to predict crystallization? .....	43
The theory of light scattering .....	43
1.4 Aims and scope of the presented work .....	46
1.5 References .....	48

CHAPTER 2 .....	61
-----------------	----

### **The Human cytosolic thymidine kinase (TK1):**

Preparation & characterization of the protein for crystallization experiments

Abstract .....	62
2.1 Introduction.....	63
2.2 Cloning & Sequencing .....	65
2.2.1 The recombinant TK1 from human lymphocytes (Clone 34).....	65
2.2.2 The deletion mutant $\Delta$ 40TK1 .....	66
2.3 Expression & Purification.....	72
2.3.1 The human TK1 (Clone 34) .....	72
2.3.2 The His-Tag-TK1 (pQE32TK1) .....	73
2.3.3 The deletion mutant $\Delta$ 40TK1 .....	76
2.3.4 The <i>E.coli</i> BL21 codon plus-RIL strain.....	79
2.4 Enzymatic activity .....	83
2.4.1 UV-Spectrophotometric activity assay .....	83
2.4.2 HPLC assay for assessing substrate specificity.....	85
2.5 Homogeneity analysis .....	86
2.5.1 FPLC.....	86
2.5.2 Dynamic light scattering measurements .....	87
2.6 Discussion .....	88
2.7 References .....	91

CHAPTER 3 .....	95
-----------------	----

### **The crystallization of thymidine kinases and enolase: Rational or black magic?**

Preface .....	96
Abstract .....	96
3.1 Introduction.....	97
3.2 The rational approach.....	102
3.2 The crystallization screenings.....	103
3.4 Crystallization of artefacts .....	104
3.4.1 Izip <sup>®</sup> protein or salt?.....	106
3.5 The effect of detergents.....	107
3.6 Precipitating agents and their concentrations .....	108
3.7 The effect of light or vibration .....	110
3.8 Seeding .....	113



3.9 N-/ C-terminus or not? .....	116
3.9.1 The active site mutant Y101F .....	116
3.9.2 The triple mutant H58LM128FY172F.....	117
3.9.3 The deletion mutant $\Delta$ 40TK1 .....	120
3.10 Enolase from <i>Alternaria alternata</i> .....	122
3.10.1 Crystallization of Enolase from <i>Alternaria alternata</i> .....	123
3.11 Discussion.....	125
Grid screening .....	125
The importance of purity .....	126
Detergents .....	126
Crystal packing .....	127
3.12 Concluding remarks .....	127
3.13 References .....	128
CHAPTER 4 .....	133

**The nucleoside binding site of *Herpes simplex* type 1 TK  
studied by X-ray crystallography**

ABSTRACT .....	135
INTRODUCTION.....	136
MATERIALS & METHODS.....	137
Materials .....	137
Preparation of the expression vectors .....	137
Protein expression and purification.....	138
Crystallization .....	138
X-ray data collection .....	139
Refinement .....	139
RESULTS AND DISCUSSION .....	141
High resolution structure of substrate-free TK.....	141
LID-domains of substrate-free TK <sub>HSV1</sub> .....	141
Superposition TK-AK.....	143
The role of water for binding of substrate .....	144
The structure of TK <sub>HSV1</sub> :9-HPA.....	146
The structure of TK <sub>HSV1-Q125N</sub> :dT .....	149
CONCLUSIONS .....	152
ACKNOWLEDGEMENT .....	152
REFERENCES.....	152

---

CHAPTER 5 .....	157
<b>Kinetics and high resolution structure of HSV1 TK and the engineered Y101F mutant in complex with an antiviral drug with conformationally restricted sugar ring pucker</b>	
ABSTRACT .....	159
INTRODUCTION .....	160
EXPERIMENTAL PROCEDURES.....	162
Materials .....	162
Protein Engineering .....	162
Sequence verification .....	163
Protein expression and purification.....	163
Assessment of phosphorylation pattern.....	164
Inhibition kinetics .....	164
Crystallization .....	164
X-ray data collection .....	165
Refinement .....	165
RESULTS.....	167
Protein engineering.....	167
Assessment of the phosphorylation pattern.....	167
Kinetics .....	168
High resolution structure of TK <sub>HSV1:(N)</sub> -MCT.....	169
Mutant Y101F .....	171
DISCUSSION .....	172
CONCLUSIONS .....	174
ACKNOWLEDGEMENT .....	174
REFERENCES.....	175
PUBLICATIONS AND TALKS.....	181
CURRICULUM VITAE.....	183
Thank you, grazie.....	184

## ABBREVIATIONS

ACV	acyclovir, 9-(2-hydroxyethoxymethyl)guanine
ACV <sup>r</sup>	acyclovir resistant
ADK	adenylate kinase
AMP	adenosine 5'-monophosphate
ADP	adenosine 5'-diphosphate
ATP	adenosine 5'-triphosphate
AV	adenovirus
AZT	3'-azido-3'-deoxythymidine
BHV	bovine herpes virus
BMT	bone marrow transplantation
CD	( <i>E.coli</i> ) cytosine deaminase
C	cytidine
dC	2'-deoxycytidine
dCK	deoxycytidine kinase
dCMP	deoxycytidine 5'-phosphate
dCDP	deoxycytidine 5'-diphosphate
dCTP	deoxycytidine 5'-triphosphate
CMV	cytomegalovirus
Δ40TK1	deletion mutant of the human cytosolic TK1 lacking the 40 C-terminal amino acids
DEAE	diethylaminoethyl
DNA	deoxyribonucleic acid
EBV	Epstein Barr Virus

---

<i>E. coli</i>	<i>Escherichia coli</i>
EHV	equine herpes virus
FHV	feline herpes virus
G	guanosine
dG	2'-deoxyguanosine
GCV (=DHPG)	ganciclovir, 9-(1,3-dihydroxy-2-propoxymethyl)guanine
dGMP	deoxyguanosine 5'-monophosphate
dGDP	deoxyguanosine 5'-diphosphate
dGTP	deoxyguanosine 5'-triphosphate
GST	glutathione S-transferase
GST-TK	fusion protein of glutathione S-transferase and herpesviral thymidine kinase
GvHD	graft versus host disease
HBPG	hydroxybutyl-N <sup>2</sup> -phenylguanine
HIV	human immune deficiency virus
HPLC	high performance liquid chromatography
HSV	herpes simplex virus
HVS	herpesvirus saimiri
IPTG	isopropyl- $\beta$ -D-thiogalactopyranoside
ITC	isothermal calorimetry
K <sub>i</sub>	inhibition constant
K <sub>m</sub>	Michaelis-Menten constant
LDH	lactate dehydrogenase
MHV	marmoset herpesvirus
MTX	methotrexate

---

NADH	nicotinamide-adenine dinucleotide, reduced form
dNTP	deoxynucleoside triphosphate
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PCV	penciclovir
PK	pyruvate kinase
PEP	phosphoenolpyruvate
PET	positron emission tomography
Pol	DNA polymerase
PRV	pseudorabies virus
rec	recombinant
RNA	ribonucleic acid
SDS	sodium dodecyl sulfate
dT	deoxythymidine, 2'-deoxythymidine
dTMP	thymidine 5'-phosphate, thymidylate
dTDP	thymidine 5'-diphosphate
dTTP	thymidine 5'-triphosphate
TmpK	thymidylate kinase
TK	thymidine kinase
TK <sup>-</sup>	TK negative
<i>tk</i>	gene encoding for TK
TK 1	human cytosolic thymidine kinase
TK 2	human mitochondrial thymidine kinase
TLC	thin layer chromatography
Tris	tris(hydroxymethyl)aminomethane
U	uridine

dU	2'-deoxyuridine
dUMP	deoxyuridine 5'-monophosphate
dUDP	deoxyuridine 5'-diphosphate
dUTP	deoxyuridine 5'-triphosphate
$V_{\max}$	rate of enzyme catalyzed reaction at infinite concentration of substrate
VV	Vaccinia Virus
VZV	Varicella Zoster Virus

The one-letter code is used for amino acids

## SUMMARY

Thymidine kinase (TK) catalyzes via the so-called thymidine salvage pathway the phosphorylation of thymidine (dT) to thymidine monophosphate in the presence of adenosine triphosphate (ATP) and magnesium ions. Based on their physico-chemical characteristics TKs are subdivided into two groups, the type I and type II TKs. A representative of the type I TK is *herpes simplex virus type 1* (HSV 1) TK that exhibits a broad substrate diversity and phosphorylates not only pyrimidine analogs (e.g. BVdU and IdU) but also guanine and adenine analogs. Furthermore, it accepts bases carrying acyclic or carbocyclic sugar moiety substituents. In contrast, the human cellular TK, a representative of the type II TK, exhibits a rather high substrate specificity and accepts only pyrimidine substrate analogs and restricted changes in the sugar moiety. The difference in substrate acceptance is crucial for therapeutic applications linked to TK. TK established itself as a major tool in many applications such as gene therapeutical treatment of cancer and AIDS, as a control system in allogeneic bone marrow transplantation and AIDS vaccine, as an expression reporter gene and is an important target in antiviral therapy.

The main aspect of this work is the exploration of the differences in substrate specificity between the viral and the host cell TKs at a structural level by X-ray crystallography. These new structural details would provide us with a base for rational design of new tailor-made substrates and mutants to improve the efficacy of antiviral therapy and anticancer gene therapy.

The human cytosolic TK1 was expressed in full-length or the C-terminally truncated mutant ( $\Delta 40$ TK1) in *E.coli* BL21 as the thrombin- or PreScission-cleavable glutathion-S-transferase-fusion protein, as well as in *E.coli* XL-1 blue as the his-tag-protein. With the exception of the deletion mutant, all other forms of the enzyme showed sequence heterogeneity, detected as double bands by SDS-PAGE, coming from the heterogeneous expression in *E.coli*. These protein forms were characterized by their tendency to aggregate thereby preventing crystallization. These issues were successfully addressed by the creation of a fully active deletion mutant, which could be purified by affinity chromatography.

The active site mutants Q125N, Y101F and H58LM128FY172F of HSV 1 TK were also expressed in *E.coli* BL21 as thrombin-cleavable glutathion-S-transferase-fusion proteins. They were purified to homogeneity in a single step procedure by affinity chromatography. For activity screenings a coupled UV-spectrophotometric assay involving pyruvate kinase and lactate dehydrogenase was applied. To quantify substrate conversion an established HPLC assay was used. The monodispersity of the proteins was analyzed either by FPLC or by dynamic light scattering (DLS).

The first target of crystallization was human TK1, which finally crystallized as the engineered fully-active deletion mutant. That allowed to prevent precipitation linked to instability or inhomogeneity. The crystals obtained did not diffract X-rays due to their small size. Attempts to increase their size were not successful. Nevertheless, the crystallization experiments with this enzyme provided us with notable experience in this field that was successfully applied in further crystallization experiments.

One of these experiments represented also the second target of this work and consisted in the elucidation of several active-site mutants of HSV 1 TK. Two of them were crystallized under conditions similar to those used for the wild type enzyme. Q125N was crystallized in complex with the natural substrate thymidine, while Y101F and the wild-type enzyme formed complexes with (N)-methanocarba-thymidine, an antiviral drug with a conformationally restricted pucker of the sugar ring. The structures of the two HSV1 TK mutants were solved to 2.4Å resolution whereas the wild-type structure was refined to 1.7Å resolution. All crystallographic data were collected under cryogenic conditions. The final models were built using molecular replacement and provided us fundamental insights into the role of water and of different amino acid residues during catalysis. Furthermore, we were able to gain a better understanding of the broad substrate acceptance of HSV 1 TK and of the effect of a conformationally restricted sugar-like moiety.

A further aspect of this work was the necessity to gain structural knowledge of the enolase from *Alternaria alternata* in order to identify the enzyme's epitope responsible for the allergic reaction with the IgE. This enzyme shares a high degree of sequence identity with known enolases from yeast and other not related



fungi, and was reported to be a major trigger of mould allergies of the respiratory tract. The enolase from *Alternaria alternata* and enolase extracts from yeast have shown immunological cross-reactivity with patients' IgE. Interestingly, the enolase from yeast was not reported to cause allergies. New insights into the structural details of this enzymes will broaden our knowledge in understanding fungal allergies.

The recombinant enolase from *Alternaria alternata* was available for us as lyophilized enzyme. Crystallization screenings were set up using standard methods. The enzyme could be crystallized, but could unfortunately not be grown to a size amendable to single-crystal diffraction.

In conclusion, the presented work was dedicated to crystallization of TK and enolase and allowed the elucidation of four new structures providing useful information for the rational design of resistance repellent drugs for antiviral therapy and for the development of new tailor made mutants and substrates for gene therapy.

Seite Leer /  
Blank leaf

## ZUSAMMENFASSUNG

Thymidinkinasen (TK) katalysieren die Phosphorylierung von Thymidin (dT) zu Thymidinmonophosphat über den sogenannten „*salvage pathway*“ unter Dephosphorylierung von Adenosintriphosphat (ATP) in Anwesenheit von Magnesiumionen. Die Thymidinkinasen sind aufgrund ihrer physiko-chemischen Eigenschaften in zwei Gruppen eingeteilt: die Thymidinkinasen vom Typ I und vom Typ II. Ein typischer Vertreter der Typ I TKs ist die *Herpes Simplex Virus Typ 1* Thymidinkinase (HSV 1 TK). Sie besitzt die Eigenschaft, nicht nur das natürliche Substrat Thymidin, sondern auch eine ganze Reihe weiterer Substrate wie zum Beispiel Guanosin- und Adenosinanaloga zu phosphorylieren. Im Gegensatz dazu besitzt die humane zytosolische Thymidinkinase, ein Vertreter der Typ II TKs, ein eher beschränktes Substratspektrum und akzeptiert nur Pyrimidinanaloga und minimale Variationen der Zuckereinheit. Dieser Unterschied bezüglich der Substratspezifität bildet den Grundstein für verschiedene Therapieansätze in der antiviralen und der Gentherapie von Tumoren.

Der Hauptaspekt dieser Arbeit ist die Erforschung der Unterschiede in der Substratspezifität zwischen viralen und zellulären Thymidinkinasen auf Strukturebene mittels Röntgenkristallographie. Diese neuen strukturellen Details liefern uns wichtige Informationen für die rationale Entwicklung neuer massgeschneiderter Substrate und Enzyme mit dem klaren Ziel, den Erfolg der existierenden therapeutischen Ansätze in der antiviralen und der Krebstherapie zu verbessern.

Die humane Thymidinkinase wurde sowohl als ganzes, als auch C-terminal verkürztes Enzym ( $\Delta 40\text{TK1}$ ) in E.-Colibakterien als Fusionsprotein exprimiert, das mit Thrombin oder einer PreScission Protease spaltbar ist. Zusätzlich wurde die humane TK1 auch als His-tag-Protein exprimiert. Alle erzeugten Enzyme, bis auf die verkürzte Form, zeigten Sequenz-Heterogenität, die mittels SDS-Gelelektrophorese in Form von Doppelbanden detektiert werden konnte. Diese Inhomogenität war auf die heterogene Expression in E.-Colibakterien zurückzuführen. Diese Proteine neigten sehr stark zur Bildung unterschiedlicher Aggregate und stellten somit ein Hindernis für die Kristallisation dar. Durch die

Konstruktion des C-terminal verkürzten Enzymes wurde dieses Hindernis überwunden und das Protein konnte via Affinitätschromatographie gereinigt werden.

Die HSV 1 TK-Mutanten Q125N, Y101F und H58LM128FY172F wurden auch in *E.-Colibakterien* als Thrombin-spaltbare Fusionsproteine exprimiert und via Affinitätschromatographie gereinigt. Die Aktivität der gereinigten Enzyme wurde UV-spektrophotometrisch mittels einer Enzym-gekoppelten Reaktion mit Pyruvatkinase und Laktatdehydrogenase getestet. Die Charakterisierung der Substratumsetzung wurde mittels HPLC durchgeführt. Die Monodispersität des Proteins wurde entweder mittels FPLC oder *Dynamic Light Scattering (DLS)* analysiert.

Das erste Zielenzym der Kristallisation war die humane Thymidinkinase, die nach Beseitigung der oben erwähnten Hindernisse wie Aggregation und Inhomogenität als C-terminal verkürztes Enzym kristallisiert werden konnte. Die Kristalle streuten allerdings nicht ausreichend gut im Röntgenstrahl, was auf ihre kleine Grösse zurückzuführen war. Alle Versuche, die gemacht wurden, um die Kristalle zu vergrössern, schlugen fehl. Trotzdem wurden mit den durchgeführten Experimenten viel Erfahrungen gesammelt, die dann bei den weiteren Kristallisationsexperimenten positiv umgesetzt werden konnten.

Eines dieser weiteren Experimente war die Kristallisation der HSV 1 TK-Mutanten. Zwei davon wurden erfolgreich kristallisiert. Die Mutante Q125N wurde als Komplex mit dem natürlich vorkommenden Substrat Thymidin kristallisiert, und die Mutante Y101F und die wild-Typ TK als Komplex mit (N)-Methanocarpa-thymidin, einem antiviralen Stoff mit einem in seiner Konformation fixierten Zuckerring. Die Strukturen der Mutanten wurden bei 2.4Å, die der wild-Typ TK bei 1.7Å gelöst. Alle kristallographischen Daten wurden unter Cryo-Bedingungen gesammelt. Die Modelle wurden unter Verwendung der Methode des „*molecular replacement*“ gebaut. Diese Strukturen lieferten uns fundamentale Erkenntnisse über die Funktion von Wasser und bestimmten Aminosäuren während der katalytischen Reaktion im aktiven Zentrum des Enzyms und führten zu einem besseren Verständnis der breiten Substratspezifität der HSV 1 TK auch in bezug auf den Effekt von rigidisierten Zuckereinheiten.

Ein weiterer Aspekt dieser Arbeit ist die Erlangung struktureller Informationen über die Enolase von *Alternaria alternata*. Hiermit wird es möglich, die Epitope des Enzyms zu charakterisieren, die für die allergische Reaktion mit IgE verantwortlich sind. Dieses Enzym zeigt hohe Sequenzidentität mit anderen Enolasen aus Hefen und anderen nicht verwandten Pilzen. Sie wurde als eine der wichtigsten Auslöser für Allergien des Respirationstraktes erkannt. Versuche mit dieser Enolase und der Enolase aus Hefeextrakten haben Kreuzreaktivität mit IgE von sensibilisierten Patienten gezeigt. Interessanterweise zeigt die Enolase von Hefe keine allergischen Eigenschaften. Neue Einsichten in strukturelle Details dieser Enzyme werden unsere Kenntnis über die Entstehung solcher Allergien erweitern.

Die rekombinante Enolase aus *Alternaria alternata* war als Lyophilisat erhältlich. Kristallisations-Experimente wurden unter Verwendung von Standardmethoden durchgeführt. Das Enzym wurde zwar kristallisiert, konnte aber nicht in einer Grösse erhalten werden, die für die Messung eines X-Ray Datensatzes geeignet wäre.

Schlussfolgernd erlaubt diese Arbeit, die der Kristallisation von Thymidinkinasen und der Enolase gewidmet war, die Diskussion von vier neuen Strukturen, die uns nützliche Informationen geliefert haben für die rationale Entwicklung von neuen antiviralen Arzneistoffen. Diese sind auch gegen resistente Viren wirksam. Ferner erhielten wir hilfreiche Erkenntnisse für die Entwicklung von massgeschneiderten Enzymen und Substraten für den Einsatz in der Gentherapie.

Seite Leer /  
Blank leaf

## SOMMARIO

Le timidinchinasi catalizzano la monofosforizzazione della timidina per via del cosiddetto “*salvage pathway*” usando adenosinatrifosfato come co-substrato in presenza di magnesio. Le timidinchinasi sono suddivise in due gruppi in base alle loro proprietà fisico-chimiche: Tipo I e tipo II. Un tipico enzima a rappresentare il tipo I è la timidinchinasi dell’ Erpes simplex virus tipo I, che è capace non solo di fosforizzare la timidina, ma anche una quantità di sostanze analoghe alla guanosina e alla adenosina. Contrariamente la timidinchinasi umana, che rappresenta il tipo II, ha uno spettro molto ristretto essendo capace di fosforizzare solamente sostanze analoghe alla pirimidina, accettando anche delle lievi modificazioni allo zucchero. Questa differenza è fondamentale e viene sfruttata sia in diversi principi di terapia antivirale sia nella terapia genetica contro il cancro.

L’aspetto principale di questo lavoro è l’esplorazione delle differenze in specificità per i substrati tra le timidinchinasi virali e cellulari a livello strutturale usando la cristallografia ai raggi X. La conoscenza di questi dettagli strutturali sarà fondamentale nello sviluppo di nuovi enzimi e substrati per migliorare le terapie antivirali e quelle contro il cancro.

La timidinchinasi umana è stata prodotta in *E.coli* sia in fusione con la glutatione-S-transferasi, sia con un *His-tag*. La proteina di fusione è poi stata rimossa per azione della trombina o di una *PreScission* proteasi. L’inomogeneità e la tendenza ad aggregarsi impedivano così la corretta cristallizzazione. Questi problemi sono stati risolti creando una timidinchinasi modificata alla parte c-terminale. In questo modo, abbiamo ottenuto un enzima omogeneo in seguito a purificazione cromatografica di affinità. Anche i mutanti Q125N, Y101F und H58LM128FY172F della timidinchinasi virale sono stati prodotti usando il sistema di espressione in *E.coli* e purificati con cromatografia di affinità. L’attività degli enzimi è stata misurata con la spettrofotometria UV usando una reazione combinata con un differente enzima. I prodotti dell’enzima sono stati caratterizzati attraverso HPLC. L’omogeneità invece è stata analizzata con FPLC e *Dynamic light scattering*.

Il primo enzima ad essere stato sperimentato nella cristallizzazione è la timidinchinasi umana. Dopo aver eliminato le inconvenienze dell’inomogeneità e dell’aggregazione, l’enzima modificato è stato cristallizzato e analizzato ai raggi X.

Date le piccole dimensioni dei cristalli, non è stato possibile registrare delle riflessioni. Tutti i tentativi fatti per incrementare le dimensioni dei cristalli sono falliti. Ugualmente, questi esperimenti hanno portato molte esperienze che hanno potuto essere usate per cristallizzare i mutanti della timidinchinasi virale.

Gli altri enzimi da analizzare nella cristallizzazione sono i mutanti della timidinchinasi virale. Il mutante Q125N è stato cristallizzato in complesso con la timidina, mentre il mutante Y101F e il "wild type" virale in complesso con la (N)-Metanocarpa-timidina. Quest'ultima rappresenta una nuova sostanza antivirale contenente uno zucchero fissato nella sua conformazione. Le strutture dei mutanti sono state risolte a risoluzione di 2.4 Å, quella del "wild type" a 1.7 Å usando la tecnica della sostituzione molecolare. Queste strutture hanno dato delle informazioni fondamentali sul ruolo dell'acqua e di certi amminoacidi nel centro attivo dell'enzima durante la reazione catalitica, e hanno aiutato a capire meglio la vasta specificità di diverse sostanze da parte dell'enzima virale.

Un altro aspetto di questo lavoro è la necessità di avere informazioni strutturali sull'enolasi di *Alternaria alternata*, una muffa che è stata riconosciuta come maggior causa di allergie delle vie respiratorie. Questo enzima dimostra un'alta percentuale di identità rispetto ad altre enolasi di lievito ed altri funghi. Gli esperimenti condotti con l'enolasi di *Alternaria alternata* e l'enolasi in estratti di lievito hanno dimostrato una certa reattività incrociata con le IgE di pazienti sensibilizzati, mentre l'enolasi di lievito non causa allergie. La conoscenza della struttura tridimensionale di questo enzima potrebbe aiutare la comprensione di queste differenze e capire lo sviluppo delle allergie ai funghi.

L'enzima enolasi di *Alternaria alternata* è stato utilizzato in forma liofilizzata. Gli esperimenti di cristallizzazione sono stati condotti usando i metodi standard. L'enzima è stato cristallizzato, ma la piccola dimensione dei cristalli non ha consentito la misurazione ai raggi X.

Concludendo, questo lavoro, che è stato dedicato alla cristallizzazione di diverse timidinchinasi e dell'enolasi, ha permesso di elucidare quattro nuove strutture, che hanno fornito delle informazioni utili che potranno essere utilizzate nello sviluppo di nuovi farmaci antivirali e nello sviluppo di nuovi enzimi e farmaci per terapia genetica contro il cancro.



# CHAPTER 1

## Introduction

## 1.1 Thymidine Kinases

Thymidine kinases were described for cellular life long before it was shown that they were also be encoded by viruses. Nonetheless, the viral thymidine kinase genes were the first to be sequenced. These enzymes have been extraordinarily useful to the researcher, serving first to help label DNA, then to get thymidine analogs incorporated into DNA for therapeutic and other purposes and more recently to move genes from one genome to another. Knowledge of the nucleotide and amino acid sequences of these enzymes has allowed some deductions about their possible three-dimensional structure, as well as the location on the polypeptide of various functions until the first crystal structure of HSV1 TK, to date still the only one, has been solved in 1995<sup>1,2</sup>. The knowledge of the sequences has also allowed their classification into two main groups: the poxviral and cellular thymidine kinases (Type II, short) and the herpesviral thymidine kinases (Type I, long)<sup>3</sup>.

In the following paragraph a review of the different type of TK and their characteristics as well as their link to therapy is presented.

### 1.1.1 The pyrimidine salvage pathway

Thymidine kinases (TK, EC 2.7.1.21) catalyze the phosphorylation of thymidine (dT) to thymidine monophosphate (dTMP) in the presence of ATP and  $Mg^{2+}$  (figure 1). The  $\gamma$ -phosphate group of ATP is transferred to the 5'-OH group of thymidine. Thus, the systematically correct term for TK is ATP-thymidine-5'-phospho-transferase. dTMP is subsequently phosphorylated by cellular thymidylate kinase to dTDP and by nucleoside diphosphate kinases into the triphosphate (dTTP), which is then used as DNA building block (Fig.1). However, the cell does not depend on TK for the synthesis of dTMP. Thymidylate synthetase, the key enzyme in the *de-novo* synthesis pathway, catalyzes the methylation of deoxyuridine monophosphate (dUMP) to dTMP in presence of methylene tetrahydrofolate. The TK is therefore designated as a salvage pathway

enzyme, recycling endogenous thymidine released from metabolic decomposition or thymidine resorbed from nutrition (figure 1).

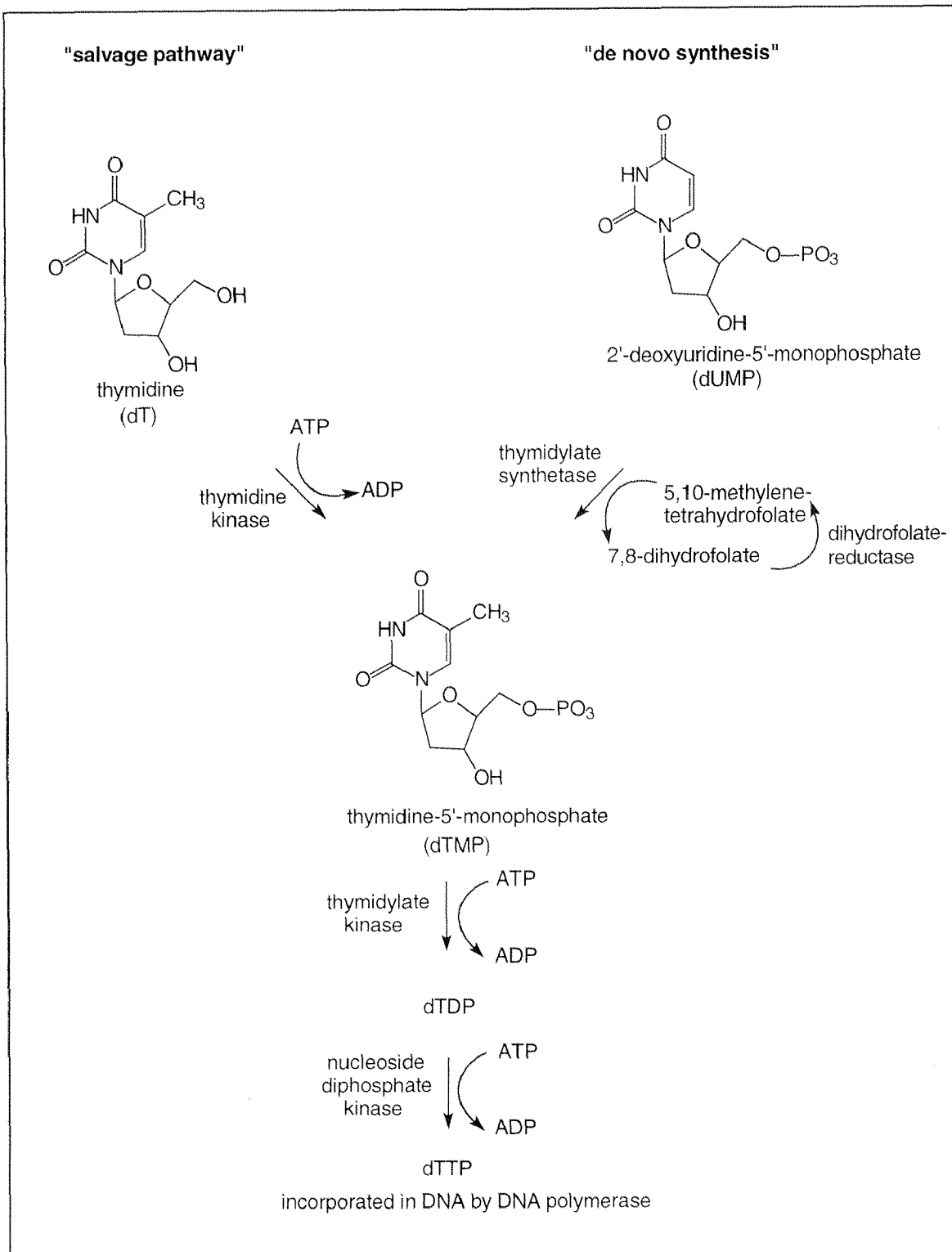


FIG. 1: Biosynthesis pathways of thymidine triphosphate

### 1.1.2 The short TKs (Type II)

TK is a highly conserved enzyme, which occurs in nearly all organisms. The group containing type II TK's, which is quite distinct from HSV-like enzymes, is comprised of thymidine kinases similar to that of VV, namely those expressed by monkeypox, variola, fowlpox, capripox, mouse, chicken, human, bacteriophage T4 and *E. coli*. Initial studies of the type II prototype TK enzyme encoded by VV suggested that it was also a dimer of 40 kDa subunits. However, when the VV TK gene was mapped and sequenced, the nucleotide sequence predicted a protein with molecular weight of 20 kDa. Since, like the type I enzymes, the functional VV TK enzyme had an apparent native molecular weight of 80 kDa, this suggested that type II enzymes must function as tetrameric complex<sup>4</sup>.

In the further text, the human enzyme will be discussed. Based on biochemical and/or electrophoretic results, three different isoenzymes of TK have been identified in human cells.

#### ***The human cytosolic TK (TK 1, fetal TK)***

The human cytosolic thymidine kinase consists of 234 aa with a subunit molecular weight of 25,5 kDa<sup>5</sup>. The native enzyme has been reported to be a tetramer in the presence of ATP and a dimer in absence of ATP. In contrast, the recombinant enzyme was found to be tetrameric independently on the presence of ATP<sup>6,7</sup>. pI values of the cytosolic TK1 between 6.3 and 8.3 have been reported to consist of different phosphorylated forms of the enzyme<sup>8</sup>.

The regulation of the human cytosolic TK is in accordance with its physiological role. High concentrations of cytosolic TK are only present in tissues with an elevated replication rate. Quiescent serum-starved or differentiated cells have low levels of TK activity, TK protein and TK mRNA. Of necessity therefore elevation of TK on growth stimulation of cells requires induction of transcription. Many studies of the past 10 years have shown that transcription of the TK gene starts in the late G<sub>1</sub> phase<sup>9-11</sup>. Although different cells exhibit very similar overall regulation of TK gene expression, promoters of the TK gene differ remarkably in different organisms. Some of these (the human and the hamster promoter) carry TATA boxes, but the others are lacking these elements. All promoters have GC boxes,

binding sites for transcription factor Sp1. The murine TK promoter furthermore has a genuine binding motif for the cell-cycle regulated transcription factor E2F, and the other promoters have sequences that more (human) or less (hamster, rat) resemble E2F sites<sup>12</sup>. The three to six fold induction of transcription was shown to be upregulated by specific protein complexes involving both transcription factors E2F and Sp1 and an additional protein DP1 at the G<sub>1</sub>/S phase border of the cell cycle<sup>13,14</sup>. The transcription has been shown to be turned off at about mid S phase by a cyclinA/cdk2 dependent phosphorylation of the two transcription factors E2F and DP1 inhibiting the further formation of the ternary complex with Sp1<sup>15</sup>. DNA tumor virus proteins were reported to interfere with this regulation causing a transactivation of the promoter<sup>13</sup>.

Furthermore, TK mRNA expression seems to be dependent on growth-specific events at different post-transcriptional levels<sup>9,11,16</sup>. Kauffman and coworkers (1991) have shown that the expression of human TK protein in resting cells is regulated by a post-transcriptional mechanism and that residues in the carboxy terminus are required for this regulation. In addition it was demonstrated that the low TK protein levels in the G<sub>1</sub> phase of the cell cycle and during differentiation are the result of repressed TK protein synthesis<sup>17-19</sup>. Reduced protein stability seems to be responsible for the disappearance of the TK protein during or shortly after mitosis and at least in part for the low levels of TK polypeptide in growth-arrested cells. The carboxy terminal residues have also been shown to be essential for the specific degradation while they do not alter the enzymatic activity significantly when abolished<sup>20-22</sup>. More recently, it has been shown, that thymidine itself exerts the stabilizing effect on the TK protein<sup>23</sup>.

### ***Mitochondrial TK (TK 2)***

This TK species is localized in the mitochondrial matrix and shows constant but low levels of TK activity during the entire cell cycle<sup>24</sup>. It is still not clear if it belongs to type I or type II. Sequence alignment of the known human deoxynucleoside kinases show 40% identity of TK2 to deoxycytidine kinase and mitochondrial deoxyguanosine kinase, while the identity to the cytosolic TK1 is very low<sup>25</sup>. TK2 is synthesized outside the mitochondria and subsequently translocated, most likely guided by a signaling peptide, which is cleaved off after having reached the mitochondrial matrix. The TK 2 is not involved in cell growth and is expressed in all tissues proportional to the mitochondrial content of the cell type. In many resting cells, such as nerve and muscle cells, TK2 is the only pyrimidine deoxynucleoside phosphorylating enzyme expressed. In these tissues, where de novo synthesis of DNA precursors is undetectable, TK 2 is likely to be responsible for the supply of deoxynucleotides required for the mitochondrial DNA synthesis, since, unlike dT, the negatively charged dTMP exhibits an extremely poor penetration of the mitochondrial membrane<sup>26</sup>. TK 2 is able to phosphorylate pyrimidine deoxynucleosides and analogs such as dT, dC, dU and AZT, as well as pyrimidine arabinosides and their analogs like AraC, AraT and fluoroarabinofuralsyl-iodouracil (FAIU)<sup>25,27</sup>. Since the capacity of resting cells to phosphorylate deoxynucleosides relies on TK 2, this enzyme is suspected to be a reason for the side effects of these drugs<sup>28</sup>. In man, the gene encoding for cytosolic TK is localized on chromosome 17 whereas the gene for the mitochondrial TK is localized on chromosome 16<sup>29</sup>.

### ***Cell line TK***

Until now, this TK variant has only been traced on certain cell lines. It probably originates from the inner mitochondrial membrane<sup>30</sup>.

### 1.1.3 TK and cancer

TK1 is one of the proteins whose abundance is most closely correlated to the proportion of S-phase cells in a tissue. According to this role, and the fact that the aggressiveness of a number of human tumors is more directly related to the proportion of S-phase and G<sub>2</sub> cells, TK1 expression both at the mRNA and protein level has been used extensively as a marker for proliferating cells. TK1 activity and proliferation have been found in excised tumor tissues, in plasma or serum samples or in mononuclear leukocytes from patients with chronic lymphatic leukemia, Hodgkin's disease and other malignancies<sup>31,32</sup>. In breast cancer patients, the serum TK concentrations correlate with the stage of the disease, the estrogen receptor status and the recurrence of the disease<sup>33</sup> and was shown to be increased in breast cancer patient compared to the healthy patients or patients with benign breast diseases and to correlate significantly with the stage of disease<sup>34,35</sup>. The activity of TK has been correlated with many different malignancies, including human mammary neoplasias, where a 14-fold increase of TK activity was demonstrated in human mammary adenocarcinomas when compared to normal mammalian tissue<sup>36</sup>. Analysis of TK activity has been shown to be a useful predictive indicator in several forms of cancer, including breast cancer, when used in conjunction with other clinical, pathological and serological factors like progesterone receptor<sup>35</sup>. Studies on lung cancer have shown a 4.3 higher TK1 activity than in uninvolved lungs. No significant correlation between the activity in tumors and clinicopathological findings such as histopathological type, grade of differentiation and tumor size was found<sup>37</sup>. The overexpression of thymidine kinase in cervical carcinoma cells was recently reported to correlate with the expression of folate receptors, leading to induction of sensitivity of these cells to AZT<sup>38</sup>. A polyclonal antibody against the C-terminal part of TK1 has been reported to be the best immunological tool for biochemical and cellular determinations of TK1, allowing to provide prognostic information on survival and to predict response to induction of chemotherapy<sup>39</sup>.

An indirect link between TK and cancer is given by the product of the reaction dTMP. Several anticancer drugs block the synthesis of dTMP inhibiting either dihydrofolate reductase or thymidylate synthetase (see figure 1). Fluorouracil is for example converted *in vivo* to F-dUMP. This dUMP analog irreversibly inhibits thymidylate synthetase by blocking the catalysis at a stage where F-dUMP is covalently bound to the enzyme. On the other hand, methotrexate (MTX), a dihydrofolate analog, is a potent competitive inhibitor of dihydrofolate reductase. However, these drugs are rather toxic because of the low selectivity towards malignant cells.

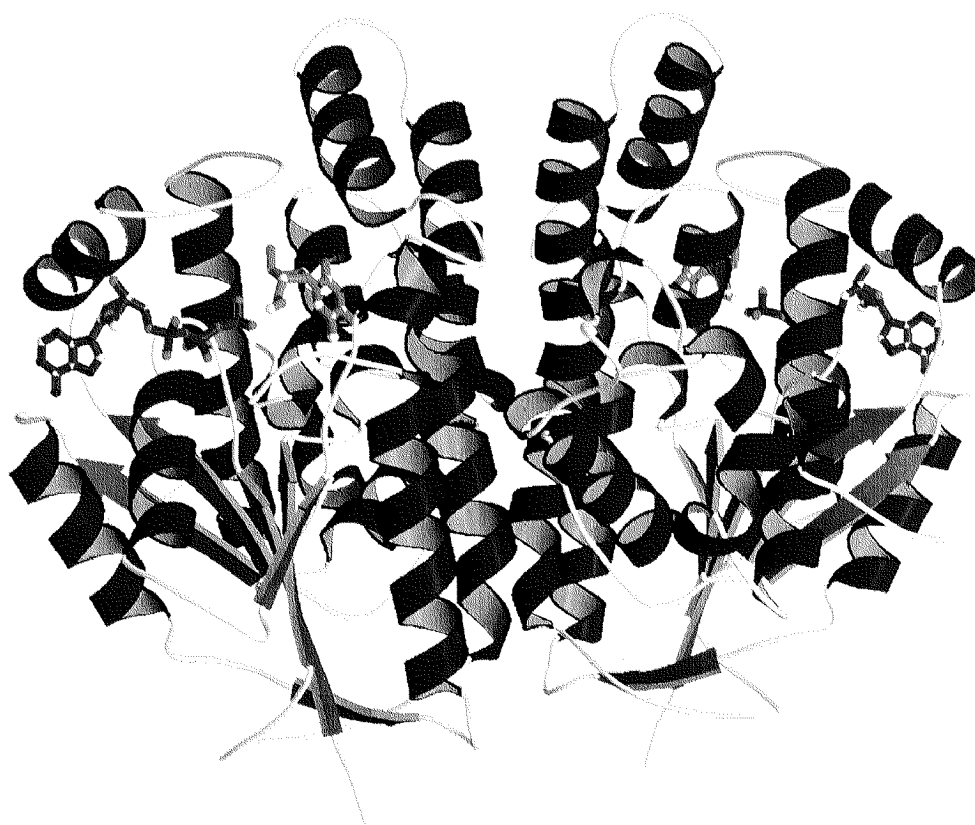
The regulation of TK is also affected by drugs that are inhibitors of the dTMP *de novo* synthesis. Therefore, the dTMP salvage pathway, involving TK, is upregulated and the cell metabolism can proceed restoring the dTMP pool. Patients treated with these anticancer drugs show, as expected, pathologically elevated serum TK levels. Hence, the salvage pathway for pyrimidine nucleotide biosynthesis decreases the therapeutic effect of agents that act by inhibiting the *de novo* synthesis as disclosed by the co-expression of HSV *tk* gene in cells that potentates MTX resistance<sup>40</sup>. Optimized therapy and dose reduction of these drugs could be achieved, if TK would be likewise inhibited.



### 1.1.4 The long TKs (Type I)

#### *The structure of the HSV1 TK*

The enzyme is present as homodimer with 376 residues per subunit related by C2 symmetry as indicated in figure 2. The recombinant TK has been crystallized as full length and truncated enzyme.



**Fig. 2:** Molscript plot of the X-ray structure of the HSV 1 TK in complex with dT and ATP solved to 2.7Å resolution. The helices are depicted in blue, the  $\beta$ -strands in light blue, while turns and coils in yellow. The substrate is in red and the co-substrate in green.

The first structure was the truncated enzyme comprising residues 34-376 with ADP, dTMP and 52 water molecules<sup>2,41</sup>. Amino acids 34-45, 150-152 and 265-279 were missing likewise in other, independently solved structures in complex with dT, GCV and further substrate analogs<sup>1,42,43</sup>. The so far highest resolution structure of HSV1 TK has been solved to 1.7Å in complex with N-MCT (this work). Nevertheless, the N-terminus is still missing in the structure. On the one hand,

residues 1-45 are not essential to explain the reaction cycle as shown by the observation that this segment is not required for catalysis<sup>44</sup>. On the other hand, a nuclear localization signal (NLS)<sup>45</sup> was detected within this N-terminal arginine-rich region of TK<sup>46</sup> pointing out the Arg 25, 26, 30, 32 and 33 to be necessary for efficient NLS functioning.

The structure of TK (PDB codes 1VTK, 2VTK and 3VTK) consists of an  $\alpha/\beta$  structure made up of 12  $\alpha$ -helices and seven  $\beta$ -strands. The five-stranded parallel  $\beta$ -sheet, which exhibits a high degree of homology with ADK, forms part of the core of the molecule, which contains the active site. The large and almost circular, planar interface of the dimer is mainly formed by three helices,  $\alpha_4$ ,  $\alpha_6$  and  $\alpha_{10}$  with only one substantial protrusion formed by Trp 310 anchoring in a hollow of the other subunit. The interface is exceptionally non-polar and despite its largeness, the crystal contacts bury only 7 % of the total dimer interface<sup>41</sup>.

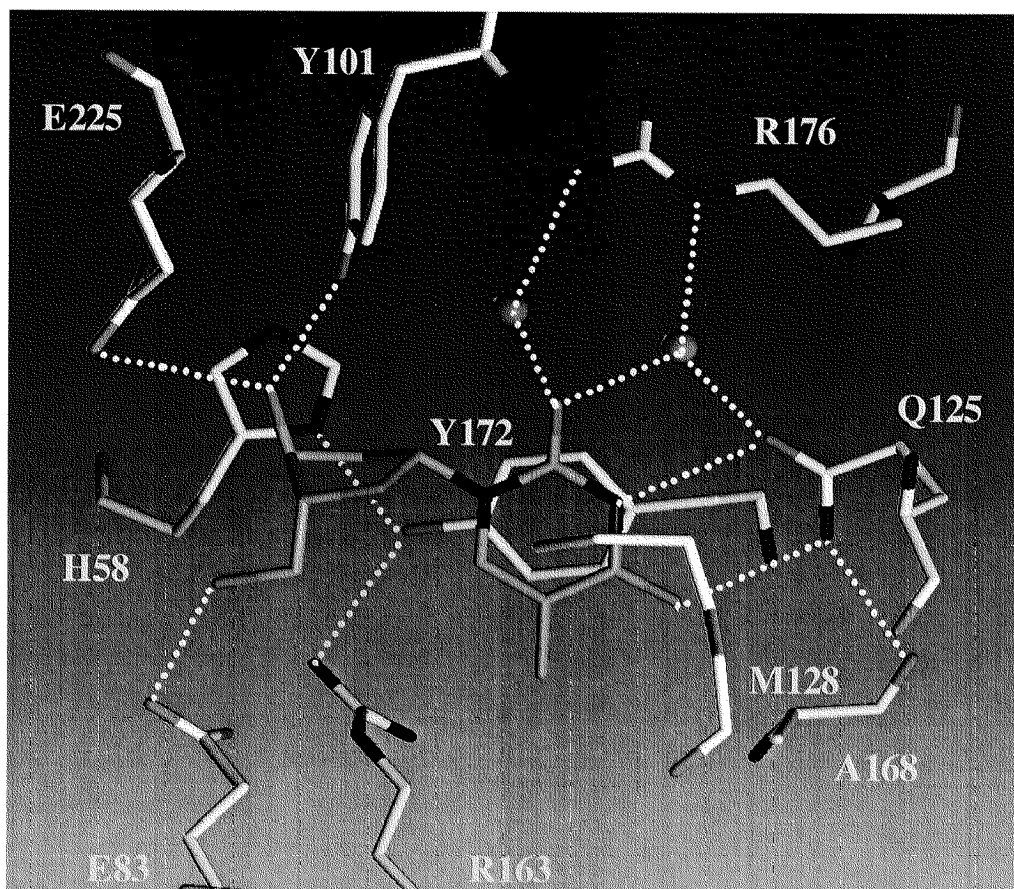


FIG. 3: View into the active site of HSV 1 TK (2VTK) showing the hydrogen bond network, which fixes the substrate thymidine (green).

A crucial part of the active center is the glycine rich loop connecting  $\beta$ -strand  $\beta 1$  with helix  $\alpha 1$ . This loop contains the sequence  $^{56}\text{G-X-X-G-X-G-K-T}^{63}$  forming a giant anion hole that accommodates the  $\beta$ - and  $\gamma$ -phosphoryl group of ATP<sup>1,2,41,43,47</sup>. The substrate thymidine is fixed within the enzyme by a sophisticated hydrogenbond network as depicted in figure 3.

The carboxamide of Q 125 hbonds with the N3 and O4 $\alpha$  atoms of thymine. For GCV binding the carboxamide is observed to be rotated by 180 $^\circ$ <sup>1</sup>. The O2 of dT is linked to R 176 via two water-mediated hbonds. The ribose moiety is orientated by Y 101 and E 225. Furthermore, dT is held between M 128 and Y 172 forming a sandwich-like complex. Unlike the solvent exposed ATP binding site ( $K_m$  12-118  $\mu\text{M}$ <sup>48,49</sup>) the dT binding site ( $K_m$  0.2  $\mu\text{M}$ <sup>49,50</sup>) is totally buried in the protein interior what would indicate an ordered mechanism of binding. Kinetic analysis of the protein indeed revealed a preferred, but not exclusive binding order of substrates i.e. dT binds prior to ATP<sup>51</sup>. Interestingly, ITC measurements showed that there is no detectable binding of ATP in absence of dT and the binding of dT is enhanced by two orders of magnitude in presence of ATP<sup>52</sup>. Although dT is tightly bound, it does not completely fill its binding pocket<sup>41</sup>.

With this constellation TK is not only able to phosphorylate dT or dU and its derivatives (figure 4) but also dTMP ( $K_m$  10 $\mu\text{M}$ )<sup>53</sup> that therefore must have two binding orientations within the binding site: one as product and one as substrate. Furthermore, TK accepts dC ( $K_m$  20  $\mu\text{M}$  to 500  $\mu\text{M}$ , dependent on the enzyme purification) and many guanine based analogs<sup>48,49,54-57</sup> and even adenine and hypoxanthine derivatives<sup>58</sup>. The phosphate donor for the monophosphorylation catalyzed by HSV 1 TK is normally ATP. But TK shows also high affinity for CTP, UTP, GTP and their deoxy analogs which is not surprising as the adenine moiety establishes only one hydrogenbond to the enzyme.  $\text{Mg}^{2+}$  is essential for the catalysis and it can exclusively be replaced by a  $\text{Mn}^{2+}$  ion. It is suspended by D 162 between the  $\beta$ - and  $\gamma$ -phosphate group of the ATP via two water molecules. To retain enzymatic activity, D 162 can only be exchanged by Q and G<sup>59</sup>. Other residues that only allow a single replacement are Y 172 which merely can be replaced by Phe<sup>60,61</sup> and Q 125 which can be exchanged by N with a ten-fold

loss in affinity for dT<sup>62</sup> or Met 128 by Ile which is a mutation that exists in nature in the HSV strain HFEM<sup>60,63</sup>.

TK accepts the L-stereoisomer of the deoxyribose of dT<sup>64</sup> as well as the antiviral drugs acyclovir or ganciclovir with their acyclic ribose analogs and the conformationally restricted sugar ring pucker of (N)-MCT<sup>65</sup>, (this work) disclosing its low stereochemical demands for the ribose moiety (figure 4). Therefore, HSV 1 TK is characterized by the extensive, broad substrate diversity for both, the base and the sugar moiety.

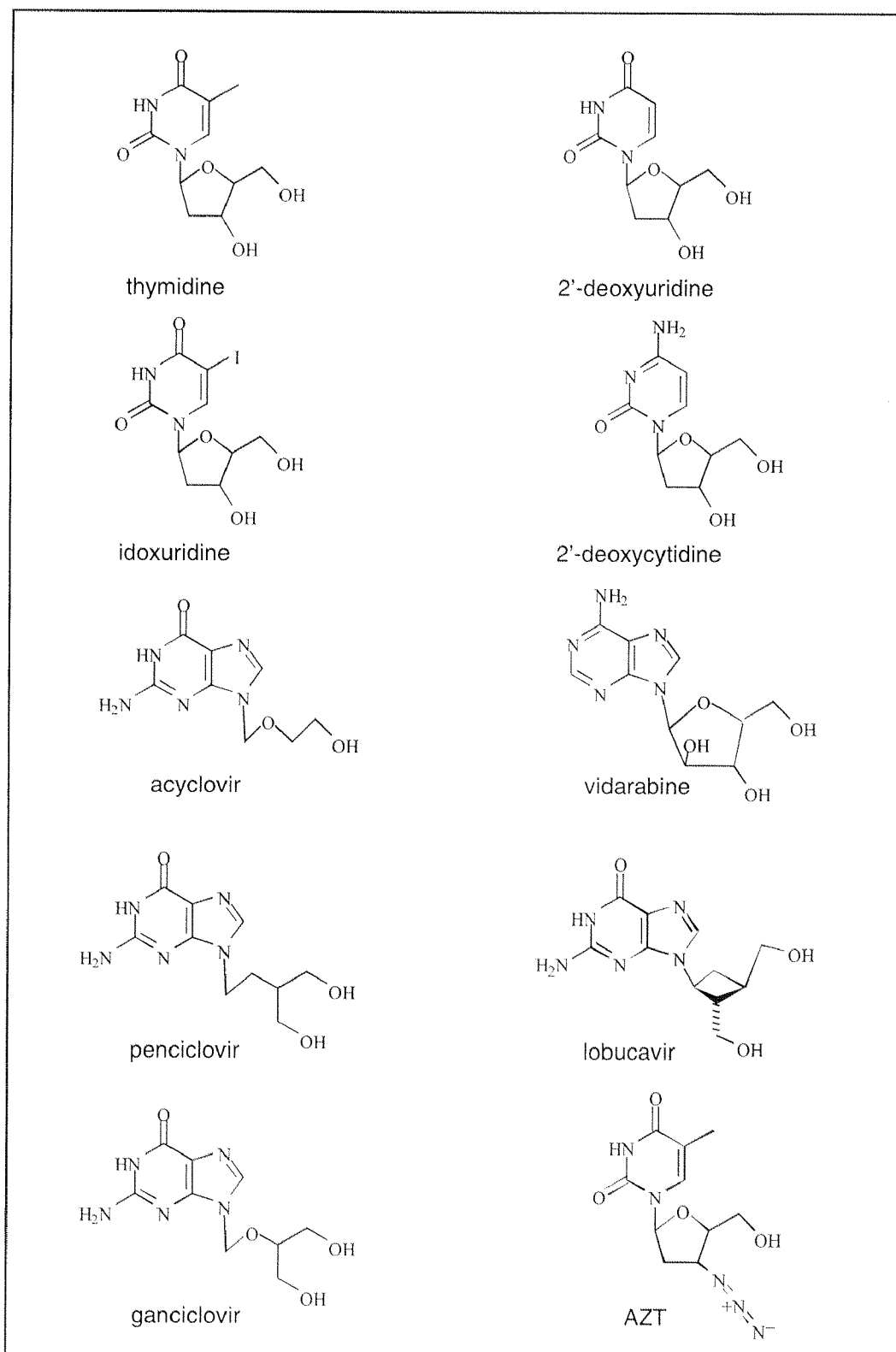


FIG. 4: Formula of HSV 1 TK substrates and substrate analogs

### 1.1.5 TK and Therapies

#### ***Antiviral therapy***

Progress made in recent years in understanding the biochemical mechanisms involved in the intracellular replication of viruses has dispelled the idea that specific antiviral therapy is impossible due to host and virus sharing the same biosynthetic pathways. Antiviral agents have been developed which interfere with virus specific functions or interact with virus-specific enzymes. Many nucleoside analogs have been investigated *in vitro* and *in vivo* for potential clinical use in antiviral therapy. Most of these analogs have a common mechanism of activation and action. Following phosphorylation to the nucleoside triphosphate form, by viral-encoded and cellular kinases in the virally infected cells, the nucleoside triphosphate may then preferentially inhibit the viral DNA polymerase or the triphosphorylated analog may be incorporated into growing viral DNA strands and so bring about the termination of DNA synthesis<sup>66</sup>. The differences in biochemical properties, particularly in substrate specificities, between cellular and viral TK now form the basis of action of several antiviral compounds. The selectivity of action of these nucleoside analogs can result from increased levels of cytosolic/virus encoded TK in the infected cells or from a higher affinity of the analog triphosphate for the viral polymerase than for cellular polymerases<sup>67</sup>.

#### ***Gene therapy***

The general concept of suicide gene therapy is based on conferring new metabolic property to target cells to enable the activation of prodrugs in these cells, artificially creating a high therapeutic effect. One of these genes is the thymidine kinase (TK) gene of herpes simplex virus (HSV-1) that express TK with different specificity than the host TK. The TK enzyme in HSV TK transfected tumor cells allows the conversion of the nontoxic nucleoside analogue ganciclovir into a phosphorylated metabolite which is then incorporated into DNA, thereby inhibiting DNA synthesis and leading to cell death<sup>68-70</sup>. In 1992, Culver and co-workers reported on the direct *in situ* introduction of the HSV1 TK gene into

proliferating tumor cells of rats bearing cerebral gliomas, subsequently followed by treatment with ganciclovir (GCV). They found that the gliomas completely regressed in the ganciclovir- and TK gene vector-treated rats<sup>69</sup>. Meanwhile, the successful treatment of experimental gliomas, sarcomas and melanomas in rats and (nude) mice with the combined HSV 1TK gene/GCV approach was confirmed by independent investigators<sup>71-77</sup>. The novel approach, using HSV 1 TK/GCV, has also been used to a variety of other tumor models i.e. Abelson leukemia virus-induced lymphomas<sup>78</sup>, colorectal adenocarcinoma<sup>79</sup>, macroscopic liver metastases<sup>80</sup>, breast tumor<sup>81</sup>. Interestingly, Grignet-Debrus and Calberg-Bacq found a dose-dependent arrest of the breast cancer MDA-MB-435 cell growth by BVDU ((E)-5-(2-bromovinyl)-2'-deoxyuridine) in nude mice, subcutaneously bearing the VZV TK gene-transfected tumor cells<sup>82</sup>.

Currently, the therapeutical approach of treating tumors with suicide genes after gene transfer is in the stage of clinical research for the treatment of recurrent pediatric malignant astrocytomas<sup>83</sup>, progressive or recurrent primary supratentorial pediatric and adult malignant brain tumors<sup>84-86</sup> and ovarian cancer<sup>87</sup>. First results on the regression of recurrent malignant brain tumors revealed a pronounced antitumor activity<sup>88</sup> with a higher rate of success in patient with small tumors. From the clinical trials performed until now the feasibility of the approach has been assessed but several efficacy limiting factors due to "technical obstacles" have been highlighted<sup>89</sup>. The first one is related to the efficacy of the enzyme-prodrug system itself and is the immunosuppressive effects of the dosages needed for tumor regression. The second one is low tissue and tumor specificity and low efficiency of transduction and is related to gene delivery and distribution and conditional expression in human cells.

Another attractive approach is the combination of different suicide gene<sup>90</sup> and/or the combination of gene therapy with oncosuppressor genes such as p21. For example a promising suicide gene and suitable partner of HSV 1 TK is represented by the cytosine deaminase (CD) gene isolated from the *E.coli* bacterium<sup>91-93</sup>. Cytosine deaminase converts the nontoxic prodrug 5-

fluorocytosine into the toxic drug 5-fluorouracil. It seems reasonable to combine the TK approach with that of CD. By using fusion genes that express both thymidine kinase and cytosine deaminase during the therapy both prodrugs ganciclovir and 5-fluorocytosine are applied and not only an additive but a synergistic effect is achieved. It bases on the accumulation of ganciclovir in the early S-phase of the CD/HSV 1 TK expressing cells, that inhibits the repairing mechanism of the broken DNA-strings caused by the 5-fluorocytosine treatment. The new approach represents an unambiguous improvement compared to the initial suicide gene therapy concept and has been already used by several groups in recent time<sup>94-97</sup>.

Furthermore, a successful application of enzyme-prodrug (HSV 1TK/GCV) gene therapy concept was recently demonstrated in bone marrow transplantation patients who developed lymphomas<sup>98</sup>. The results showed that HSV1 TK gene transduction of donor lymphocytes may increase the efficacy and safety of allogeneic bone marrow transplantation and that enzyme-prodrug gene therapy approach may have a broad application in the clinic, even beyond the treatment of cancer.



## 1.2 Enolase

Enolase, 2-phospho-D-glycerate hydrolase or phosphopyruvate hydratase (EC 4.2.1.11) is a highly expressed key enzyme of glycolysis and gluconeogenesis in all prokaryotic and eukaryotic cells that is needed for the conversion of D-glycerate-2-phosphate (DGP) to phosphoenolpyruvate (PEP) and vice versa (Fig. 5)

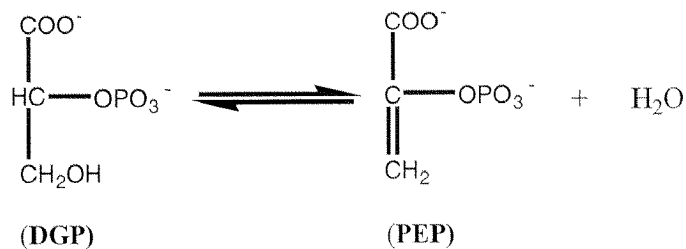


Fig. 5 Scheme of the catalytic reaction of enolases.

The enzyme is a homodimer and requires two divalent cations per active site for catalytic activity. The first one is often referred to as the „conformational“ ion because its binding induces a conformational change. The second ion binds only in presence of the substrate (DGP) and is necessary for enzymatic activity.

As one of the enzymes involved in glycolysis and fermentation, enolase is likely to be almost ubiquitous in the biological world<sup>99</sup>.

### 1.2.1 The sequence homology

The alignment of fungal enolase sequences (Fig.6) shows a very high degree of amino acid identity (at least 77%), even among fungi that are not close relatives according to current systematics.

Immunological cross-reactivity of recombinant enolases of *Cladosporium herbarum* and *Alternaria alternata* and the enolases in extracts from *Saccharomyces* and *Candida* with patients' IgE was shown by Breitenbach and co-workers<sup>100</sup>.

This cross-reactivity seems to be well correlated with the high sequence homology.

<i>A. alternata</i>	m t i t K I H A R	S V Y D S R G N P T	V E V D i v T E t G	L h R a I V P S G A	S T G s H E A c E L	R D G D K S K W g G
<i>C. herbarum</i>	m p i S K I H s R	y V Y D S R G N P T	V E V D i v T E t G	L h R a I V P S v A	S T G s H E A c E L	R D G D K S K W a G
<i>S. cerevisiae</i> c.1	- a v S K v y A R	S V Y D S R G N P T	V E V e l T T E K G	v F R S I V P S G A	S T G V H E A L E m	R D G D K S K W m G
<i>S. cerevisiae</i> c.2	- a v S K v y A R	S V Y D S R G N P T	V E V e l T T E K G	v F R S I V P S G A	S T G V H E A L E m	R D e D K S K W l G
<i>C. albicans</i>	M s y a t K I H A R	y V Y D S R G N P T	V E V D f T T d K G	L F R S I V P S G A	S T G V H E A L E L	R D G D K S K W l G
<i>A. alternata</i>	K G V t K A V A N V	N D t I A P R L I K	e k L D V K D Q s A	V D a F L n k L D G	T t N K t n L G A N	A I L G V S M A i A
<i>C. herbarum</i>	K G V t K A V A N V	N e i I A P A L I K	e N L D V K D Q a A	V D a F L n k L D G	T t N K t K I G A N	A I L G V S M A v A
<i>S. cerevisiae</i> c.1	K G V t h A V K N V	N D v I A P A I v K	A N i D V A D Q k A	V D d F L i S L D G	T a N K S K L G A N	A I L G V S l A A s
<i>S. cerevisiae</i> c.2	K G V m n A V n N V	N n v I A a A I v K	A N L D V K D Q k A	V D d F L i S L D G	T a N K S K L G A N	A I L G V S M A A A
<i>C. albicans</i>	K G V I K A V A N V	N D I I A P A L I K	A k i D V v D Q a k	I D e F L i S L D G	T p N K S K L G A N	A I L G V S l A A A
<i>A. alternata</i>	k A A A A E K g V P	L Y a H I s D L a g	t K K - P Y V L P V	P F q N V L N G G S	d A G G r L A F Q E	F M I v P c e A p T
<i>C. herbarum</i>	k A A A A E K r V P	L Y a H I s D L S g	t K K - P f V L P V	P F m n v N G G S	H A G G r L A F Q E	F M I v P s G A p s
<i>S. cerevisiae</i> c.1	r A A A A E K n V P	L Y k H I A D L S k	s K t s P Y V L P V	P F i N V L N G G S	H A G G A L A l Q E	F M I A P T G A k T
<i>S. cerevisiae</i> c.2	r A A A A E K n V P	L Y q H I A D L S k	s K t s P Y V L P V	P F i N V L N G G S	H A G G A L A l Q E	F M I A P T G A k T
<i>C. albicans</i>	n A A A A a g I P	L Y k H I A n l S n	a K K g k f V L P V	P F q N V L N G G S	H A G G A L A F Q E	F M I A P T G v s T
<i>A. alternata</i>	F s E A M R q G a E	V Y q k L K a L a K	K t Y G O S A G N V	G D E G G V A P D I	Q T A E E A L D L I	t k A I c e A G Y t
<i>C. herbarum</i>	F t E A M R q G a E	V Y q k L K S L T K	K R Y G O S A G N V	G D E G G V A P D I	Q T A E E A L D L I	t D A I c c A G Y t
<i>S. cerevisiae</i> c.1	F a E A l R I G S E	V Y H N L K S L T K	K R Y G a S A G N V	G D E G G V A P n I	Q T A E E A L D L I	v D A I k a A G H d
<i>S. cerevisiae</i> c.2	F a E A M R I G S E	V Y H N L K S L T K	K R Y G a S A G N V	G D E G G V A P n I	Q T A E E A L D L I	v D A I k a A G H d
<i>C. albicans</i>	F s E A l R I G S E	V Y H N L K S L T K	K k Y G O S A G N V	G D E G G V A P D I	K T p k E A L D L I	m D A I d k A G Y k
<i>A. alternata</i>	G K i K I A M D V A	S S E F Y K a d e k	K Y D L D F K N P d	S D K S K W L T y e	Q L A e m Y k S L a	e k Y P I V S I E D
<i>C. herbarum</i>	G q i K I A M D V A	S S E F Y K a d e k	K Y D L D F K N P d	S D K S K W L T y e	Q L A d q Y n e L a	a k Y P I V S I E D
<i>S. cerevisiae</i> c.1	G K V K I g l D c A	S S E F f K - - D G	K Y D L D F K N P n	S D K S K W L T G p	Q L A D I Y h S l m	k r Y P I V S I E D
<i>S. cerevisiae</i> c.2	G K V K I g l D c A	S S E F f K - - D G	K Y D L D F K N P e	S D K S K W L T G v	e L A D m Y h S l m	k r Y P I V S I E D
<i>C. albicans</i>	G K V g I A M D V A	S S E F Y K - - D G	K Y D L D F K N P e	S D p S K W L s G p	Q L A D I Y e q L i	s e Y P I V S I E D
<i>A. alternata</i>	P F A E D D W E A W	S y F F K T y d g q	- - I V G D D L T V	T N P e f I K k A I	E l K s c N A L L L	K V N Q I G T i T E
<i>C. herbarum</i>	P F A E D D W E A W	S y F y K T s G s d	f Q I V G D D L T V	T N P e f I K k A I	E t K c A N A L L L	K V N Q I G T i T E
<i>S. cerevisiae</i> c.1	P F A E D D W E A W	S H F F K T a G -	I Q I V a D D L T V	T N P k r I a T A I	E K K A A d A L L L	K V N Q I G T L s E
<i>S. cerevisiae</i> c.2	P F A E D D W E A W	S H F F K T a G -	I Q I V a D D L T V	T N P a R I a T A I	E K K A A d A L L L	K V N Q I G T L s E
<i>C. albicans</i>	P F A E D D W d A W	v H F F e r v G d k	I Q I V G D D L T V	T N P t R I K T A I	E K K A A N A L L L	K V N Q I G T L T E
<i>A. alternata</i>	a I q A A k D a F g	A G W G V M V S H R	S G E T E D v t I A	D i F V G L R s G Q	I K T G A P A R S E	R L A K L N Q I L R
<i>C. herbarum</i>	a I n A A k D S F A	A G W G V M V S H R	S G E T E D v t I A	D i F V G L R a G Q	I K T G A P A R S E	R L A K L N Q I L R
<i>S. cerevisiae</i> c.1	S I k A A q D S F A	A G W G V M V S H R	S G E T E D T F I A	D L F V G L R t G Q	I K T G A P A R S E	R L A K L N Q I L R
<i>S. cerevisiae</i> c.2	S I k A A q D S F A	A n W G V M V S H R	S G E T E D T F I A	D L F V G L R t G Q	I K T G A P A R S E	R L A K L N Q I L R
<i>C. albicans</i>	S I q A A n D S y A	A G W G V M V S H R	S G E T E D T F I A	D L s V G L R s G Q	I K T G A P A R S E	R L A K L N Q I L R
<i>A. alternata</i>	I E E E L G D k r l	Y A G n N F r t A v	n L *			
<i>C. herbarum</i>	I E E E L G D k r l	Y A G d N F r t A j	n L *			
<i>S. cerevisiae</i> c.1	I E E E L G D n A v	f A G e N F h g d	k L *			
<i>S. cerevisiae</i> c.2	I E E E L G D k A v	Y A G e N F h g d	k L *			
<i>C. albicans</i>	I E E E L G s e A l	Y A G k d F q k A s	q L *			

**Fig. 6 Multiple amino acid sequence alignment of five fungal enolases.** The consensus is shown by capital letters and indicates that the amino acid residue is identical in at least three sequences.

A detailed analysis of the few amino acids that are directly involved in IGE contact will be only possible when the structure of the fungal enolase will be elucidated.

### 1.3 Crystallography

Until the 1930s, and indeed that for many years, the rationale for crystallizing proteins, particularly enzymes, was to supply a technique for purifying a specific protein from a complex extract, or to demonstrate, in the classical chemists sense, the purity of a preparation. Throughout this period, crystallinity was associated with purity. In the late 1930s, however, certain X-ray diffractionists, such as Astbury, Bernal, Bragg, Crowfoot, Frankuchen and Perutz, began to turn their attention to protein crystals as a source of structural information about biological macromolecules<sup>101</sup>.

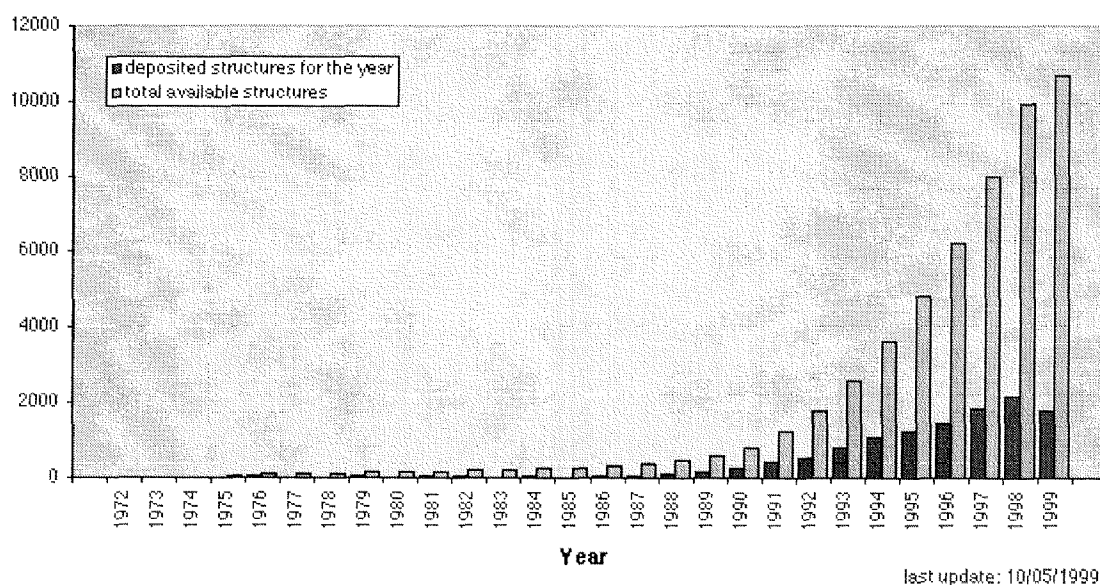
When Max Perutz and his colleagues built the first atomic model of hemoglobin in the 1960s, it took up 16 square feet of floor space at the Medical Research Council's laboratories in Cambridge, United Kingdom. Completing that model was a monumental task, begun in 1937 with Perutz's arrival in Cambridge, and it stands as a monument to the way things will never be again in the burgeoning field of structural biology<sup>102</sup>. Perutz described the arduousness of that decades-long saga of making crystals, interpreting X-ray diffraction patterns, solving the phase problem and creating a structure out of those data after his 80<sup>th</sup> birthday as follows: "First of all, you had the computer output of numbers giving the density distribution of the hemoglobin molecule in three dimensions. Then you had contour maps of the electron density drawn by hand, transferred onto plexiglass, and stacked up, like microtome sections through tissue. Then I had to measure the coordinate of peaks on these sections. Then I'd build a three-dimensional model"<sup>102</sup>.

Recent advances in molecular biology and protein crystallization, along with improvements in hardware and software, have merged to vastly increase the pace at which new structures of complex biological molecules are solved. The use of recombinant DNA techniques has allowed crystallographers to select a specific protein of interest, clone its gene, and make large amounts of relatively pure protein, which is necessary (although not, sufficient) for getting good crystals.

Wayne Hendrickson says: " In the early days, it was so difficult that when you got a structure done, much of the interest in the molecule may have already passed.

Now you are going from a description of activity to the structure within a year or so. Now crystallographic work is quick enough and so incisive that it's become really central. Timeliness is having a huge impact"<sup>102</sup>.

A measure of this acceleration comes from Brookhaven. Two decades ago there were 15 to 25 structures a year deposited in the Protein Data Bank. Currently new structures are being registered at a rate of about 180 per month. If current trends continue, the data bank will swell to 30,000 structures by 2000.



**Fig.7 Diagram showing the PDB content growth during the past two decades<sup>103</sup>.**

Given our current expertise, and the certain future developments in genetically altering organisms to produce proteins of modified structure and function, the concept of protein engineering is nearing reality. Similarly, our ability to describe and utilize protein structure and to define interactions with ligands has made possible the rational design of new drugs and pharmacological events. Even in the absence of any intention toward applied use or value, the correlation of regulation, mechanism, and function of proteins with their detailed molecular structure has now become a primary concern of modern biochemistry and molecular biology.

### 1.3.1 A historical background of macromolecular Crystallization

Crystal growth, which is a very old activity, has always intrigued mankind, and many philosophers and scientists have compared it with the biological process of reproduction, and it has even been speculated that the duplication of genetic material would occur through crystallization like mechanisms<sup>104</sup>. The history of protein crystallization contains a number of lessons that are still of value to us today. It is rich in methods long forgotten and now ignored and at the same time shows us that many of our modern approaches are only reinvented. The literature is filled with tricks and quirks that illustrate the subtleties, frustrations, and vagaries of the crystal growth process. Pioneers in the field generally were forced, by the primitive techniques available to them, to work exclusively with proteins directly obtainable in very large amounts<sup>101</sup>.

More than one century ago the first reports on crystallization of proteins were published by Hünefeld, Reichert, Leydig, Kölliker and Budge when they crystallized hemoglobin from the blood of different vertebrates and invertebrates and demonstrated several important points<sup>105-109</sup>. Their observations suggested for the first time that protein crystals could be obtained by the controlled evaporation (slow dehydration) of a concentrated protein solution from relatively crude physiological conditions. It further suggested that proteins form crystals in much the same way as do conventional small molecules and that a high degree of purification was not always essential. Through 1850, all of the blood crystals reported were observed to have grown more or less fortuitously and no investigator had suggested any general procedure for their directed growth. The first person to actually devise successful and reproducible methods for the growth of hemoglobin crystals was Fünke who began publishing a series of articles on the purposeful growth of hemoglobin crystals in 1851<sup>110</sup>. In 1909 Reichert and Brown published a monumental monograph on the extensive investigation of hemoglobin crystal growth, where they thoroughly had explored methods for obtaining hemoglobin crystals from the blood of several hundred different animals following the thesis that the outward form of a crystal reflected the composition and structure of the molecules that composed it, and that hemoglobin from different

species of animal would contain differences of composition and molecular structure commensurate with their phylogenetic differences<sup>109</sup>.

The crystallization of hemoglobin was followed by the crystallization of hen-egg albumin and a series of plant seed proteins, principally globulins<sup>101</sup>. In 1926, after nine years of intermittent work, Sumner crystallized the jack bean urease by extracting jack bean meal with acetone, filtering, and then cooling the still filtering solution to about 2°C<sup>111</sup>. At almost the same time the first polypeptide hormone, insulin, was also crystallized by Abel et al<sup>112</sup>. This was one of the first examples of crystal growth promoted by the addition of divalent metal ions. Other notable examples of the use of metal ions to induce protein crystal growth were the crystallization of the iron storage protein, ferritin<sup>113</sup>, and the ribonuclease from bovine pancreas<sup>114</sup>. Another scientific achievement arose in 1935 when Stanley crystallized Tobacco mosaic virus. In 1937 Sumner and Dounce crystallized the heme-containing enzyme catalase<sup>115</sup>, but the greatest number of crystalline enzymes, however were provided during the 1930s and 1940s by the remarkable work of Northrup, Herriott and Kunitz<sup>114</sup>. Working primarily with hydrolytic enzymes extracted from the pancreas, they succeeded in crystallizing trypsin and trypsinogen, chymotrypsin and chymotrypsinogen, pepsin and pepsinogen, both pancreatic and soybean trypsin inhibitors and ribonuclease. From other sources, they obtained crystals of diphtheria toxin and hexokinase. The methods that were developed to crystallize these proteins describe for the first time the exploitation of several approaches now in common use and are given in the book *Crystalline Enzymes*<sup>114</sup>: Temperature variation under constant solution conditions, dialysis against low ionic strength to take advantage of the salting in property of many proteins, and a further use of organic solvents as precipitating or crystallizing agents.

In 1946 Sumner, Stanley and Northrop shared the Nobel Prize for their remarkable research. This remained the only Nobel Prize awarded cultivators of

protein crystals until Deisenhofer, Huber and Michel received the prize in 1988 for the first crystalline membrane protein<sup>101</sup>.

### 1.3.1 Definition & principles of protein crystallization

Alexander McPherson defines Macromolecular crystallization as a matter of searching, as systematically as possible, the ranges of the individual parameters that impact upon crystal formation, finding a set or multiple sets of these factors that yield some kind of crystals, and then optimizing the variable sets to obtain the best possible crystals for X-ray analysis<sup>116</sup>.

In his book, Jan Drenth considers protein crystallization mainly a trial-and-error procedure in which the protein is slowly precipitated from its solution<sup>117</sup>. Obtaining suitable single crystals is the least understood step in the X-ray structural analysis of a protein. The science of protein crystallization is an underdeveloped area, although interest is growing, spurred especially by microgravity experiments in space flights<sup>118</sup>.

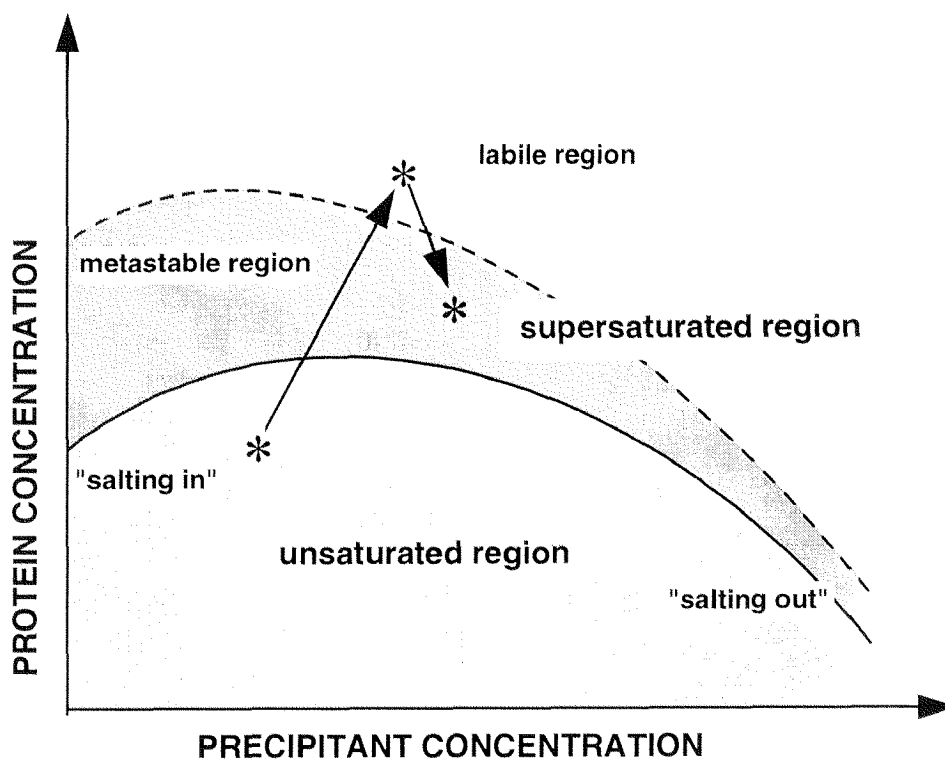
Procedures for growing crystals of macromolecules are generally empirical, owing to the present inadequate level of understanding of the complex physical processes involved. Crystallization is usually achieved by varying those physical parameters affecting solubility of the macromolecule in order to attain a state of supersaturation<sup>119</sup>.

Macromolecular crystals are relatively small in comparison with conventional crystals, rarely exceeding a millimeter on an edge. Because only one stereoisomer of a biological macromolecule exists in nature, they do not form crystals possessing inversion symmetry and, therefore, generally exhibit simple shapes that lack the polyhedral character of small molecule crystals. They are extremely fragile, often crushing at the touch, they degrade when outside of a narrow range of temperature, ionic strength, or pH; they generally exhibit weak optical properties; and they diffract X-rays to resolutions far of the theoretical limit. The reason for most of these character deficiencies is that macromolecular crystals

incorporate large amounts of solvent in their lattices, ranging from about 30%, at the lower limit, to 90% or more, in the most extreme cases<sup>118</sup>.

Additionally there are two further crucial differences between macromolecular and conventional crystal growth, which have important practical consequences. Firstly, macromolecular crystals are usually nucleated at extremely high levels of supersaturation, often several hundred to a thousand percent. In contrast small molecule crystals usually nucleate at only a few percent of supersaturation. Though, every quantitative aspect of crystal growth can be observed as a direct function of supersaturation<sup>120</sup>. The second crucial point is the fact, that supersaturated macromolecular solutions produce amorphous precipitates in addition to crystal nuclei. Therefore, competition between crystals and precipitate exists at both the nucleation and growth stages and is particularly acute because it is promoted by high levels of supersaturation. Amorphous precipitates are of higher energy state and being kinetically favored, they tend to dominate the solid phase and inhibit or influence crystal formation. Though once the solution is brought to supersaturation and nuclei are formed, crystal growth is then achieved by reducing supersaturation to a lower level. Maintaining a high supersaturation would result in the formation of too many nuclei and therefore too many small crystals (Fig.8)<sup>116,117</sup>.





**Fig. 8 Phase diagram showing the solubility of a hypothetical protein as a function of precipitant concentration.** The solid line represents the maximum solubility or saturation curve of the protein. The supersaturated region lies above the maximum solubility curve and is divided into the metastable region and the labile region. In the metastable region crystal nuclei can only grow, while in the labile region they can spontaneously form and grow.

Initially, the parameters that one wishes to establish as rapidly as possible are optimal concentration for each precipitant used, optimal pH for solubilization and crystallization, and the effect of temperature. Even though nowadays many screenings are available on the market, it's worth to start to screen conditions using the two major classes of precipitant in use, namely ammonium sulfate and PEG. Depending on the available quantities of protein one can extend the investigations using organic solvents and short chain alcohols, e.g. ethanol and MPD. Initially, a wide pH range of 3.5 to 9.0 should be analyzed in  $\Delta\text{pH}$  intervals of 0.5 against different precipitant concentrations. The trials should be set up at two different temperature (room temperature and 4°C for example). The aim of this wide range screen is to get an initial idea in which zone one should concentrate the further experiments.



### 1.3.2 Crystallization techniques

Laboratories devoted to protein crystallization in the late 1920s and 1930s used a few pieces of simple but special equipment. One of the most important of these items was a large refrigerator or better yet, a walk-in cold room. A good centrifuge was also essential, preferably one which would process at least a liter of material at a time. Refrigerated centrifuges were rare; usually a room-temperature centrifuge was used in a walk-in cold room. A good microscope, preferably one with photographic attachments was necessary in working with crystallizable proteins. The graduated cylindrical pipette was another essential item; smaller varieties had graduations down to 0.02 or 0.01 ml. Other items of necessary equipment included grinding and mincing machines for use with starting materials. Still another piece of laboratory equipment rewired was the glass electrode. Additional special equipment consisted of glass stills assembled from available and hand-made glassware<sup>121</sup>.

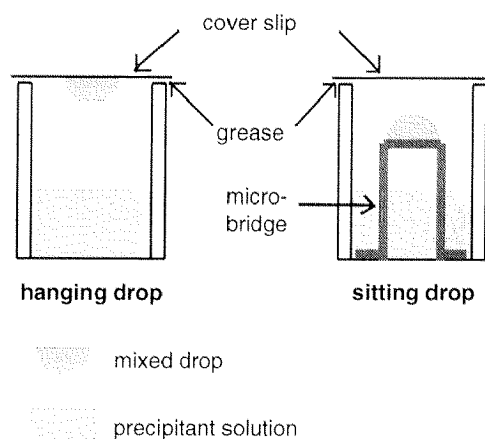
In practice the amount of protein available for crystallization experiments is often very small. To determine the best crystallization conditions it is usually necessary to carry out a great number of experiments. To get an impression of that imagine you have 1 mg purified protein at an usual starting protein concentration of 10 mg/ml, depending on the method you can perform maximally 50 crystallization experiments with it.

The four commonly used methods to perform the crystallization trials are vapor diffusion, free interface diffusion, batch, and dialysis and are briefly discussed below.

#### ***Vapor diffusion***

The vapor diffusion technique is probably the most widely used crystallization method throughout the world. The principle by which supersaturation of macromolecules is achieved is very simple: Drops containing a mixture of protein and precipitant solution (normally 1:1) are equilibrated in a sealed system against a reservoir containing the precipitant solution, either by hanging or sitting drop (fig. 10). The difference in precipitant concentration between the drop and the well

solution is the driving force which causes then diffusion of the solvent until an equilibrium between the vapor pressure of the drop and the reservoir is achieved.



**Fig. 10 Schematic representation of the hanging and sitting drop techniques.** The mixture of protein and precipitant solution is suspended as a drop above the precipitant solution in a sealed system.

Vapor diffusion is the optimal technique to use either when screening a large number of conditions (by varying the composition of each well solution) or when dearth of protein prevents the use of other methods. Furthermore, this method can be used to increase or decrease the concentration of protein in the equilibrated state relative to its initial concentration. This is done by varying the volume of protein mixed with the well solution when the drop is initially setup. Since the drop equilibrates so that the precipitant concentration matches that of the well solution, the final volume of the drop will always equal that of the initial amount of well solution mixed with the protein.

### ***Free interface diffusion***

In this method the protein solution and the solution containing the precipitant are layered on top of each other in a capillary (small-bore capillary or melting point capillary) and are let slowly diffuse into each other. The two should be made up so that at equilibrium the concentration of the precipitant is still high enough to promote crystal growth. To avoid rapid mixing the crystallizing agent can be frozen before the protein solution is layered above. Nucleation and crystal growth generally occurs at the interface between the two layers, at which both the concentration of salt and the concentration of protein are at their highest values. This method can be used as a means of growing large crystals<sup>122,123</sup>.

### ***Batch Crystallization***

This method is the oldest and simplest one. A concentrated precipitant solution is added to a concentrated protein solution to produce suddenly a state of high supersaturation, which induces crystal growth without any further processing. This can be done with up to ml amounts of solution and typically results in larger crystals due to the larger volumes of solute present and the lower chance of impurities diffusing to the face of the crystal. Needless to say, this technique is by far the most expensive in terms of consumption of the solute macromolecule, and thus should not generally be used to screen initial conditions for crystallization. An automated system for microbatch crystallization has been designed by Chayen and coworkers<sup>124</sup>.

### ***Dialysis***

Dialysis is familiar to nearly all biochemists as a means of changing some properties of a protein-containing solution. Dialysis techniques utilize diffusion and equilibration of small precipitant molecules through a semipermeable membrane as a means of slowly approaching the supersaturation state at which the macromolecule crystallizes. It has the advantage that by liquid-liquid diffusion through a semi-permeable membrane, a protein solution can be exposed to a continuum of potential crystal-producing conditions without actually altering directly the mother liquor. Diffusion through the membrane is slow and controlled. Because the rate of exchange of substituents in the mother liquor is proportional to the gradient of concentrations across the membrane, the nearer the system approaches equilibrium, the more slowly it changes.

This method was generally applicable only on a large scale until it has been adapted to microtechniques by crystallographers using capillaries or small plexiglass buttons<sup>125</sup>. The method confines a protein solution to the interior of a glass capillary, or the microcavity of a small plexiglass button. Dialysis buttons themselves are available in a variety of sizes from 5-350  $\mu\text{l}$ . The cavity of the button or the ends of the microcapillary tube are then closed off by a semipermeable dialysis membrane and submerged in a much larger volume of an

exterior liquid and the whole system kept within a closed vessel such as a test tube or vial.

The advantage of dialysis over other methods is in the ease with which the precipitating solution can be varied, simply by moving the entire dialysis button, capillary or tube from one condition to another. Protein can thus be continuously recycled until the correct conditions for crystallization are found<sup>126</sup>. One drawback of this method is that it does not work at all with concentrated PEG solutions, as they tend to draw all the water out of the button or sack faster than PEG dialyzes across the membrane, thus resulting in precipitated protein.

### **1.3.3 Preparation of Crystals for the X-ray experiment**

The high solvent content of the crystals causes problems in handling them due to destabilization by solvent evaporation. Therefore, protein crystals should always be kept in their mother liquor, in the saturated vapor of their mother liquor, or at a sufficiently low temperature to prevent evaporation of the solvent<sup>117</sup>. For the first collection of X-ray data the crystal should be mounted in a thin-walled glass capillary of borosilicate glass or quartz as described in<sup>117</sup> and should be measured at or near room temperature. To reduce the liability of the crystals to radiation damage it's worth to collect the data set by keeping the crystal in a stream of cold nitrogen gas at a temperature in the range of 100-120 K (cryocooling). Therefore, the crystal must first be soaked into optimized cryoprotectant conditions, containing high amounts of glycerol or MPD<sup>127</sup>. The crystal is lifted from the cryoprotectant solution by a loop made by a thin fiber, which has approximately the size of the crystal or a little larger ("lasso technique") and is then flash frozen directly on the cryo stream<sup>117</sup>. The same procedure can be used for soaking compounds into the protein crystals, such as inhibitors or different substrates. (see chapter 4 & 5)

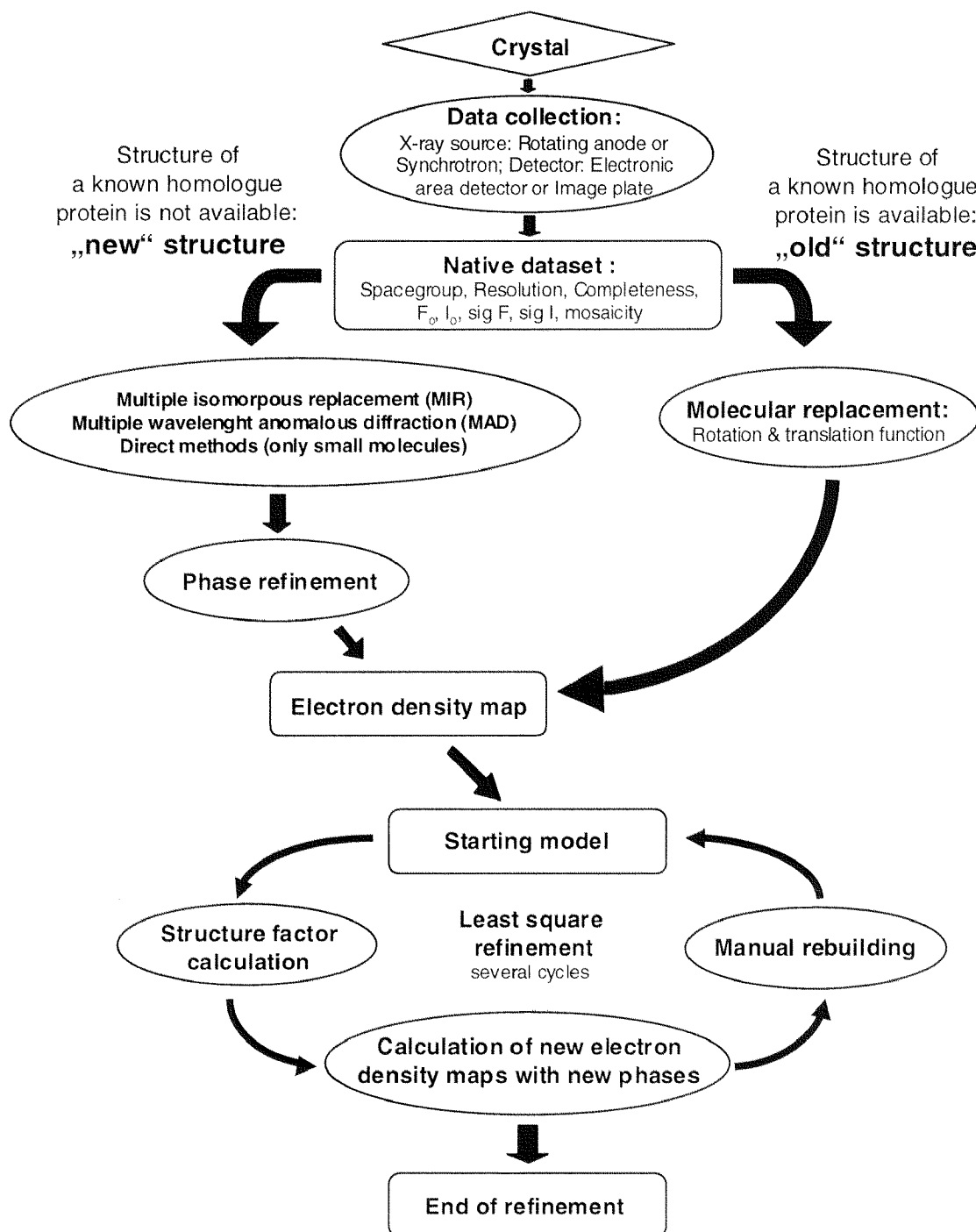
### **1.3.4 From data collection to solving the structure**

The first structures, myoglobin and haemoglobin, were elucidated only around 1960, despite the fact that crystals of proteins have been known since the beginning of the century. The development of the isomorphous replacement

technique made it possible. Since then, a lot of theoretical and technical advances have been made: among them, the use of anomalous dispersion, molecular replacement, and the development of oscillation techniques in data collection of crystals with large cell dimensions. At the end of the 1970s the growth of computing power made possible the refinement in reciprocal space. In the meantime, the introduction of powerful and affordable graphic software, allowed faster interpretation of electron density maps and model building. Only in very recent years the availability of synchrotron radiation and two-dimensional detectors has made data collection a routine procedure.

X-ray crystallography is an experimental technique that exploits the fact that X-rays are diffracted by crystals. It is not an imaging technique. X-rays have the proper wavelength (in the Ångström range,  $\sim 10^{-8}$  cm) to be scattered by the electron cloud of an atom of comparable size. Based on the diffraction pattern obtained from the periodic assembly of molecules or atoms in the crystal, the electron density can be reconstructed. Additional phase information extracted either from the data of a known structure or from supplementing experiments must be supplied to complete the reconstruction (phase problem). A model is then progressively built into the experimental electron density, refined against the data and the result is a quite accurate molecular structure<sup>128</sup>.

The knowledge of accurate molecular structures is a prerequisite for rational drug design and for structure based functional studies to aid the development of effective therapeutic agents and drugs. Crystallography can reliably provide the answer to many structure related questions, from global folds to atomic details of bonding. In contrast to NMR, which is an indirect spectroscopic method, no size limitation exists for the molecule or complex to be studied. The price for the high accuracy is that a good crystal must be found, and that only limited information about the molecule's dynamic behavior is available from a single diffraction experiment<sup>128</sup>. To illustrate and give an impression of the work needed for solving a protein crystal structure the most important steps from data collection to the three-dimensional structure are illustrated schematically in the following flow-chart:



**Fig. 11 Summarized steps involved in the resolution of the three-dimensional structure of a biological macromolecule** The steps and the possible different pathways are briefly described in the text.



### 1.3.5 Dynamic light scattering to predict crystallization?

While common purity analyses such as chromatography and electrophoresis may provide data on the presence of foreign molecules, they cannot provide immediate data on aggregation and denaturation which cause problems in crystallization. Light scattering studies have confirmed the importance of monodispersity in allowing crystallization of macromolecules. People involved in protein crystallization is following the consensus that the chance of obtaining crystals would be >80% if by some means monodispersity could be insured<sup>129-132</sup>. Therefore it is usual to determine the uniformity of the protein solution prior to crystallization trials<sup>133-136</sup>.

#### *The theory of light scattering*

The light source for a light scattering setup in a laboratory environment is usually a laser. Lasers are preferred because they provide a monochromatic, intense, well-defined beam of light. When a single molecule is illuminated by a laser, the molecule re-irradiates part of the light in all directions due to the interaction of its dipole moment with the electric field vector. Depending on the wavelength of the laser and the dimensions of the molecule, this scattered light can show very characteristic intensity patterns. Small molecules scatter light equally in all directions. On the other hand, particles as big as the wavelength of the incident light scatter more in certain directions than others. For a small molecule at low concentration, the intensity in the plane perpendicular to the incident polarization at an angle  $\theta$  is given by equation 1:

$$R(\theta) = MP(\theta)K^* \times c \quad (\text{eq. 1})$$

- $R(\theta)$  = Rayleigh ratio (describes the absolute intensity scattered at an angle  $\theta$  in excess of the light scattered by the pure solvent)  
 $M$  = molecular mass of the scattering molecule [g/mol]  
 $c$  = concentration of the molecule [g/ml]  
 $P(\theta)$  = form factor (ratio of scattered intensity at angle  $\theta$  to intensity at angle 0)  
 $K^*$  = optical constant (contains the refractive index of the solvent, Avogadro's number, the wavelength of the incident light and the specific refractive index increment of the sample molecule).

Thus by measuring the scattered intensity of a sample one can determine the average molecular weight. This technique, known as static light scattering, however does not provide any information about the size distribution of the scattering sample. Therefore a more versatile approach to the sizing problem is possible with the technique of dynamic light scattering. Instead of the average value of the scattered intensity at one time, one looks at the fluctuations of the scattered light as a function of time.

When several small particles are illuminated by a plane wave the scattered light can be seen as a “speckle” pattern. These speckles would, for example, appear on a screen placed behind a sample illuminated by a laser beam. The speckle pattern originates from interferences between the light scattered from different individual small particles, leading to a collection of bright (constructive interference) and dark (destructive interference) spots. In time, this intensity speckle pattern changes because of the positions of the small particles with respect to each other change. The particles move due to Brownian diffusion. These speckle fluctuations happen very fast for the smallest particles but more slowly for bigger particles. By observing the intensity fluctuations as a function of time (autocorrelation function  $G(\tau)$ ) one can find the translational diffusion constant of the particles ( $D_T$ ) that leads to those fluctuations.

$$G(\tau) = 1 + \exp(-2D_T q^2 \tau) \quad (\text{eq. 2})$$

$G(\tau)$  = autocorrelation of the scattered light as a function of time delay

$D_T$  = diffusion constant

$q$  = wave vector, depends on the scattering angle  $\theta$ , the refractive index  $n_s$  of the solvent, and the wavelength  $\lambda$  of the incident radiation ( $q = 4\pi n_s \sin(\theta/2) / \lambda$ )

$\tau$  = time delay

Knowing the translational diffusion coefficient and assuming a hard-sphere model for the particles, one can find their sizes with the Stokes-Einstein relationship:

$$D_T = \frac{k_B T}{6\pi\eta R_H} \quad (\text{eq. 3})$$

- $k_B$  = Boltzmann constant
- $T$  = absolute temperature in Kelvin
- $\eta$  = viscosity of the solvent
- $R_H$  = hydrodynamic radius of the average scattering particle

## 1.4 Aims and scope of the presented work

The difference in substrate acceptance between the human cytosolic TK and the HSV 1 TK is the crucial molecular basis for therapeutic applications linked to TK. HSV 1 TK is an important target in antiviral therapy and was established as major tool in many applications such as gene therapeutical treatment of cancer and AIDS, as control system for graft versus host disease in allogeneic bone marrow transplantation (allo-BMT) and AIDS vaccine, as expression reporter gene.

The clinical applications of enzyme-prodrug gene therapy based on HSV 1 TK has shown limitation linked to the enzyme-prodrug itself such as immunosuppressive side-effects of the GCV dosages required for tumor regression, immunogenicity caused by HSV 1 TK when the enzyme is used as control system or GCV incompatibility between CMV treatment and TK controlled GvHD approach in allo-BMT as well as the emerging problem of resistance by the antiviral therapy.

Based on this background it is clear that the system has to be optimized and new compounds have to be found to address the mentioned limitations. This optimization procedure requires new and more detailed structural knowledge on the human cytosolic enzyme as well as HSV 1 TK and its mutants.

Therefore, the first aim of the presented work was to solve the three-dimensional structure of the human cytosolic thymidine kinase, by X-ray crystallography. This would provide us new general information about TK's and helpful features to improve current therapeutic approaches. As an example, structural knowledge on the human TK1 would help to construct a humanized HSV 1 TK or human TK with characteristic of the wild type HSV 1 TK (broad substrate specificity) overcoming the problem of immunogenicity and improving the efficacy and safety of the allo-BMT strategy against tumor.

The second aim was to extend the structural knowledge on the binding mode of putative prodrug within the HSV 1 TK active site.

The third aim was the elucidation of the structure of HSV 1 TK mutants with an increased specificity towards the prodrug.

A further enzyme which was analyzed in this work was the enolase of *Alternaria alternata*. It was reported to be a major cause of mould allergies of the respiratory tract. Interestingly, the IgE's of previously sensitized patients showed cross-reactivity to enolase extracts from yeast, even though its enolase was reported to not be allergenic.

Therefore, the fourth aim was the crystallization and solution of the three dimensional structure of the enolase from *Alternaria alternata*, which would provide us with the knowledge of the enzyme's epitopes responsible for the allergic reaction with IgE.

To achieve these goals, different efforts were undertaken in optimizing expression, purification and crystallization systems.

In this context, **chapter 2** gives an overview of the applied methods in expression and purification of the human cytosolic thymidine kinase including site-directed mutagenesis.

**Chapter 3** is a presentation of all practical experiences made in the crystallization experiments with critical discussion of the results, which lead to the resolution of four new crystal structures of HSV 1 TK and mutants.

The new crystal structures are presented and discussed in detail in **chapter 4 and 5**.

## 1.5 References

1. Brown, D.G. *et al.* Crystal structures of the thymidine kinase from herpes simplex virus type-1 in complex with deoxythymidine and ganciclovir. *Nat Struct Biol* **2**, 876-81 (1995).
2. Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett* **368**, 289-92 (1995).
3. Gentry, G.A. Viral thymidine kinases and their relatives. *Pharmacol Ther* **54**, 319-55 (1992).
4. Black, M.E. & Hruby, D.E. Quaternary structure of vaccinia virus thymidine kinase. *Biochem Biophys Res Commun* **169**, 1080-6 (1990).
5. Bradshaw, H.D., Jr. Molecular cloning and cell cycle-specific regulation of a functional human thymidine kinase gene. *Proc Natl Acad Sci U S A* **80**, 5588-91 (1983).
6. Munch-Petersen, B., Tyrsted, G. & Cloos, L. Reversible ATP-dependent transition between two forms of human cytosolic thymidine kinase with different enzymatic properties. *J Biol Chem* **268**, 15621-5 (1993).
7. Munch-Petersen, B. *et al.* Altered substrate and inhibitor specificity of purified human adult thymidine kinases (TK2) from leukemic cells. *Adv Exp Med Biol* **370**, 253-6 (1994).
8. He, Q., Skog, S., Wu, C., Johansson, A. & Tribukait, B. Existence of phosphorylated and dephosphorylated forms of cytosolic thymidine kinase (TK1). *Biochim Biophys Acta* **1289**, 25-30 (1996).
9. Coppock, D.L. & Pardee, A.B. Control of thymidine kinase mRNA during the cell cycle. *Mol Cell Biol* **7**, 2925-32 (1987).
10. Hofbauer, R., Mullner, E., Seiser, C. & Wintersberger, E. Cell cycle regulated synthesis of stable mouse thymidine kinase mRNA is mediated by a sequence within the cDNA. *Nucleic Acids Res* **15**, 741-52 (1987).
11. Stewart, C.J., Ito, M. & Conrad, S.E. Evidence for transcriptional and post-transcriptional control of the cellular thymidine kinase gene. *Mol Cell Biol* **7**, 1156-63 (1987).

12. Wintersberger, E. Regulation and biological function of thymidine kinase. *Biochem Soc Trans* **25**, 303-8 (1997).
13. Ogris, E., Rotheneder, H., Mudrak, I., Pichler, A. & Wintersberger, E. A binding site for transcription factor E2F is a target for trans activation of murine thymidine kinase by polyomavirus large T antigen and plays an important role in growth regulation of the gene. *J Virol* **67**, 1765-71 (1993).
14. Dou, Q.P. *et al.* G1/S-regulated E2F-containing protein complexes bind to the mouse thymidine kinase gene promoter. *J Biol Chem* **269**, 1306-13 (1994).
15. Krek, W. *et al.* Negative regulation of the growth-promoting transcription factor E2F-1 by a stably bound cyclin A-dependent protein kinase. *Cell* **78**, 161-72 (1994).
16. Seiser, C., Teixeira, S. & Kuhn, L.C. Interleukin-2-dependent transcriptional and post-transcriptional regulation of transferrin receptor mRNA. *J Biol Chem* **268**, 13074-80 (1993).
17. Knofler, M., Waltner, C., Wintersberger, E. & Mullner, E.W. Translational repression of endogenous thymidine kinase mRNA in differentiating and arresting mouse cells. *J Biol. Chem.* **268**, 11409-11416 (1993).
18. Kauffman, M.G. & Kelly, T.J. Cell cycle regulation of thymidine kinase: residues near the carboxyl terminus are essential for the specific degradation of the enzyme at mitosis. *Mol Cell Biol* **11**, 2538-46 (1991).
19. Gross, M.K. & Merrill, G.F. Regulation of thymidine kinase protein levels during myogenic withdrawal from the cell cycle is independent of mRNA regulation. *Nucleic Acids Res* **16**, 11625-43 (1988).
20. Carozza, M.A. & Conrad, S.E. Regulation of thymidine kinase protein stability in serum-stimulated cells. *Cell Growth Differ* **5**, 901-8 (1994).
21. Sutterluety, H., Bartl, S., Karlseder, J., Wintersberger, E. & Seiser, C. Carboxy-terminal residues of mouse thymidine kinase are essential for rapid degradation in quiescent cells. *J Mol Biol* **259**, 383-92 (1996).
22. Kauffman, M.G., Rose, P.A. & Kelly, T.J. Mutations in the thymidine kinase gene that allow expression of the enzyme in quiescent (G0) cells. *Oncogene* **6**, 1427-35 (1991).

23. Sutterluety, H. & Seiser, C. Thymidine inhibits the growth-arrest-specific degradation of thymidine kinase protein in transfected L fibroblasts. *J Mol Biol* **265**, 153-60 (1997).
24. Kit, S. Thymidine kinase, DNA synthesis and cancer. *Molecular & Cellular Biochemistry* **11**, 161-82 (1976).
25. Wang, L. *et al.* Human thymidine kinase 2: molecular cloning and characterisation of the enzyme activity with antiviral and cytostatic nucleoside substrates. *FEBS Lett* **443**, 170-4 (1999).
26. Arner, E.S. & Eriksson, S. Mammalian deoxyribonucleoside kinases. *Pharmacol Ther* **67**, 155-86 (1995).
27. Johansson, M. & Karlsson, A. Cloning of the cDNA and chromosome localization of the gene for human thymidine kinase 2. *Journal of Biological Chemistry* **272**, 8454-8 (1997).
28. Lewis, W. & Dalakas, M.C. Mitochondrial toxicity of antiviral drugs. *Nature Medicine* **1**, 417-22 (1995).
29. Willecke, K., Teber, T., Kucherlapati, R.S. & Ruddle, F.H. Human mitochondrial thymidine kinase is coded for by a gene on chromosome 16 of the nucleus. *Somatic Cell Genetics* **3**, 237-45 (1977).
30. Kit, S. & Leung, W.C. Submitochondrial localization and characteristics of thymidine kinase molecular forms in parental and kinase-deficient HeLa cells. *Biochem Genet* **11**, 231-47 (1974).
31. Hallek, M., Wanders, L., Strohmeyer, S. & Emmerich, B. Thymidine kinase: a tumor marker with prognostic value for non- Hodgkin's lymphoma and a broad range of potential clinical applications. *Ann Hematol* **65**, 1-5 (1992).
32. Gronowitz, J.S., Hagberg, H., Kallander, C.F. & Simonsson, B. The use of serum deoxythymidine kinase as a prognostic marker, and in the monitoring of patients with non-Hodgkin's lymphoma. *British Journal of Cancer* **47**, 487-95 (1983).
33. O'Neill, K.L. *et al.* Breast tumour thymidine kinase levels and disease recurrence. *Medical Laboratory Sciences* **49**, 244-7 (1992).



34. Mansour, O., Motawi, T., Khaled, H. & el-Ahmady, O. Clinical value of thymidine kinase and tissue polypeptide specific antigen in breast cancer. *Disease Markers* **11**, 171-7 (1993).
35. Romain, S. *et al.* Prognostic value of cytosolic thymidine kinase activity as marker of proliferation in breast cancer. *Int. J. Cancer* **61**, 7-12 (1995).
36. Sakamoto, S., Ebuchi, M. & Iwama, T. Relative activities of thymidylate synthetase and thymidine kinase in human mammary tumours. *Anticancer Res* **13**, 205-7 (1993).
37. Yusa, T. *et al.* [A study of thymidine kinase activity in lung cancer tissue]. *Nihon Kyobu Shikkan Gakkai Zasshi* **32**, 211-5 (1994).
38. Sun, X.L. *et al.* Modulation of the cytotoxicity of 3'-azido-3'-deoxythymidine and methotrexate after transduction of folate receptor cDNA into human cervical carcinoma: identification of a correlation between folate receptor expression and thymidine kinase activity. *Cancer Res* **59**, 940-6 (1999).
39. He, Q., Skog, S., Wang, N., Eriksson, S. & Tribukait, B. Characterization of a peptide antibody against a C-terminal part of human and mouse cytosolic thymidine kinase, which is a marker for cell proliferation. *Eur J Cell Biol* **70**, 117-24 (1996).
40. Mineishi, S. *et al.* Co-expression of the herpes simplex virus thymidine kinase gene potentiates methotrexate resistance conferred by transfer of a mutated dihydrofolate reductase gene. *Gene Therapy* **4**, 570-6 (1997).
41. Wild, K., Bohner, T., Folkers, G. & Schulz, G.E. The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue. *Protein Sci* **6**, 2097-106 (1997).
42. Bennett, M.S. *et al.* Structure to 1.9 Å resolution of a complex with herpes simplex virus type-1 thymidine kinase of a novel, non-substrate inhibitor: X-ray crystallographic comparison with binding of aciclovir. *FEBS Lett* **443**, 121-5 (1999).
43. Champness, J.N. *et al.* Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands. *Proteins* **32**, 350-61 (1998).

44. Halpern, M.E. & Smiley, J.R. Effects of deletions on expression of the herpes simplex virus thymidine kinase gene from the intact viral genome: the amino terminus of the enzyme is dispensable for catalytic activity. *Journal of Virology* **50**, 733-8 (1984).
45. Robbins, J., Dilworth, S.M., Laskey, R.A. & Dingwall, C. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. *Cell* **64**, 615-23 (1991).
46. Degreve, B., Johansson, M., De Clercq, E., Karlsson, A. & Balzarini, J. Differential intracellular compartmentalization of herpetic thymidine kinases (TKs) in TK gene-transfected tumor cells: molecular characterization of the nuclear localization signal of herpes simplex virus type 1 TK. *Journal of Virology* **72**, 9535-43 (1998).
47. Dreusicke, D. & Schulz, G.E. The glycine-rich loop of adenylate kinase forms a giant anion hole. *FEBS Letters* **208**, 301-4 (1986).
48. Waldman, A.S., Haeusslein, E. & Milman, G. Purification and characterization of herpes simplex virus (type 1) thymidine kinase produced in *Escherichia coli* by a high efficiency expression plasmid utilizing a lambda PL promoter and cI857 temperature-sensitive repressor. *Journal of Biological Chemistry* **258**, 11571-5 (1983).
49. Chen, M.S., Walker, J. & Prusoff, W.H. Kinetic studies of herpes simplex virus type 1-encoded thymidine and thymidylate kinase, a multifunctional enzyme. *Journal of Biological Chemistry* **254**, 10747-53 (1979).
50. Michael, M., Fetzter, J. & Folkers, G. Site-directed mutagenesis clarifies the substrate position within the three-dimensional model of the active site of herpes simplex virus type- 1 thymidine kinase. *Eur J Biochem* **226**, 219-26 (1994).
51. Kussmann-Gerber, S. *et al.* Interaction of the Recombinant Herpes Simplex Virus Type 1 Thymidine Kinase with Thymidine and Aciclovir: A Kinetic Study. *Nucleosides & Nucleotides* **18**, 311-330 (1999).
52. Perozzo, R., Folkers, G. & Scapozza, L. manuscript in preparation. (1999).

53. Chen, M.S. & Prusoff, W.H. Association of thymidylate kinase activity with pyrimidine deoxyribonucleoside kinase induced by herpes simplex virus. *Journal of Biological Chemistry* **253**, 1325-7 (1978).
54. Smee, D.F., Martin, J.C., Verheyden, J.P. & Matthews, T.R. Anti-herpesvirus activity of the acyclic nucleoside 9-(1,3-dihydroxy-2-propoxymethyl)guanine. *Antimicrobial Agents & Chemotherapy* **23**, 676-82 (1983).
55. Larder, B.A. & Darby, G. Properties of a novel thymidine kinase induced by an acyclovir-resistant herpes simplex virus type 1 mutant. *Journal of Virology* **42**, 649-58 (1982).
56. Larder, B.A., Cheng, Y.C. & Darby, G. Characterization of abnormal thymidine kinases induced by drug-resistant strains of herpes simplex virus type 1. *Journal of General Virology* **64 Pt 3**, 523-32 (1983).
57. Keller, P.M. *et al.* Enzymatic phosphorylation of acyclic nucleoside analogs and correlations with antiherpetic activities. *Biochemical Pharmacology* **30**, 3071-7 (1981).
58. Pilger, B.D. & Scapozza, L. manuscript in preparation. (1999).
59. Black, M.E. & Loeb, L.A. 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**, 11618-26 (1993).
60. Pilger, B.D. *et al.* Substrate diversity of herpes simplex virus thymidine kinase. Impact Of the kinematics of the enzyme [In Process Citation]. *J Biol Chem* **274**, 31967-73 (1999).
61. Munir, K.M., French, D.C., Dube, D.K. & Loeb, L.A. Permissible amino acid substitutions within the putative nucleoside binding site of herpes simplex virus type 1 encoded thymidine kinase established by random sequence mutagenesis [corrected] [published erratum appears in J Biol Chem 1992 Jul 25;267(21):15258]. *Journal of Biological Chemistry* **267**, 6584-9 (1992).
62. Kussmann-Gerber, S., Kuonen, O., Folkers, G., Pilger, B.D. & Scapozza, L. Drug resistance of herpes simplex virus type 1--structural considerations at

- the molecular level of the thymidine kinase. *Eur J Biochem* **255**, 472-81 (1998).
63. Gompels, U. & Minson, A. The properties and sequence of glycoprotein H of herpes simplex virus type 1. *Virology* **153**, 230-47 (1986).
  64. Spadari, S. *et al.* L-thymidine is phosphorylated by herpes simplex virus type 1 thymidine kinase and inhibits viral growth. *Journal of Medicinal Chemistry* **35**, 4214-20 (1992).
  65. Marquez, V.E. *et al.* Nucleosides with a twist. Can fixed forms of sugar ring pucker influence biological activity in nucleosides and oligonucleotides? *J Med Chem* **39**, 3739-47 (1996).
  66. Hannigan, B.M., Barnett, Y.A., Armstrong, D.B., McKelvey-Martin, V.J. & McKenna, P.G. Thymidine kinases: the enzymes and their clinical usefulness. *Cancer Biother* **8**, 189-97 (1993).
  67. Furman, P.A. *et al.* Phosphorylation of 3'-azido-3'-deoxythymidine and selective interaction of the 5'-triphosphate with human immunodeficiency virus reverse transcriptase. *Proc Natl Acad Sci U S A* **83**, 8333-7 (1986).
  68. Germann, C., Shields, A.F., Grierson, J.R., Morr, I. & Haberkorn, U. 5-Fluoro-1-(2'-deoxy-2'-fluoro-beta-D-ribofuranosyl) uracil trapping in Morris hepatoma cells expressing the herpes simplex virus thymidine kinase gene. *J Nucl Med* **39**, 1418-23 (1998).
  69. Culver, K.W. *et al.* In vivo gene transfer with retroviral vector-producer cells for treatment of experimental brain tumors [see comments]. *Science* **256**, 1550-2 (1992).
  70. Moolten, F.L. Drug sensitivity ("suicide") genes for selective cancer chemotherapy. *Cancer Gene Ther* **1**, 279-87 (1994).
  71. Bonnekoh, B. *et al.* Inhibition of melanoma growth by adenoviral-mediated HSV thymidine kinase gene transfer in vivo. *J Invest Dermatol* **104**, 313-7 (1995).
  72. Chen, S.H., Shine, H.D., Goodman, J.C., Grossman, R.G. & Woo, S.L. Gene therapy for brain tumors: regression of experimental gliomas by adenovirus-mediated gene transfer in vivo. *Proc Natl Acad Sci U S A* **91**, 3054-7 (1994).

73. Boviatsis, E.J. *et al.* Long-term survival of rats harboring brain neoplasms treated with ganciclovir and a herpes simplex virus vector that retains an intact thymidine kinase gene. *Cancer Res* **54**, 5745-51 (1994).
74. Perez-Cruet, M.J. *et al.* Adenovirus-mediated gene therapy of experimental gliomas. *J Neurosci Res* **39**, 506-11 (1994).
75. Takamiya, Y. *et al.* Gene therapy of malignant brain tumors: a rat glioma line bearing the herpes simplex virus type 1-thymidine kinase gene and wild type retrovirus kills other tumor cells. *J Neurosci Res* **33**, 493-503 (1992).
76. Ezzeddine, Z.D. *et al.* Selective killing of glioma cells in culture and in vivo by retrovirus transfer of the herpes simplex virus thymidine kinase gene. *New Biol* **3**, 608-14 (1991).
77. Moolten, F.L. & Wells, J.M. Curability of tumors bearing herpes thymidine kinase genes transferred by retroviral vectors. *J Natl Cancer Inst* **82**, 297-300 (1990).
78. Moolten, F.L., Wells, J.M., Heyman, R.A. & Evans, R.M. Lymphoma regression induced by ganciclovir in mice bearing a herpes thymidine kinase transgene. *Hum Gene Ther* **1**, 125-34 (1990).
79. Lechanteur, C. *et al.* HSV-1 thymidine kinase gene therapy for colorectal adenocarcinoma- derived peritoneal carcinomatosis. *Gene Ther* **4**, 1189-94 (1997).
80. Caruso, M. *et al.* Regression of established macroscopic liver metastases after in situ transduction of a suicide gene. *Proc Natl Acad Sci U S A* **90**, 7024-8 (1993).
81. Sacco, M.G. *et al.* Local regression of breast tumors following intramammary ganciclovir administration in double transgenic mice expressing neu oncogene and herpes simplex virus thymidine kinase. *Gene Ther* **2**, 493-7 (1995).
82. Grignet-Debrus, C. & Calberg-Bacq, C.M. Potential of Varicella zoster virus thymidine kinase as a suicide gene in breast cancer cells. *Gene Ther* **4**, 560-9 (1997).
83. Raffel, C. *et al.* Gene therapy for the treatment of recurrent pediatric malignant astrocytomas with in vivo tumor transduction with the herpes

- simplex thymidine kinase gene/ganciclovir system. *Hum Gene Ther* **5**, 863-90 (1994).
84. Oldfield, E.H. *et al.* Gene therapy for the treatment of brain tumors using intra-tumoral transduction with the thymidine kinase gene and intravenous ganciclovir. *Hum Gene Ther* **4**, 39-69 (1993).
  85. Kun, L.E. *et al.* Stereotactic injection of herpes simplex thymidine kinase vector producer cells (PA317-G1Tk1SvNa.7) and intravenous ganciclovir for the treatment of progressive or recurrent primary supratentorial pediatric malignant brain tumors. *Hum Gene Ther* **6**, 1231-55 (1995).
  86. Culver, K.W. *et al.* Gene therapy for the treatment of malignant brain tumors with in vivo tumor transduction with the herpes simplex thymidine kinase gene/ganciclovir system. *Human Gene Therapy* **5**, 343-79 (1994).
  87. Freeman, S.M. *et al.* The treatment of ovarian cancer with a gene modified cancer vaccine: a phase I study. *Hum Gene Ther* **6**, 927-39 (1995).
  88. Ram, Z. *et al.* Therapy of malignant brain tumors by intratumoral implantation of retroviral vector-producing cells [see comments]. *Nat Med* **3**, 1354-61 (1997).
  89. Encell, L.P., Landis, D.M. & Loeb, L.A. Improving enzymes for cancer gene therapy. *Nature Biotechnology* **17**, 143-7 (1999).
  90. Nishihara, E. *et al.* Treatment of thyroid carcinoma cells with four different suicide gene/prodrug combinations in vitro. *Anticancer Res* **18**, 1521-5 (1998).
  91. Rogulski, K.R., Kim, J.H., Kim, S.H. & Freytag, S.O. Glioma cells transduced with an Escherichia coli CD/HSV-1 TK fusion gene exhibit enhanced metabolic suicide and radiosensitivity. *Hum Gene Ther* **8**, 73-85 (1997).
  92. Hoganson, D.K., Matsui, H., Batra, R.K. & Boucher, R.C. Toxin gene-mediated growth inhibition of lung adenocarcinoma in an animal model of pleural malignancy. *Hum Gene Ther* **9**, 1143-56 (1998).
  93. Haberkorn, U. *et al.* Monitoring gene therapy with cytosine deaminase: in vitro studies using tritiated-5-fluorocytosine. *J Nucl Med* **37**, 87-94 (1996).

94. Uckert, W. *et al.* Double suicide gene (cytosine deaminase and herpes simplex virus thymidine kinase) but not single gene transfer allows reliable elimination of tumor cells in vivo. *Hum Gene Ther* **9**, 855-65 (1998).
95. Freytag, S.O., Rogulski, K.R., Paielli, D.L., Gilbert, J.D. & Kim, J.H. A novel three-pronged approach to kill cancer cells selectively: concomitant viral, double suicide gene, and radiotherapy [see comments]. *Hum Gene Ther* **9**, 1323-33 (1998).
96. Aghi, M., Kramm, C.M., Chou, T.C., Breakefield, X.O. & Chiocca, E.A. Synergistic anticancer effects of ganciclovir/thymidine kinase and 5-fluorocytosine/cytosine deaminase gene therapies [see comments]. *J Natl Cancer Inst* **90**, 370-80 (1998).
97. Blackburn, R.V., Galoforo, S.S., Corry, P.M. & Lee, Y.J. Adenoviral-mediated transfer of a heat-inducible double suicide gene into prostate carcinoma cells. *Cancer Res* **58**, 1358-62 (1998).
98. Bonini, C. *et al.* HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia [see comments]. *Science* **276**, 1719-24 (1997).
99. Wold, F. *The Enzymes*, 499-583 (Academic Press, New York, 1970).
100. Breitenbach, M. *et al.* Enolases are highly conserved fungal allergens. *Int Arch Allergy Immunol* **113**, 114-7 (1997).
101. McPherson, A. A brief history of protein crystal growth. *J Cryst Growth* **110**, 1-10 (1991).
102. Hall, S.S. Protein images update natural history [news] [see comments]. *Science* **267**, 620-4 (1995).
103. <http://www.rcsb.org/pdb/index.html>. Protein Data Bank. (Research Collaboratory for Structural Bioinformatics (RCSB), online).
104. Ducruix, A. & Giegé, R. *Crystallization of nucleic acids and proteins : a practical approach*, xxiv, 331 (IRL Press at Oxford University Press, Oxford [England] ; New York, 1992).
105. Budge. *Sitzungsber Niederrh Ges Natur Heilkunde* (1850).
106. Hünefeld. *Der Chemismus in der tierischen Organisation*. Leipzig (1840).
107. Kölliker. *Z Wissensch Zoologie* **1**, 266 (1849).

108. Leydig. *Z Wissensch Zoologie* **1**, 116 (1849).
109. Reichert, K.E. *Müllers Arch Anat Physiol Wissensch Medizin* , 197-251 (1849).
110. Fünke, O. *Z Rat Medizin (NF)* **1**, 185; 198, 288 (1851).
111. Sumner, J.B. *J Biol Chem* **69**, 435-441 (1926).
112. Abel, J.J. & et al. *J. Pharmacol Exptl Therap* **31**, 65 (1927).
113. Laufberger, M. *Bull Soc Chim Biol* **19**, 1575 (1937).
114. Northrop, J.H., Kunitz, M. & Herriott, R.M. *Crystalline Enzymes*, (Columbia University Press, New York, 1946).
115. Sumner, J.B. & Dounce, A.L. *J.Biol.Chem.* **121**, 417-424 (1937).
116. McPherson, A. Current approaches to macromolecular crystallization. *Eur J Biochem* **189**, 1-23 (1990).
117. Drenth, J. *Principles of protein x-ray crystallography*, xv, 341 (Springer, New York, 1999).
118. McPherson, A., Malkin, A.J. & Kuznetsov, Y.G. The science of macromolecular crystallization. *Structure* **3**, 759-68 (1995).
119. Skelly, J.V. & Madden, C.B. Overexpression, isolation, and crystallization of proteins. *Methods Mol Biol* **56**, 23-53 (1996).
120. Chernov, A.A. Crystal Growth. in *Modern Crystallography III* (Springer-Verlag, Berlin, 1984).
121. Dounce, A.L. & Allen, P.Z. Fifty years later: recollections of the early days of protein crystallization. *Trends Biochem Sci* **13**, 317-20 (1988).
122. Althoff, S.M. Bachelors Thesis., University of Illinois (1987).
123. Althoff, S., Zambrowicz, B., Liang, P., Glaser, M. & Phillips, G.N., Jr. Crystallization and preliminary X-ray analysis of Escherichia coli adenylate kinase [letter]. *J Mol Biol* **199**, 665-6 (1988).
124. Chayen, N.E. New developments of the IMPAX small-volume automated crystallization system. *Acta Cryst* **D50**, 456-458 (1994).
125. Zeppezauer, M., Eklund, H. & Zeppezauer, E.S. Micro diffusion cells for the growth of single protein crystals by means of equilibrium dialysis. *Arch Biochem Biophys* **126**, 564-73 (1968).



126. Carter, C.W.J., Baldwin, E.T. & Frick, L. Statistical design of experiments for protein crystal growth and the use of a precrystallization assay. *J. of Cryst. Growth* **90**, 60-73 (1988).
127. Garman, E.F. & Mitchell, E.P. Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *J Appl Cryst* **29**, 584-587 (1996).
128. Rupp, B. Crystallography 101. (<http://www-structure.llnl.gov/Xray/101index.html>, online).
129. Wilson, W.W. Monitoring crystallization experiments using dynamic light scattering: Assaying and monitoring protein crystallization in solution. *METHODS: A Companion to Methods in Enzymology* **1**, 110-117 (1990).
130. Mikol, V., Hirsch, E. & Giege, R. Diagnostic of precipitant for biomacromolecule crystallization by quasi-elastic light-scattering. *J Mol Biol* **213**, 187-95 (1990).
131. D'Arcy, A. Crystallizing proteins - a rational approach? *Acta Cryst D* **50**, 469-471 (1994).
132. Baldwin, E.T., Crumly, K.V. & Carter, C.W. Practical, rapid screening of protein crystallization conditions by dynamic light scattering. *Biophys. J.* **49**, 47 (1986).
133. Ferre-D'Amare, A.R. & Burley, S.K. Dynamic light scattering in evaluating crystallizability of macromolecules. *Methods in Enzymology* **276**, 157-166 (1997).
134. Chene, C. *et al.* Crystallization of the complex of human IFN-gamma and the extracellular domain of the IFN-gamma receptor. *Proteins* **23**, 591-4 (1995).
135. Mittl, P.R. *et al.* Structure of recombinant human CPP32 in complex with the tetrapeptide acetyl-Asp-Val-Ala-Asp fluoromethyl ketone. *J Biol Chem* **272**, 6539-47 (1997).
136. Kadima, W., McPherson, A., Dunn, M.F. & Jornak, F.A. Characterization of precrystallization aggregation of canavalin by dynamic light scattering. *Biophys J* **57**, 125-32 (1990).

Seite Leer /  
Blank leaf

## CHAPTER 2

### **The Human cytosolic thymidine kinase (TK1):**

Preparation & characterization of the protein for crystallization  
experiments

## Abstract

Based on the necessity to know more about the structure at atomic resolution of the human cytosolic thymidine kinase the aim of the work described in this chapter was to produce large amounts of high quality protein samples allowing the crystallization and solution of the structure by X-ray crystallography. Therefore, different efforts were done to achieve this goal.

During expression and purification we recognized different properties of the protein to be potential negative factors for crystallization experiments. These obstacles were stepwise eliminated before we could really start with crystallization itself. First attempts to get pure recombinant cytosolic TK using thrombin-cleavable Fu-pro-strategy produced insufficient amounts of pure protein. Impurities were co-purified and the protein was present in two different length. Furthermore, the protein tent to aggregate preventing to reach the needed high concentrated solutions. The co-purified proteins were eliminated by improving the procedure with additional washing steps. The problem of solubility and quantity was addressed using a his-tag protein. Hence larger amounts of protein could be purified, but the quality of the protein was not sufficient for crystallization trials. The protein was more inhomogeneous than previously encountered and let hypothesize that the problem would lye in a heterogene expression in *E.coli*. To overcome this problem a new strategy was developed for the construction of a deletion mutant missing the c-terminal 40 amino acid residues but with conserved wild type activity. The elimination of this part of the protein, being important as recognition tail for protein degradation during the cell cycle but not for enzymatic activity, was expected to lead to a more homogeneous and stable product. The deletion mutant was produced by site directed mutagenesis using the mis-match primer method. The mutant could be purified as fusion protein as well as cleaved enzyme and both showed full enzymatic activity. The resulting batches of protein had the suitable quality and homogeneity while the quantity was still low for crystallization trials. To improve the amount of protein of good quality the expression was performed using a new bacterial strain, which was especially created for the expression of eucaryotic proteins in *E.coli*. The quality and quantity of the expressed protein were improved and the crystallization experiments could be performed and resulted in fast small crystals of  $50 \times 50 \times 50 \mu\text{m}^3$ .

## 2.1 Introduction

Thymidine kinase (TK) is the key enzyme of the pyrimidine salvage pathway. It has been found in all species with the exception of fungi. The enzyme catalyzes the monophosphorylation of the 5'-hydroxy-group of thymidine using ATP as phosphate donor. As previously described in chapter 1 the human cytosolic TK1 belongs to the short TKs and to date no structural information is available for them, besides a three-dimensional model, which was constructed using computer aided structure prediction techniques based on comparisons with isofunctional proteins<sup>1</sup>. This model showed that regions from amino acid 81 to 94 and the FQRKP region at position 128 of human TK1 play a critical role for substrate specificity. Additionally a domain important for dimerization of the human TK1 which is directly linked to the biological function of the enzyme has been identified, offering a very economical way to modulate the TK1 activity at the post-translational level only by modification, without the need for specific proteolytic degradation. [Hofbauer, 1986, E. NAR, 741-752; Hofbauer, 1993, Methods in molecular genetics 2 2-25]

Limitations encountered in clinical application of HSV 1 TK in enzyme prodrug gene-therapy of different types of cancer because of the immunosuppressive side-effects of the GCV dosages required for tumor regression<sup>2-4</sup> show the necessity of new structural knowledge on the human enzyme. These structural details would allow to design new tailor made substrates and mutants for this kind of therapy based on the available commonness and differences to known structures of wild type and ACV/GCV resistant mutants of HSV 1 TK.

Several limitations are also known in clinical study of HSV1 TK gene for the control of GvHD after allogeneic bone marrow transplantation (allo-BMT), where donor lymphocytes carrying the TK gene play a central therapeutic role in both graft-versus-tumor (GvT) activity and immune reconstitution<sup>5,6</sup>. First patients developed specific CD8 cytotoxic immune response against the viral TK gene, which resulted in the immediate elimination of HSV 1 TK transduced cells<sup>7</sup>. This acute side effect due to immunogenicity of HSV1 TK limits the success of the clinically used anticancer treatment based on the GvT activity. Structural knowledge on the human TK1 would help to construct a humanized HSV 1 TK or human TK with characteristic of the wild type HSV 1 TK (broad substrate specificity) overcoming

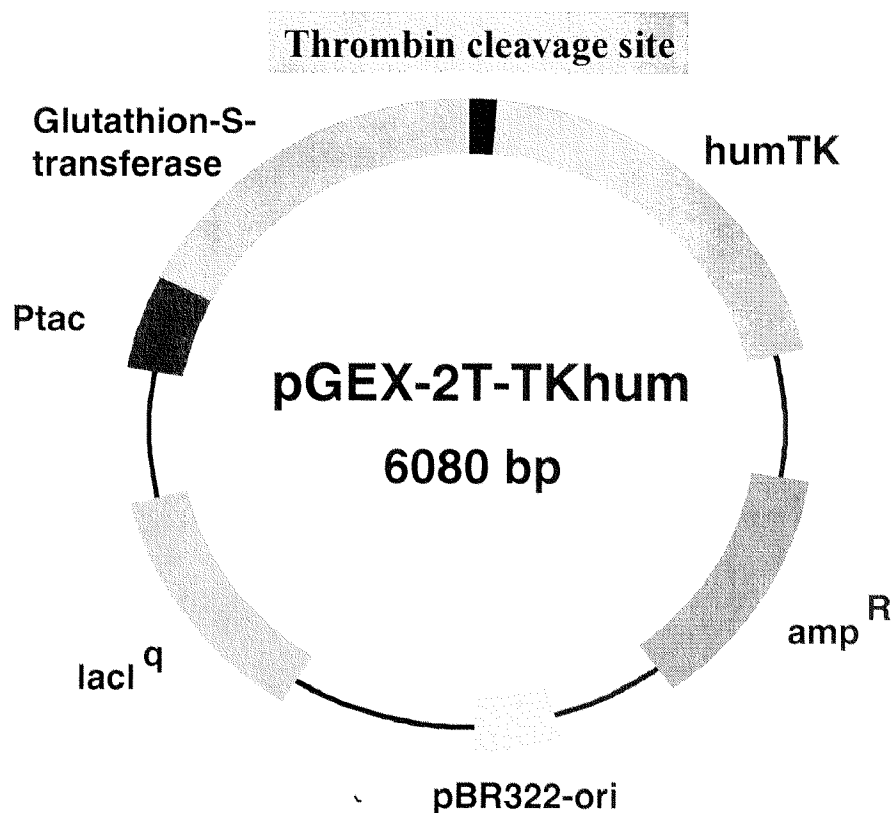
the problem of immunogenicity and improving the efficacy and safety of the allo-BMT strategy against tumor. A further limitation of this strategy is based on the fact that the prodrug GCV is also an essential antiviral tool for the control of CMV infections, a disease frequently associated with allo-BMT. To avoid the elimination of HSV 1 TK transduced lymphocytes, the actual immune effectors of anti-viral and anti-tumor activity, GCV can not be used in patients treated with HSV 1 TK transduced lymphocytes. In the frequent case of foscarnet-resistant CMV infections, patients have to be treated with GCV resulting in the elimination of transduced donor lymphocytes and reduced GvT activity of allo-BMT. The development of an enzyme/prodrug system which is characterized by reduced immunogenicity and restricted substrate specificity of the enzyme will allow GCV treatment of allo-BMT patients with CMV infections, thus enhancing the efficacy and safety of GvT therapy.

Based on this background the necessity for new structural knowledge on this field becomes evident. Therefore, the aim of the presented work was to solve the three-dimensional structure of the human cytosolic thymidine kinase, which would provide us new general information about TK's and helpful features to improve current therapeutic approaches. In this chapter steps such as cloning, subcloning, different expression systems, as well as purification procedures and preliminary protein analysis tools, which are necessary for crystallographic studies are described.

## 2.2 Cloning & Sequencing

### 2.2.1 The recombinant TK1 from human lymphocytes (Clone 34)

Starting material for the following expression, purification and crystallization experiments as well as for the cloning experiments of the  $\Delta 40$  TK1 was the pGEX-2T-TKhum, which was constructed by Fetzer J. and Bohner T. according to standard methods<sup>8</sup>. The *KpnI-HindIII* fragment from pGEM-TK (rec TK1 from lymphocytes, generously provided by R. Hofbauer, Vienna) was subcloned into pBluescript. The resulting pBluescript-TK was cleaved by *BamHI* and *EcoRI* and the fragment inserted into the *BamHI-EcoRI* cleaved pGEX-2T yielding pGEX-2T-TKhum. Direct subcloning from pGEM-TK into pGEX-2T in compliance with the correct open reading frame was not possible due to the limited polycloning site in pGEX-2T.

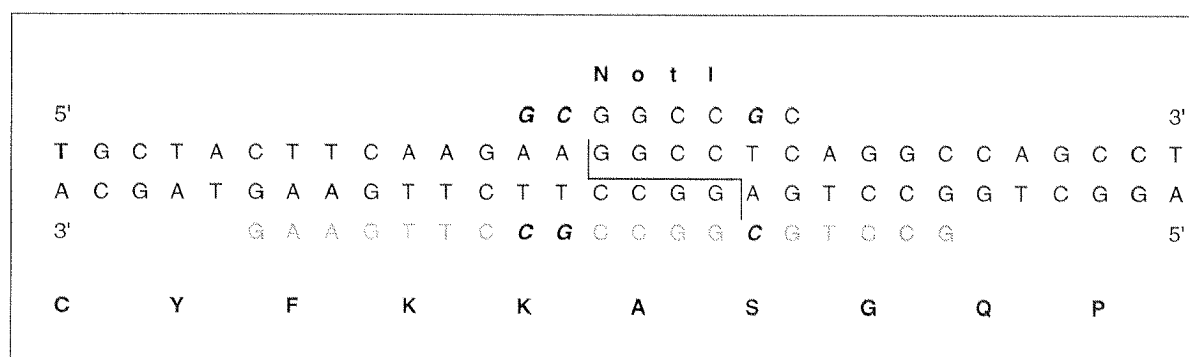


**Fig.1 The expression plasmid pGEX-2T-Tkhum** The plasmid contains a gene under the control of a ptac promoter which is made up of the codons for the glutathione binding domain of glutathione-S-transferase followed by the codons for a thrombin recognition site and the open reading frame of human cytosolic thymidine kinase.

### 2.2.2 The deletion mutant $\Delta 40\text{TK1}$

Different crystallization experiments of the human cytosolic TK had been carried out before with the purified inhomogeneous protein without any success [see Chapter 3]. The failure was thought to be mainly caused by the inhomogeneity caused by proteolytic cleavage of the enzyme during expression and the high grade of aggregation of the cleaved protein. It is known that minimizing the termini of a protein is very often crucial for successful crystallization<sup>9</sup>. The experiments with the HSV1 TK, performed in our laboratory, had shown in the past that an N-terminal truncated TK exhibited the same kinetic properties like the full length enzyme<sup>10</sup>. Studies on the cell cycle regulation of thymidine kinase revealed the residues near the carboxyl terminus to be essential for the specific degradation of the enzyme at mitosis<sup>11-13</sup>. In an earlier work it was demonstrated that residues in an unstructured region at the extreme carboxyl terminus of the protein are important for determining its proteolytic susceptibility in *E.coli*<sup>14</sup>.

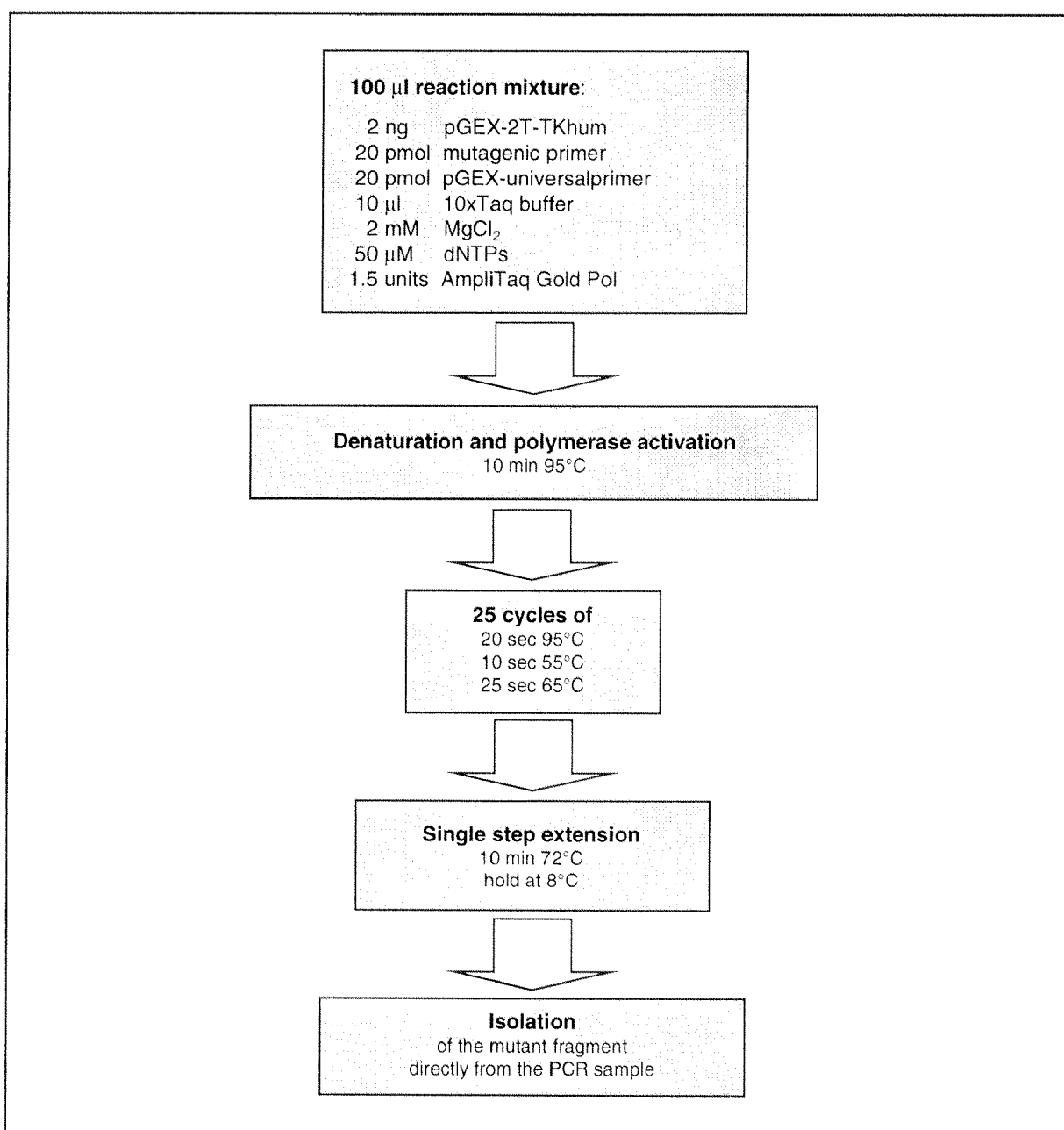
Based on these observations a new strategy was developed with the aim to improve the stability and the homogeneity of the enzyme for crystallization studies by creating a full enzymatic active deletion mutant of the human TK lacking the c-terminal 40 amino acids. Therefore, we had to introduce an additional restriction site for *NotI* in the coding sequence in order to be able to clone the desired fragment into the expression vector pGEX-6P-2.



**Fig. 2 PCR strategy for the creation of the deletion mutant  $\Delta 40\text{TK1}$ .** The recognition site for *NotI* is shown on the top of the cDNA segment of clone 34 from bp 563 to 592. The mutagenesis primer for the PCR reaction is on the bottom of the cDNA segment. The mismatches are in italic bold.



The fragment for the deletion mutant  $\Delta 40TK1$  was generated by site-directed mutagenesis using a standard PCR protocol, where the vector pGEX-2T-TKhum (template) was amplified using an universal pGEX-Primer (N-Term) and a mutagenic Primer (*NotI* $\Delta 40$ ) containing three mismatches (fig. 2). The general strategy of amplification is outlined in fig. 3.

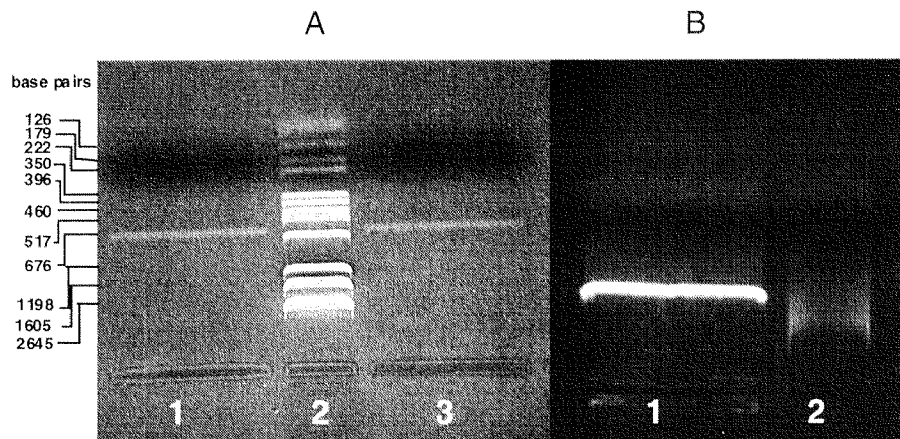


**Fig. 3** Scheme of the standard PCR procedure used for the amplification of the mutant fragment for the deletion mutant  $\Delta 40TK1$

The resulting fragment comprising the mutation was isolated directly from the PCR reaction mixture using the QIAquick<sup>®</sup> PCR purification kit (Qiagen).

After digestion with the restriction enzymes *Bam*HI and *Not*I during 1.5 h the mutant fragment and the digested pGEX-6P-2 were isolated from a 2% resp. 0.6% agarose gel and purified using the QIAquick gel extraction kit.

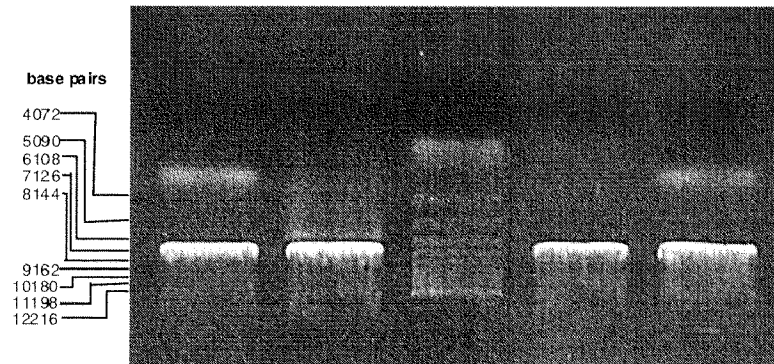
The digested fragments are shown in fig. 4:



**Fig. 4 Agarose gels of the digested fragments** **A:** 2% Agarose gel with two digested mutant fragments from the PCR reaction: Lane 1 and 3 Amplified mutant fragment of the human TK; Lane 2 D15 DNA-marker, Novex; **B:** 0.6% Agarose gel with the *Bam*HI/*Not*I digested pGEX-6P-2 vector (Lane 1) and Kb ladder (Lane2).

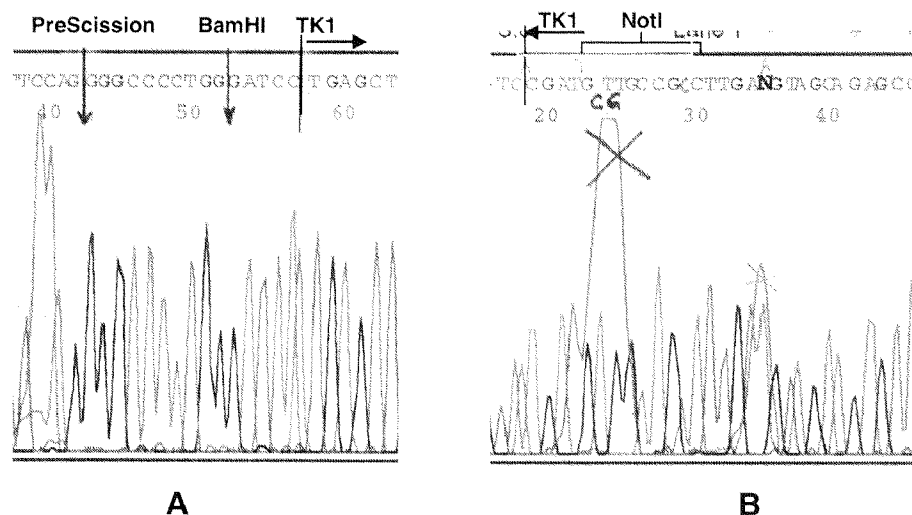
The ligation was performed over night at 4°C according to the standard protocol supplied by Promega<sup>®</sup> for the T4 DNA Ligase. These samples were then used to transform competent *E.coli* strains BL21 for the selection of colonies on Ampicillin-Agar plates. Six colonies which grew over night at 37°C were further checked for the presence of the new plasmid pGEX-6P-2-Δ40TK1 by PCR using the standard amplification protocol as described in fig.3 with the same primers. Two colonies contained the newly ligated expression vector and were used to grow over night cultures. The amplified vector was isolated from the bacteria using the QIAprep Spin Plasmid Kit (Qiagen) and used for the control digestion and sequence analysis.

The plasmids isolated from both cultures could clearly be linearized by the control digestion with *Not*I or *Bam*HI as shown in fig. 5:



**Fig.5 Control digestion of the isolated plasmid revealed the presence of a *NotI* and *BamHI* restriction site** Lane 1: *NotI* digestion of colony 1; Lane 2: *NotI* digestion of colony 4; Lane 3: Kb ladder; Lane 4: *BamHI* digestion of colony 1; Lane 5: *BamHI* digestion of colony 4

To verify that the correct fragment was present and no additional mutation and frameshift occurred, the new ligated expression vector was sequenced using the fluorescence (dye-) labeled dideoxy terminator method (ABI PRISM™ 310) as described by<sup>15</sup>. For the cycle sequencing reaction the pGEX universal primer were used. In both isolated vectors the presence of the right fragment sequence could be demonstrated and no additional mutations or frameshift were detected as illustrated in fig.6.



**Fig. 6 Sequence verification of the established deletion mutant  $\Delta 40TK1$**  The electropherograms are extracts from the sequence of colony 1 showing in A (pGEX-forward) the presence of the coding region for the PreScission protease recognition site, the restriction site for *BamHI* and the star of the coding sequence for TK1 in frame. In B the electropherogram (pGEX-backward) indicates the presence of the *NotI* restriction site at the end of the open reading frame for the deletion mutant.

The complete DNA sequence in pGEX-6P-2 coding for the deletion mutant  $\Delta 40TK1$  and the translated amino acid sequence are presented in fig. 7. The amino acid sequence is almost identical with that of clone 34 with two main changes caused by the mutation. One difference is found at the N-terminus where due to the PreScission cleavage site, three additional amino acids are present.

```

1  GATCCATGAGCTGCATTAACCTGCCCACTGTGCTGCCCGGCTCCCCCAGC
  +-----+-----+-----+-----+-----+-----+
a  G P L G S M S C I N L P T V L P G S P S -
  AAGACCCGGGGCAGATCCAGGTGATTCTCGGGCCGATGTTCTCAGGAAAAAGCACAGAG
61  +-----+-----+-----+-----+-----+-----+ 120
  TTCTGGGCCCCCGTCTAGGTCCACTAAGAGCCCGGCTACAAGAGTCCCTTTTTCGTGTCTC
a  K T R G Q I Q V I L G P M F S G K S T E -
  TTGATGAGACGCGTCCGTTCGCTTCCAGATTGCTCAGTACAAGTGCCTGGTGATCAAGTAT
121  +-----+-----+-----+-----+-----+-----+ 180
  AACTACTCTGCGCAGGCAGCGAAGGTCTAACGAGTCATGTTACCGGACCACTAGTTCATA
a  L M R R V R R F Q I A Q Y K C L V I K Y -
  GCCAAAGACACTCGCTACAGCAGCAGCTTCTGCACACATGACCGGAACACCATGGAGGCG
181  +-----+-----+-----+-----+-----+-----+ 240
  CGGTTTCTGTGAGCGATGTCTGTCGTCGAAGACGTGTGTACTGGCCTTGTGGTACCTCCGC
a  A K D T R Y S S S F C T H D R N T M E A -
  CTGCCCCGCTGCCTGCTCCGAGACGTGGCCCCAGGAGGCCCTGGGCGTGGCTGTCATAGGC
241  +-----+-----+-----+-----+-----+-----+ 300
  GACGGGCGGACGGACGAGGCTCTGCACCGGTCTCCGGGACCCGCACCGACAGTATCCG
a  L P A C L L R D V A Q E A L G V A V I G -
  ATCGACGAGGGGACGTTTTTCCCTGACATCGTGGAGTTCTGCGAGGCCATGGCCAACGCC
301  +-----+-----+-----+-----+-----+-----+ 360
  TAGCTGCTCCCCGTCAAAAAGGGACTGTAGCACCTCAAGACGCTCCGGTACCGGTTGCGG
a  I D E G Q F F P D I V E F C E A M A N A -
  GGAAGACCGTAATTGTGGCTGCACTGGATGGGACCTTCCAGAGGAAGCCATTTGGGGCC
361  +-----+-----+-----+-----+-----+-----+ 420
  CCCTTCTGGCATTAAACACCGACGTGACCTACCCTGGAAGGTCTCCTTCGGTAAACCCCGG
a  G K T V I V A A L D G T F Q R K P F G A -
  ATCCTGAACCTGGTGCCGCTGGCCGAGAGCGTGGTGAAGCTGACGGCGGTGTGCATGGAG
421  +-----+-----+-----+-----+-----+-----+ 480
  TAGGACTTGGACCACGGCGACCGGCTCTCGCACCACTTCGACTGCCGCCACACGTACCTC
a  I L N L V P L A E S V V K L T A V C M E -
  TGCTTCCGGGAAGCCGCTATACCAAGAGGCTCGGCACAGACAAGGAGGTGAGGTGATT
481  +-----+-----+-----+-----+-----+-----+ 540
  ACGAAGGCCCTTCGGCGGATATGGTCTCCGAGCCGTGTCTGTTCCCTCCAGCTCCACTAA
a  C F R E A A Y T K R L G T D K E V E V I -
  GGGGAGCAGACAAGTACCCTCCGTGTGTCGGCTCTGCTACTTCAAGGCATCG
541  +-----+-----+-----+-----+-----+-----+ 600
  CCCCTCGTCTGTTTCATGGTGAGGCACACAGCCGAGACGATGAAGTCCGCCGGTAGC
a  G G A D K Y H S V C R L C Y F K A A A S -
  TGACTGACTGAC
601  +-----+-----+-----+-----+-----+-----+ 612
  ACTGACTGACTG
a  * * * -

```

**Fig. 7 Complete open reading frame of  $\Delta 40TK1$  in pGEX-6P-2 with the corresponding amino acid sequence.** The restriction sites for *Bam*HI and *Not*I are highlighted in grey, while the additional and mutated amino acids are in italic bold. After PreScission cleavage of the fusion protein three additional amino acid will be present at the N-terminus with respect to the original sequence.

Furthermore the insertion of the *NotI* restriction site caused three mutations at the C-terminus, namely K192A, S194A and G195S. All these mutations were not expected to alter the enzymatic activity, since it is known from sequence alignments that both ends are outside the conserved regions throughout the most TK's. Additionally, it has been shown that the 45 residue long amino terminal segment, which is not resolved in any of the published crystal structures, is not necessary for catalytic activity<sup>16</sup> but plays a role in migration within the cell<sup>17</sup>.

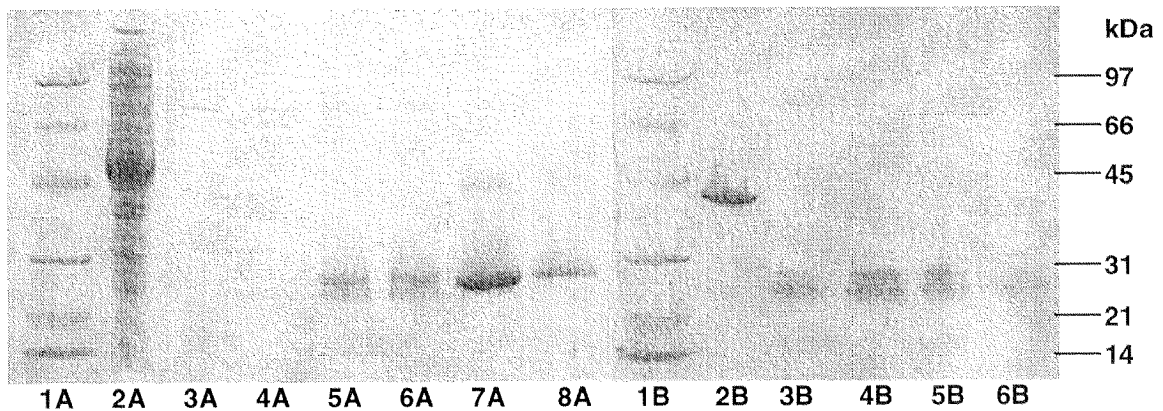
## 2.3 Expression & Purification

### 2.3.1 The human TK1 (Clone 34)

The plasmid PGEX-2T-Tkhum was used to transform both competent *E.coli* KY895 and BL21 strains. Bacteria were grown over night at 37°C in LB-medium containing 100 µg/ml ampicillin, the diluted 1:10 in fresh medium and grown to an OD<sub>600</sub> of 0.6 before reducing the temperature to 25°C and adding IPTG to a final concentration of 0.1 mM to induce the production of the GST-TK fusion protein. After 22 hours at 25°C, bacteria were harvested by centrifugation, frozen, thawed and lysed in buffer (50 mM Tris pH 7.5, 1 mM PMSF, 10 mM DTT, 10% glycerol and 1% Triton X-100) in the presence of 150 µg/ml lysozyme and 2000 units DNaseI (10 mM MgCl<sub>2</sub>, 1 mM MnCl<sub>2</sub> and 10 mM EDTA for inactivation of DNaseI afterwards) for 30 min at 4°C and by additional flow sonication (peristaltic pump 5x10, pulse –32, 0.7"). The lysate was clarified by centrifugation at 12,000xg for 20 min, filtered by a 0.45 µm filter and frozen at –70°C.

The crude extract was purified using the previously described method for HSV 1 TK<sup>10</sup>. After thrombin cleavage on the GSH-agarose column, the eluted fractions containing the human TK1 were pooled, diluted 1:10 in ATP-buffer A (50 mM MOPS pH 7.5, 1 mM EDTA, 10% Glycerol, 0.1% Triton X-100, 1 mM DTT and 50 µM dT) and incubated over night at 4°C. The diluted sample was then applied on an ATP-agarose column, washed with ATP-buffer B (50 mM MOPS pH 7.5, 10% Glycerol, 0.1% Triton X-100, 1 mM DTT and 50 µM dT) and eluted with 5 mM ATP

in ATP-buffer B. The recovery after ATP-agarose column was in a range of 35% with minute fluctuations ( $S_{rel} = \pm 28\%$ ). The purification is indicated in fig. 8.



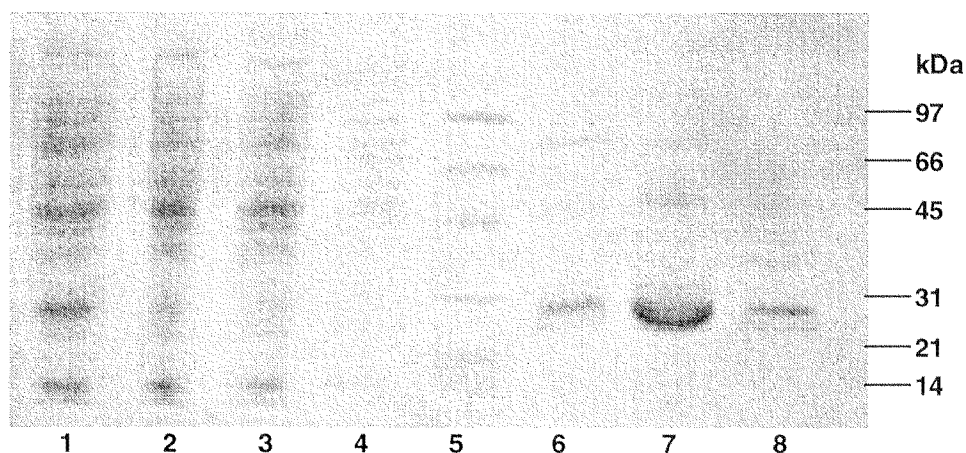
**Fig. 8 Purification of the human cytosolic TK1 clone 34.** **A.** Lane1: SDS low-range protein standards, Lane2: Crude extract after GSH-column, Lane3&4: dnaK-elution, Lane5&6: Fractions containing TK1 after Thrombin cleavage, Lane7&8: Fractions containing GST after regeneration of the column with GSH. **B.** Lane 1: SDS low-range protein standards, Lane2: HSV1 TK, Lane3-6: Fractions containing TK1 after ATP-column

This further purification step allowed to isolate the active form of the TK1. The pooled fractions containing protein were concentrated by ultrafiltration using Centricon 30<sup>®</sup> concentrators. During this last step most of the protein precipitated to major extent disallowing to reach higher concentrations than 6 mg/ml. The always co-purified dnaK could be separated partially by an additional washing step on the GSH-column following the procedure described by<sup>18</sup> The human TK1 was always purified in two different sizes after both columns. This heterogeneity was expected to influence the crystallization. Nevertheless, the concentrated samples were used for crystallization experiments.

### 2.3.2 The His-Tag-TK1 (pQE32TK1)

The bacterial strains XIL-1 blue transfected with the expression vector pQE32-TK1 clone 1 and 6 were generously supplied by R. Hofbauer, Vienna. Starting material for the construction of this vector has been the PCR-fragment RH60-RH39 amplified from the cDNA of clone 39-5 TK1, which was first subcloned into pGEM-T to allow the *Bam*HI/*Pst*I cleavage of the fragment to be inserted in the pQE32 vector. (Personal communication R. Hofbauer) Expression of the His-Tag-TK1 was performed as described 2.3.1 for the clone 34. The lysis was carried out

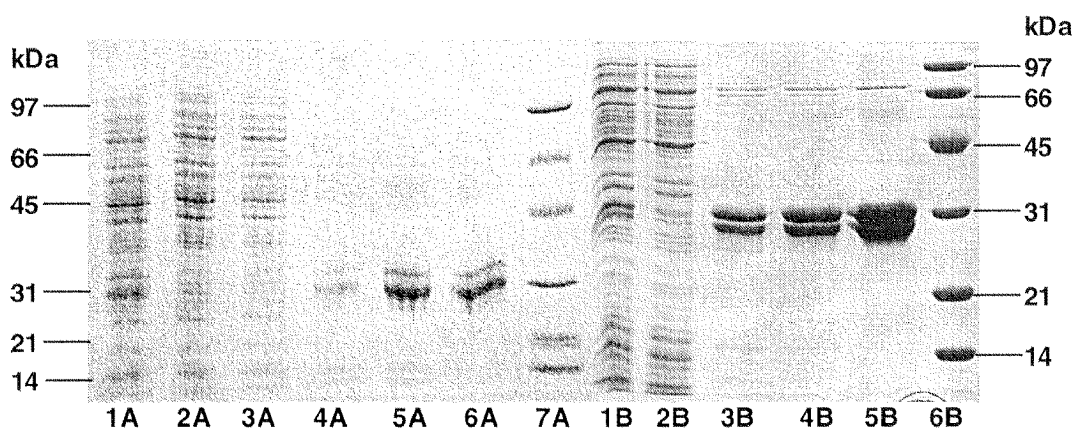
according to the protocols in<sup>19</sup>. The pellet was resuspended in PBS ,centrifuged and resuspended in five times the pellet volume of lysis buffer (50 mM Tris pH 8.0, 10 mM  $\beta$ -mercaptoethanol, 500 mM NaCl ,0.1% Triton X-100 and 10 mM imidazole). After addition of 150  $\mu$ g/ml lysozyme the suspension was stirred on ice for 15 min, sonicated for 5 min at 0.7" pulse and centrifuged at 12'000xg 4°C for 30 min. The supernatant was applied five times on a 2 ml Ni-NTA column (QIAGEN<sup>®</sup>) at room temperature. After a washing step with 50 mM Tris pH 6.0, 10 mM  $\beta$ -mercaptoethanol, 500 mM NaCl ,0.1% Triton X-100 and 40 mM imidazole to elute unspecific bound proteins the TK was eluted with a gradient of imidazole from 100 to 500 mM in elution buffer (50 mM Tris pH 8.0, 10 mM  $\beta$ -mercaptoethanol, 500 mM NaCl , 0.1% Triton X-100) yielding 10 to 65 mg of protein. The purification was monitored by SDS-PAGE. The removal of imidazole from the eluted fractions by Sephadex G-25 diminished the stability of the protein. In the presence of imidazole the protein was observed to be stable at 4°C during several weeks, while after removal of the imidazole almost the whole protein precipitated over night under the same conditions. For these reasons the imidazole was first not removed and the eluted proteins were concentrated and setup for crystallization.



**Fig. 9 Typical purification profile for his-tag TK1 using Ni-column** The whole his-tag-TK1 could be removed from the crude extract. After the washing step the column bound protein was eluted with imidazole. The purity of the eluted fractions was estimated to be >90%. Lane 1: Crude extract, Lane 2: Flow through, Lane 3 + 4: Wash steps, Lane 5: Low range marker, Lane 6-8: Fraction 3, 5 and 8 eluted with imidazole.



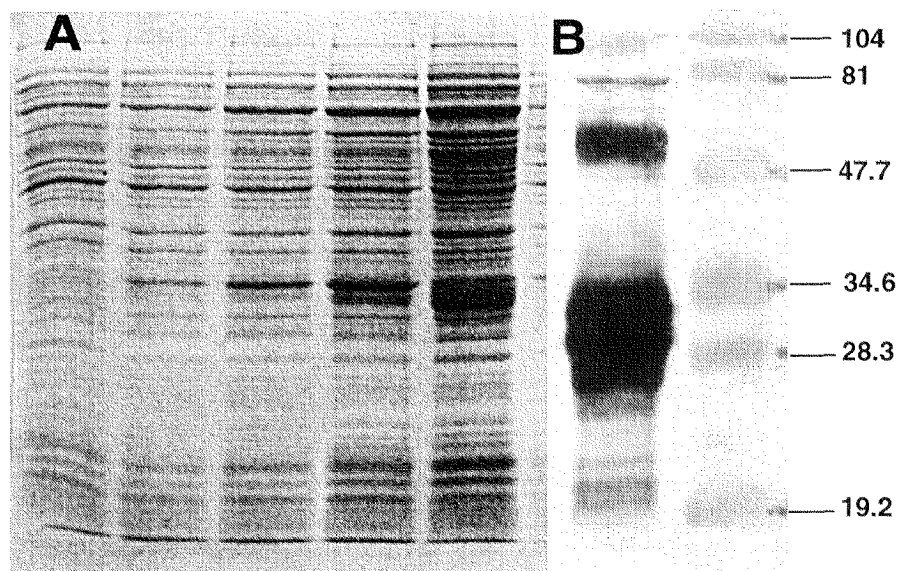
Further purification steps to remove the imidazole and the present impurities using ATP affinity chromatography were impossible. After elution from the Ni-column the samples containing the His-tag TK1 were first incubated with thymidine and then applied on an ATP-sepharose column. The column bound protein could not be eluted from the column neither with ATP gradients from 1 to 5 mM at pH 7.5 nor at pH 4.0. Further elution trials with 10 mM NiSO<sub>4</sub> also failed. After applying a sample of protein dye reagent on the column we could observe the protein to be retained on the upper bed support of the column. The protein could only be eluted with 8N urea. The questions that arose were the following: Does the protein interact with this support? Does it precipitate on the column? The first answer was clear no, because the used support was the same material as used for the Ni-column. For the second question the answer is still open. To avoid elution with imidazole a further elution trial with elution buffer containing 5 mM EDTA was performed. The protein could be eluted from the column, but less pure than with imidazole. Additionally, independently on the elution method the protein was never purified in a homogeneous length. At least two major sizes could be detected. The comparison of the eluted proteins is shown fig. 10.



**Fig. 10 Purification of the Histag-TK1 by Ni-affinity chromatography: A. Elution with imidazole:** Lane1: Crude extract, Lane2: Flow through, Lane3: Wash step, Lane4-6: Eluted protein fractions, Lane7: SDS low-range protein standards

In the western blot analysis both protein bands were recognized from purified polyclonal antibodies against the human TK (generously supplied by R. Hofbauer, Vienna).

Further investigations were performed to see, whether these different length were the result of protein degradation during expression. Therefore samples were taken from a bacterial culture before, after 1,3,6 and 24 h induction. Over the whole time range a clear progression towards different length could be monitored by SDS-PAGE. The western blot and the time experiment are presented in fig.11.



**Fig. 11 A. Control of His-tag-TK1 expression at different times after induction.** Lane1: t=0, Lane2: t=1h, Lane3: t=3h, Lane4: t=6h, Lane5: t=24h **B. Western blot with the His-tag-TK1.** Lane1: His-tag-TK1, Lane2: Prestained protein marker

These results let assume a proteolytic cleavage during expression in *E.coli*.

### 2.3.3 The deletion mutant $\Delta 40TK1$

The expression of the deletion mutant was performed as described for clone 34 in 2.3.1. GST affinity chromatography was then used to purify either the fusion protein or the cleaved protein.

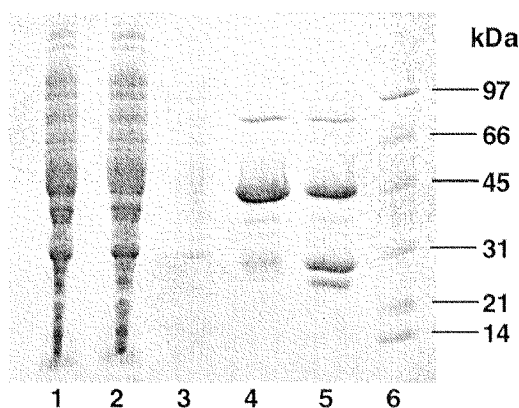
The crude extract was applied several times on a fresh regenerated GSH agarose column during 60-90 min by using a peristaltic pump. The column bound protein was then extensively washed with 25 ml buffer A, 25 ml buffer B, 10 ml buffer C and 10 ml buffer D.(see Tab.2.3.3.1) The column was then equilibrated with PreScission buffer (Buffer E). For the elution of the fusion protein 1ml fractions were collected with PreScission buffer containing 1mM GSH. The proteolytic

cleavage was performed directly on the column by incubating the GSH agarose with PreScission buffer containing 80 units PreScission protease over night at 4°C. The cleaved protein was then eluted with PreScission buffer. The fractions containing protein were pooled and concentrated using Millipore Ultrafree 4 Biomax 10K to reach the desired concentration for crystallization trials.

**Tab.2.3.3.1 Purification buffers for GSH-affinity chromatography**

Lysis buffer	50 mM TrisHCl, pH 7.5, 10% Glycerol, 5 mM EDTA
Buffer A	1N NaCl in lysis buffer 10 mM DTT , 0.1% Triton X-100
Buffer B	250 mM KH <sub>2</sub> PO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> , 150 mM NaCl, 0.1% Triton X-100, 10 mM DTT
Buffer C	250 mM KH <sub>2</sub> PO <sub>4</sub> /Na <sub>2</sub> HPO <sub>4</sub> , 150 mM NaCl, 0.1% Triton X-100, 10 mM DTT, 10 mM MgATP, pH 7.33
Buffer D	Lysis buffer, 100 μM dT, 10 mM DTT
Buffer E	50 mM TrisHCl, pH 7.0, 150 mM NaCl, 1 mM EDTA, 100 μM dT, 10 mM DTT

The first purification of the deletion mutant was performed for the fusion protein. One sample of the fusion protein was cleaved for 1h in solution at 4°C to check the presence of the PreScission cleavage site.(Fig. 12)

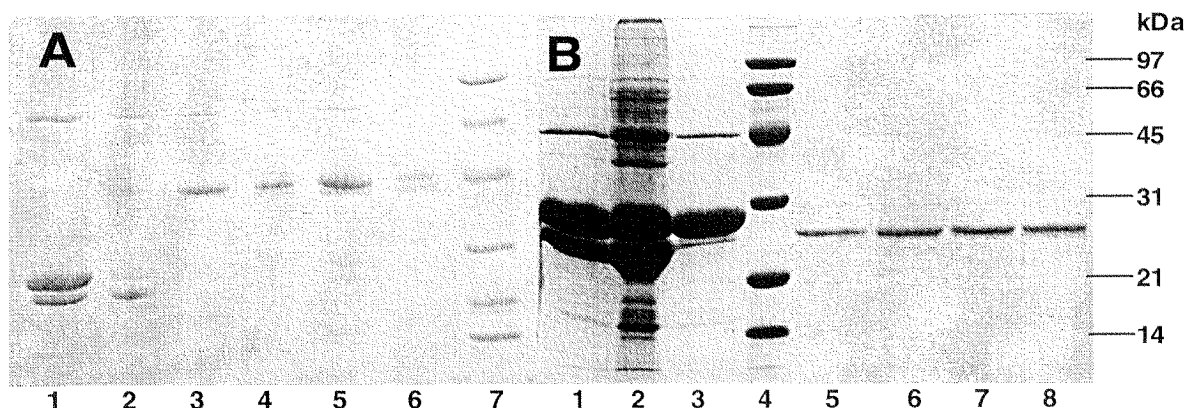


**Fig. 12 SDS-PAGE to monitor the purification and the presence of the PreScission cleavage site.** Lane 1: Crude extract; Lane 2: Flow through; Lane 3: Washing step with buffer A; Lane 4: Eluted fusion protein of  $\Delta 40$  TK1; Lane 5: PreScission cleavage of the fusion protein in solution; Lane 6: SDS low-range protein standards

One sample of fusion protein was retained for characterization of the enzyme (see 2.4.1)The rest of the fusion protein was cleaved over night at 4°C before the next purification step. During cleavage over night a precipitation of the protein was observed.

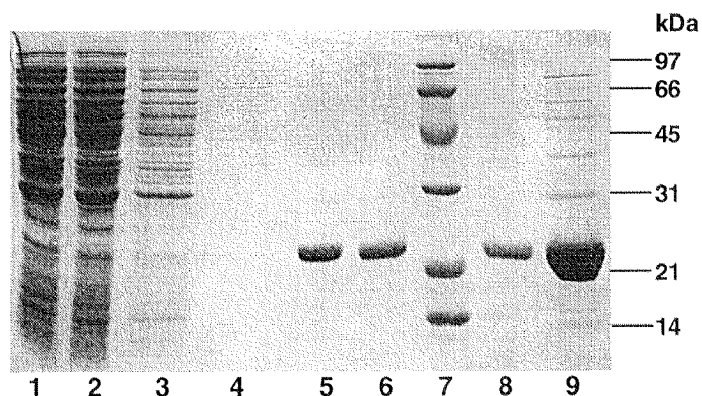
After centrifugation, the supernatant of the cleaved fraction was incubated with 50  $\mu$ M dT for 2 h before it was applied 10 times on an ATP agarose column. With the cleaved fraction poor binding on the column could be observed. After a washing step with lysis buffer the small amount of pure protein was eluted with elution buffer containing 5 mM ATP. Further binding experiments with the fusion protein showed that most of the applied protein was retained on the column but only very little amounts could be eluted with elution buffer containing 5 mM ATP. The rest of the protein could only be eluted from the column with 8N urea. This seemed to be the same problem like in the his-tag experiment.

An additional purification step using dT-agarose was performed without any success: No protein could be bound on the column. The ATP- and dT-agarose experiments are shown in fig. 13.



**Fig. 13 SDS-PAGE to monitor further purification steps of the deletion mutant.** LEFT GEL: Lane 1: Complete cleavage over night; Lane 2: Precipitate during cleavage at 4°C over night; Lane 3: Fusion protein before ATP agarose column; Lane 4: Fusion protein solution after application on dT-Sepharose; Lane 5: Fusion protein after elution from ATP agarose (0.7 mg/ml); Lane 6: Regeneration of the ATP agarose by 8M urea; Lane 7: SDS low-range protein standards. RIGHT GEL: Lane 1: Complete cleavage over night at 4°C (supernatant); Lane 2: Precipitate during cleavage at 4°C over night; Lane 3: Supernatant after ten fold application on ATP agarose column (no activity could be measured by UV-Spectrophotometry); Lane 4: SDS low-range protein standards; Lane 5-8: Fractions eluted with 5mM ATP from the column containing totally 400  $\mu$ g of pure  $\Delta$ 40 TK1

Since by an additional purification step using ATP affinity chromatography only small amounts of the protein were obtained, the first purification system was optimized to achieve enough high purity of the protein for crystallization trials. An optimized purification monitored by SDS-PAGE is shown in fig. 14.



**Fig. 14 SDS-PAGE of an optimized one step purification using GSH agarose.** Lane 1: Crude extract; Lane 2: Flow through; Lane 3: Washing step with buffer A; Lane 4: Washing step with buffer B; Lane 5&6: Eluted fractions of  $\Delta 40$  TK1 after cleavage on the column; Lane 7: SDS low-range protein standards; Lane 8: Unified fractions containing  $\Delta 40$  TK1 before concentration ( $c = 0.26$  mg/ml); Lane 9: Unified fractions containing  $\Delta 40$  TK1 after concentration ( $c = 2.5$  mg/ml)

The pooled fractions could be concentrated up to 7.5 mg/ml and were used for characterization and crystallization experiments.

#### 2.3.4 The *E.coli* BL21 codon plus-RIL strain

The transformation into the new *E.coli* strain was performed using the standard protocol from Stratagene with some minor modifications. The transformation was carried out with the following three plasmids: pGEX-6P-2-TK1, pGEX-2T-TK1 and pGEX-6P-2- $\Delta 40$ TK1. A control was also performed with the pUC18 plasmid (Stratagene<sup>®</sup>).

50  $\mu$ l aliquots of competent *E.coli* Codon plus-RIL strains were transferred into pre-chilled ED-cups. 1  $\mu$ l of an 1:10 dilution of XL 10-Gold mercaptoethanol (Stratagene<sup>®</sup>) were added and the bacterial suspension was let stand on ice for 10 minutes by gently mixing every two minutes. After adding 1  $\mu$ l of a DNA solution (0.1 ng/ $\mu$ l) the probe was gently mixed and incubated for 30 minutes on ice. After twenty seconds heat shock at 42°C the probe was incubated for additional two minutes on ice before adding 200  $\mu$ l of LB-medium containing 10 mM KCl and 10 mM MgSO<sub>4</sub>. The mixture was then incubated at 37°C for 45 minutes.

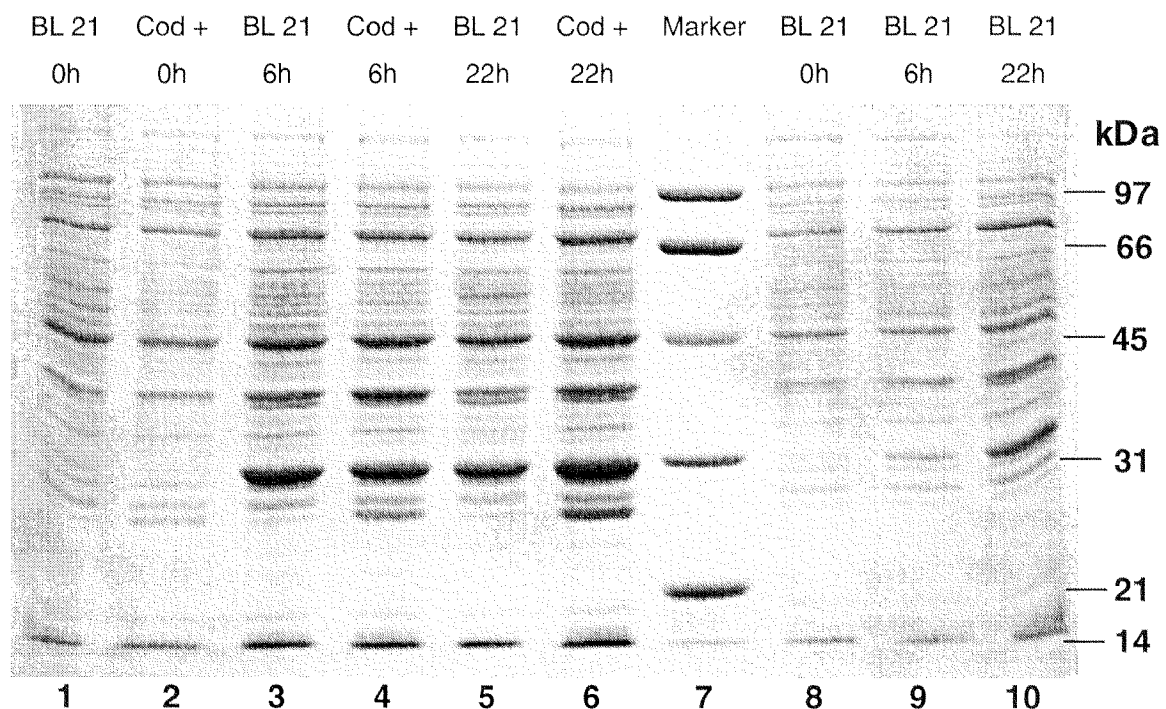
After transformation, the 200  $\mu$ l bacterial culture was directly transferred into 100 ml LB-Amp medium containing additional 34  $\mu$ g/ml chloramphenicol. The

presence of chloramphenicol in the medium was necessary to achieve a selective growth of strains carrying the additional RIL-plasmid.

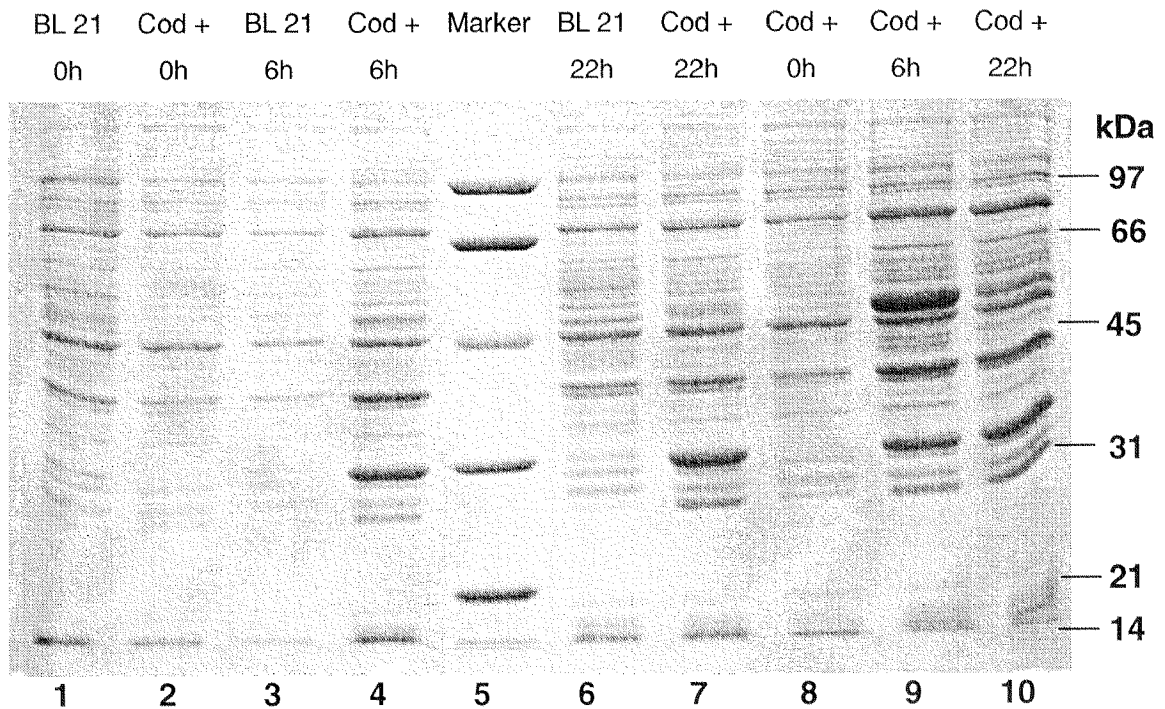
After dilution 1:10 with LB-Amp/Chl the cultures were grown for additional 90 –120 minutes at 37°C to reach an OD<sub>600</sub> between 0.6 and 1.0. Protein expression was then induced with 0.1 mM IPTG. After 22 hours at 25°C, bacteria were harvested by centrifugation, frozen, thawed and lysed as previously described in 2.3.1.

To evaluate the efficacy of the expression in the new strains, the same experiments were performed in parallel with the normal BL21 strains (Pharmacia®).

For qualitative analysis 1ml samples of the different cultures were taken before, after 6h and 22h of induction. The samples were centrifuged for 2 min at 14'000xg and the pellet was re-suspended / dissolved into 50 µl 8M Urea by mixing during 30 min with additional 5 pulse of sonication. After further 2 min centrifugation at 14'000xg the SDS-PAGE samples were prepared by mixing 5 µl supernatant with 15 µl water and 10 µl of 3x sample buffer.



**Figure 15.** SDS-PAGE for the control of the expression in the two different E.coli strains BL21 and BL21 Codon plus-RIL. Lane 1-6:  $\Delta 40$  TK1; Lane 7: SDS-Marker low range; Lane 8-10 pGEX-2T-TK. Strain and induction time is indicated on the top of the lanes. (BL21 = E.coli strain BL21; Cod+ = E.coli strain BL21 Codon plus-RIL)



**Figure 16.** SDS-PAGE for the control of the expression in the two different *E.coli* strains BL21(BL21) and BL21 Codon plus-RIL (Cod+). Lane 1-4,6 and 7: pGEX-6P-2-TK1; Lane 5: SDS-Marker low range; Lane 8-10 pGEX-2T-TK. Strain and induction time are indicated on the top of the lanes.

This experiment showed interesting results. The clearest result could be observed with the pGEX-2T-TK1. After 6h of induction a clear overexpression of a >45 kDa protein could be monitored in the codon plus strains, but not in the normal BL21 strains. After 22 h of induction this band disappeared indicating a possible proteolytic degradation of the protein during expression. The BL21 strains did never achieve a clear overexpression of the protein.

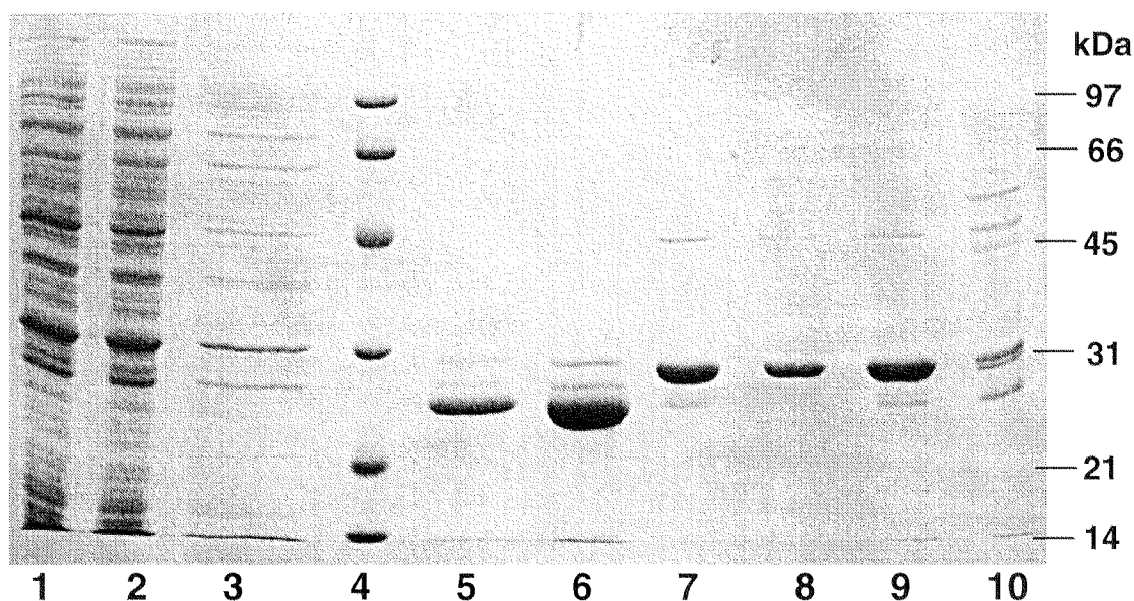
With the pGEX-6P-2-TK1 we could observe a similar effect, but weaker.

For the pGEX-6P-2- $\Delta$ 40TK1 the overexpression was slightly higher, but not that clear as observed in the pGEX-2T-TK1. In contrast, after 22 h even an intensification of the overexpressed protein band could be monitored.

These results were not really surprising and could have been expected partially. The difference between the full length and the deletion mutant can easily be explained by comparison of the DNA sequences of both enzymes. Two of totally five seldom codons for Arg present in the sequence for the human TK are located in the segment coding for the c-terminal 40 aa of the enzyme. Therefore the

pGEX-6P-2- $\Delta$ 40TK1, missing the last 40 aa, should not be expected to show a big improvement of protein expression by using the *E.coli* Codon plus-RIL strain.

On the other hand, the c-terminal part of proteins has been shown to play an important role for proteolytic degradation in *E.coli* and cells, taking the role of a possible recognition site for proteases<sup>11-14</sup>. This could explain the fact, that degradation could only be observed during expression of the full length enzyme and not for the deletion mutant.



**Figure 17.** SDS-PAGE monitoring the purification of  $\Delta$ 40TK1 expressed in *E.coli* strain BL21 Codon plus-RIL. Lane 1: Crude extract; Lane2: Flow through; Lane3: Washing step; Lane4: SDS-Marker low range; Lane 5: Associated fractions of  $\Delta$ 40TK1 after cleavage on the column; Lane 6: Concentrated fractions for crystallization; Lanes 7&8: Eluted GST after regeneration of the column with GSH; lanes 9&10: Regeneration of the column with 8M Urea.

The crude extracts were purified by GSH-affinity chromatography as previously described (see 2.3.3). Although using the same purification procedure the quality of the protein could be improved because of the new expression system (fig. 17).

The regeneration of the column indicated that the cleavage was almost complete. Only a small amount of fusion protein could be monitored by SDS-PAGE.

The pooled fractions of  $\Delta$ 40TK1, which contained 2.4 mg of protein, were concentrated to a final protein concentration of 10.5 mg/ml.

This enzyme solution was then used for characterization and crystallization trials.



## 2.4 Enzymatic activity

Besides purity, activity is a further important feature needed for crystallization of a macromolecule. Denatured macromolecules affect adversely crystal growth more than do unrelated molecules, especially when structural heterogeneities concern domains involved in crystal packing<sup>20</sup>. For this purpose the activity of the enzyme must be assessed before setting up crystallization experiment.

### 2.4.1 UV-Spectrophotometric activity assay

For a quick check of activity an UV-spectrophotometric assay based on pyruvate kinase and lactate dehydrogenase was mainly used<sup>21</sup> where the change in  $A_{340}$  due to oxidation of NADH was monitored over time and corresponds to the ADP formation during the phosphorylation reaction. The enzyme cascade is displayed in figure 18. The advantage of this method is its rapid and easy application.

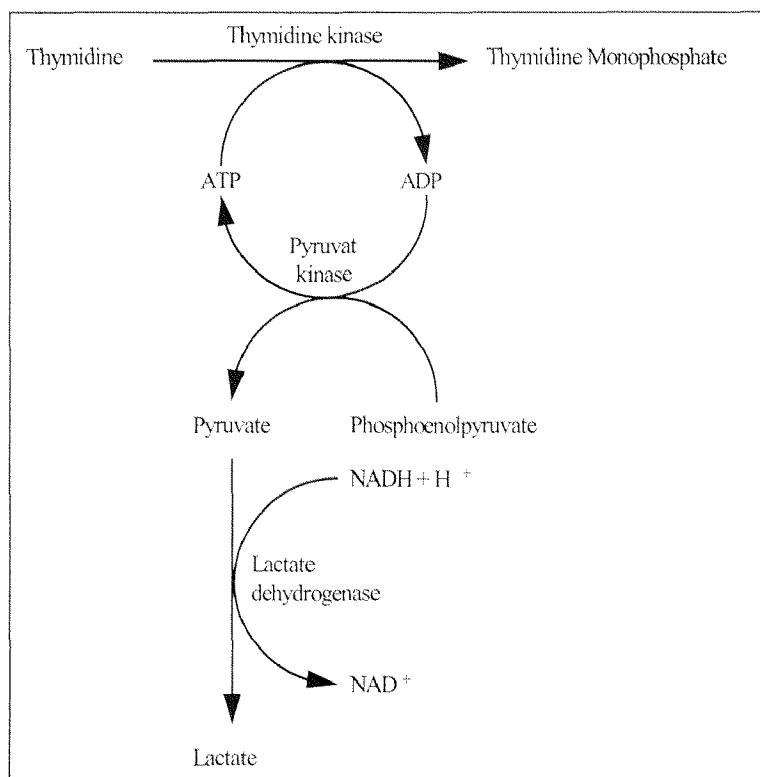


Fig. 18: Coupled enzyme assay based on ADP formation for thymidine kinase activity.

Yet, the method is rather quickly disturbed as the enzymatic turnover is monitored indirectly, i.e. some TK mutants might have lost kinase activity but exhibit

phosphatase activity, thus with this assay one can not be distinguished between hydrolyzed ATP and ADP formed during dT phosphorylation. Additionally, the association with an enzyme cascade does not allow modifications of the reaction conditions such as pH. The UV-approach is also a method of choice for ATP kinetics and for thymidine kinetics of mutants with very low affinity to dT. For good analytical measurements a certain amount of substrate must be converted to reach a change in absorbance of at least 0.05 to 0.1 in the spectrophotometer. Thus, the method is appropriate for kinetic measurements of substrates and enzymes exhibiting  $K_m$  values higher than 80  $\mu\text{M}$ .

**Tab. 2.4.1 Specific activities measured for different human TK1 by UV-Spectrophotometric assay**

ENZYME	Specific activity [pmol/min/ $\mu\text{g}$ ]	Stdev [pmol/min/ $\mu\text{g}$ ]	$S_{\text{rel}}$ [%]
Human TK1(clone 34)	80.8	29.6	36.6
FUPRO Human TK1(clone 34)	109.4	33.8	30.9
pQE32 TK1	1369.5	446.1	32.6
$\Delta 40\text{TK1}$	4161.5	1224.2	29.4

The eluted enzymes showed different activities indicated in table 2.4.1. Since the activity measurements were performed qualitatively no absolute conclusions could be made. A significant increase in activity was observed between the human TK1 clone 34 and the deletion mutant.

One may argue if activity and purity were always the same. The purity of the measured enzymes was different and the specific activity is related to the total amount of protein in the measured sample. Therefore we do not know the amount of correct folded protein. Only by knowing these details we could hypothesize that the C-terminus has effect on catalysis. The solution of the crystal structure will perhaps give us enough information allowing us to understand more about the structure-function relationship of this enzyme.

### 2.4.2 HPLC assay for assessing substrate specificity

For qualitative evaluation of the characteristics of the deletion mutant a high performance liquid chromatography (HPLC) system was applied to monitor ADP and nucleoside mono- or diphosphate formation during the enzymatic reaction. The HPLC system is based on ion-pair chromatography and uses a modified version of a previously published protocol<sup>22</sup> (Column: RP-18; Solvent: 0.2 M  $\text{NaH}_2\text{PO}_4$ , 25 mM Tetrabutyl-ammonium-hydrogensulfate, 3% Methanol; Flow: 1.1 ml/min; Detection: UV 254 nm).

Reactions were carried out in a final volume of 70  $\mu\text{l}$  containing 50 mM Tris pH 7.2, 5 mM ATP, 5 mM  $\text{MgCl}_2$ , 2 to 5 mM substrate and 1,5  $\mu\text{g}$   $\Delta 40$  TK1. The reaction was stopped after one hour at 37 °C by a ten fold dilution in water. The samples were then directly injected without removing the protein by acid precipitation prior to injection. The detection limit for phosphorylated substrate lies under 20 nmol making this method more sensitive than the UV-assay.

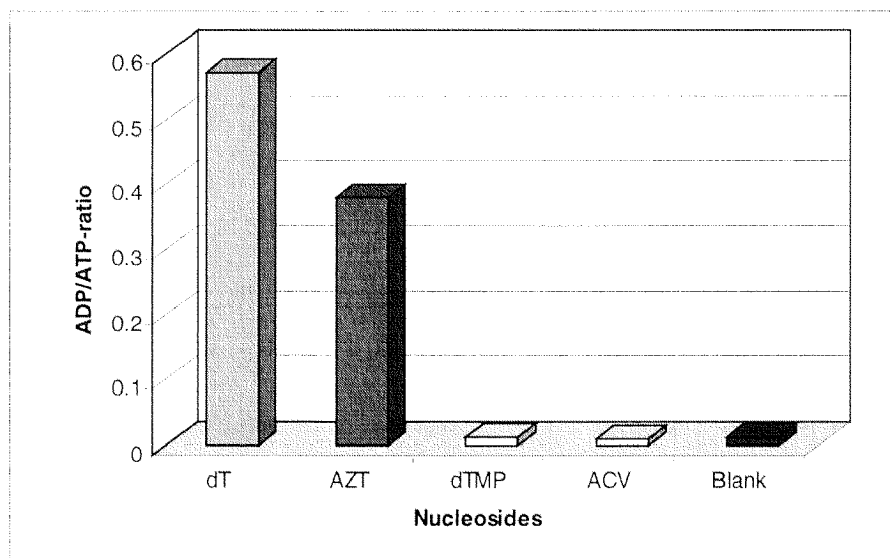


Fig. 19 ADP/ATP-ratios determined by HPLC to check some properties of the deletion mutant  $\Delta 40$ TK1

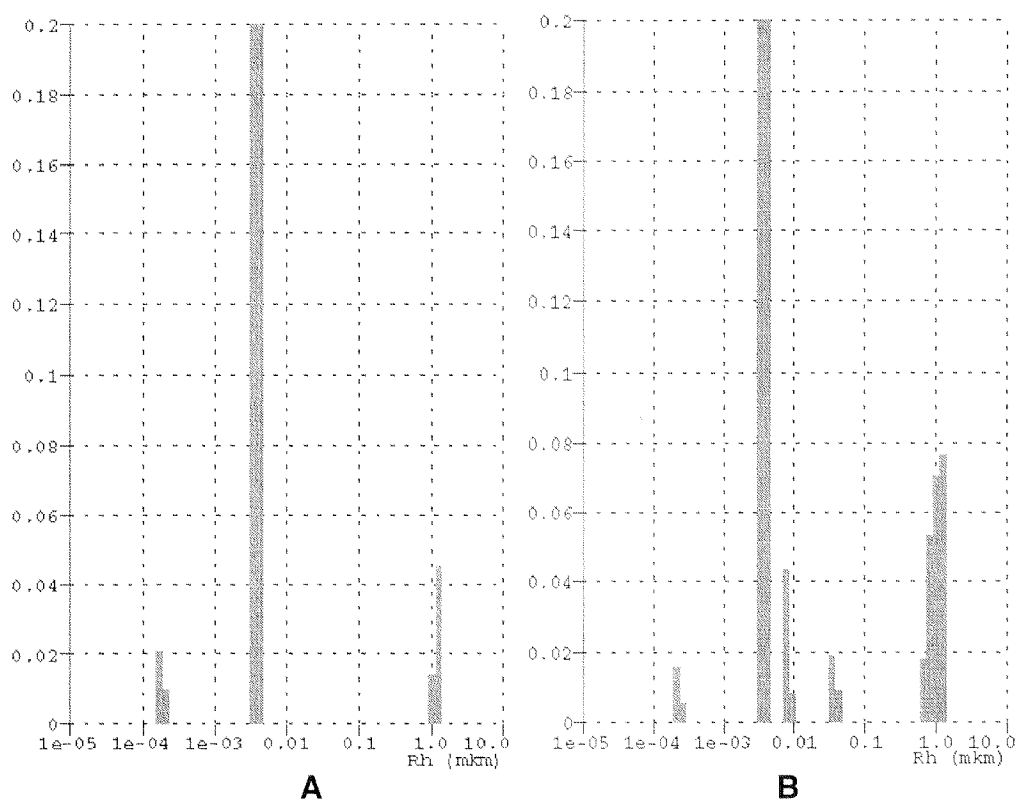
The deletion mutant showed similar substrate specificity as the full length human TK1 being able to phosphorylate dT and AZT but not dTMP or ACV<sup>23,24</sup>. The values of these two nucleosides are in the range of the blank monitoring the reaction independent ATP-hydrolysis. The relative standard deviations were smaller than 3%.



### 2.5.2 Dynamic light scattering measurements

Light scattering was performed with a Dynapro-801 molecular sizing instrument (Protein Solutions, Inc., Charlottesville, VA), generously made available by Prof. Dr. M. Grütter at the Department of Biochemistry (University of Zurich, Winterthurerstrasse 190, CH 8057 Zurich). Data analyses were performed with Protein Solutions Auto Pro Data analysis software. The purified protein was injected through a 0.02- $\mu\text{m}$  syringe filter at room temperature into the Dynapro-801 detector. The sample was measured at a protein concentration of 1.0 mg/ml.

The size distribution function of the  $\Delta 40\text{TK1}$  is shown in fig. 21.



**Fig. 21** Size distribution of a sample  $\Delta 40\text{TK1}$  (1.0 mg/ml) measured after filtration (A) and without filtration (B). The filtered sample shows a major monodispersity with some minor higher aggregates. The non filtered sample seems to be more heterodisperse showing clearly the presence of more higher aggregates.

The measurement A indicate clearly the presence of a mainly monodisperse sample with a mean particle radius of  $3.82 (\pm 0.4)$  nm in the presence of a small population of higher aggregates. The same experiment performed without filtering the sample (B) shows the presence of other two populations with particle radii of 7.9 , 36.4 nm beside the higher aggregates. We cannot decide, whether the

protein alone is responsible for this heterodispersity or not, since it is not known what kind of impurities are present in the sample. In both samples the main population has a radius of 3.8 nm. The theoretical molecular weight calculated with this radius is 82 kDa. This would confirm, in accordance with the FPLC-data, the  $\Delta 40\text{TK1}$  to be present as a tetramer under these conditions.

## 2.6 Discussion

Following the aim to obtain suitable protein for crystallization studies, different expression & purification systems were assessed for the human cytosolic thymidine kinase. The sequences of the analyzed TKs are aligned in the following figure compared to the sequence of the native enzyme.

The purification of the full length human TK1 as a thrombin cleavable GST-fusion protein resulted in a heterogeneous product consisting of two different length and additional impurities. The enzyme tend to aggregate. The quantities obtained with this system were not sufficient to set up a wide range of crystallization experiments. By using the his-tag-system we could improve the amounts of purified protein, but not the quality. The protein resulted to be more heterogeneous than in the GST-FUPRO system. Time experiment performed for the control of TK1 expression in *E.coli* let assume a proteolytic degradation that occurs during expression.

Having a closer look at the c-terminus of the his-tag-TK1 we recognize a quite hydrophobic sequence which could also play an important role for the degradation<sup>14</sup>. The double band has also been described recently as due to the effects of heterogeneous properties of the protein from the bacterial expression and the strong charge of the tag, as well as the overloading of the gel, without any further investigations<sup>26</sup>. Our results sustain the hypothesis of the heterogeneous properties and the hydrophobic sequence. Double bands were always present independently on the sample concentration and the presence of the tag. For this reason current approaches are being assessed in the lab to clone the open reading frame of the human cytosolic thymidine kinase without any sequence coding for additional amino acids into a pQE30 vector, in order to obtain a his-tag-

TK1 of homogeneous length. These major changes could improve the amount and quality of the protein for crystallization experiments. If not the system should be changed completely in a system that allows a better expression of eucaryotic proteins, for example yeast, CHO or insect cells.

		1				50
P04183	KITH_HUMAN	.....	...MSCINLP	TVLPGSPSKT	RGQIQVILGP	MFSGKSTEML
	pQE32-TK1	<b>MRGSHHHHHH</b>	<b>GIH</b> MSCINLP	TVLPGSPSKT	RGQIQVILGP	MFSGKSTEML
	pGEX-2T-TKhum	.....	.GSMSCINLP	TVLPGSPSKT	RGQIQVILGP	MFSGKSTEML
	D40TK1	.....	<b>GP</b> LGSMSCINLP	TVLPGSPSKT	RGQIQVILGP	MFSGKSTEML
	Consensus	.....	. <b>gs</b> MSCINLP	TVLPGSPSKT	RGQIQVILGP	MFSGKSTEML
		51				100
P04183	KITH_HUMAN	RRVRRFQIAQ	YKCLVIKYAK	DTRYSSSFCT	HDRNTMEALP	ACLLRDVAQE
	pQE32-TK1	RRVRRFQIAQ	YKCLVIKYAK	DTRYSSSFCT	HDRNTMEALP	ACLLRDVAQE
	pGEX-2T-TKhum	RRVRRFQIAQ	YKCLVIKYAK	DTRYSSSFCT	HDRNTMEALP	ACLLRDVAQE
	D40TK1	RRVRRFQIAQ	YKCLVIKYAK	DTRYSSSFCT	HDRNTMEALP	ACLLRDVAQE
	Consensus	RRVRRFQIAQ	YKCLVIKYAK	DTRYSSSFCT	HDRNTMEALP	ACLLRDVAQE
		101				150
P04183	KITH_HUMAN	ALGVAVIGID	EGQFFPDIVE	FCEAMANAGK	TVIVAALDGT	FQRKPFGAIL
	pQE32-TK1	ALGVAVIGID	EGQFFPDIVE	FCEAMANAGK	TVIVAALDGT	FQRKPFGAIL
	pGEX-2T-TKhum	ALGVAVIGID	EGQFFPDIVE	FCEAMANAGK	TVIVAALDGT	FQRKPFGAIL
	D40TK1	ALGVAVIGID	EGQFFPDIVE	FCEAMANAGK	TVIVAALDGT	FQRKPFGAIL
	Consensus	ALGVAVIGID	EGQFFPDIVE	FCEAMANAGK	TVIVAALDGT	FQRKPFGAIL
		151				200
P04183	KITH_HUMAN	NLVPLAESVW	KLTAVCMECF	REAAATKRLG	<b>TE</b> KEVEVIGG	ADKYHSVCRL
	pQE32-TK1	NLVPLAESVW	KLTAVCMECF	REAAATKRLG	<b>TE</b> KEVEVIGG	ADKYHSVCRL
	pGEX-2T-TKhum	NLVPLAESVW	KLTAVCMECF	REAAATKRLG	<b>TD</b> KEVEVIGG	ADKYHSVCRL
	D40TK1	NLVPLAESVW	KLTAVCMECF	REAAATKRLG	<b>TD</b> KEVEVIGG	ADKYHSVCRL
	Consensus	NLVPLAESVW	KLTAVCMECF	REAAATKRLG	<b>T#</b> KEVEVIGG	ADKYHSVCRL
		201				250
P04183	KITH_HUMAN	CYFKK <b>ASG</b> QP	AGEDNKENC	VPC <b>K</b> PGEAVA	ARKLFAPQQI	LQC <b>SPAN</b> ...
	pQE32-TK1	CYFKK <b>ASG</b> QP	AGEDNKENC	VPC <b>K</b> PGEAVA	AKKLFAPQQI	LQC <b>MPESLVR</b>
	pGEX-2T-TKhum	CYFKK <b>ASG</b> QP	AGEDNKENC	VPC <b>R</b> PGEAVA	ARKLFAPQQI	LQC <b>MQA</b> ....
	D40TK1	CYFKK <b>AS</b> AS..	.....	.....	.....	.....
	Consensus	CYFKK <b>asg</b> QP	agpdnkencp	vpg.pgeava	a.klfapqqi	lqc.....
		251				
P04183	KITH_HUMAN	.....				
	pQE32-TK1	<b>PPAVQA</b>				
	pGEX-2T-TKhum	.....				
	D40TK1	.....				
	Consensus	.....				

**Fig. 22 Multiple alignment of the sequences of the analyzed TKs compared to the native one.** The differences in the sequence are highlighted in bold.

The experiments performed with the deletion mutant  $\Delta 40\text{TK1}$  revealed to be the most promising ones, leading to a homogeneous product with full enzymatic activity and reduced tendency to aggregate. The enzyme could be purified as fusion protein and as PreScission cleaved protein. Both the fusion protein and the cleaved did bind to ATP agarose but were poorly eluted with ATP. The major amount of protein was removable only with 8N Urea. Therefore a one step purification procedure was optimized. The always co-purified heat shock protein from *E.coli*, dnaK was reduced when the TK was expressed in *E.coli* BL21 codon plus RIL strains. After cleavage on the column maximal concentrations of the protein between 7 and 10 mg/ml could be achieved in the presence of high amounts of DTT. UV- and HPLC Activity tests performed to verify some characteristics of the enzyme indicated that the  $\Delta 40\text{TK1}$  conserved the typical properties of the human cytosolic TK1, being able to phosphorylate dT and AZT, but not dTMP and ACV<sup>23,24</sup>. Activity of the enzyme was still recovered after storage of several month at  $-70^{\circ}\text{C}$ . Homogeneity and quaternary structure analysis confirmed the tetrameric state of the recombinant enzyme as previously described<sup>25</sup>.

After the one step purification procedure using GSH-affinity chromatography we do not know the amount of correct fold protein, which will be responsible for activity and crystallization. The purity of the protein is normally determined by SDS-PAGE. We generally consider that a major band on a Coomassie blue stained gel is sufficient for initial screening in crystallization. This does not exclude the possibility that subsequent purification may be necessary to improve crystal quality. Therefore dynamic light scattering was performed. This method allows us to detect the presence of aggregates in the protein solution which may be detrimental to crystal nucleation or growth. The measurements were performed with the deletion mutant  $\Delta 40\text{TK1}$ . For this enzyme we could measure the main population to be monodisperse and consequently covering the ideal premise for crystallization experiments. Similar pattern have been measured for the active site mutant Y101F (see chapter 3) which has been successful crystallized (chapter 5). The question remained whether dynamic light scattering can predict crystallization or not. Size distribution is certainly a good tool to determine the quality of the protein sample.



In a study where 44 of 66 analyzed proteins were crystallized, 83% of them were narrow unimodal, 15% broad unimodal and only 2% multimodal<sup>27</sup>. Nevertheless, once this value for the protein in solution is known the major work still remains to find the ideal conditions for crystal growth, even though the chances to succeed are higher.

Looking at all this aspects I think to have shown on this chapter that crystallization does not start in the hanging or sitting drop, but at a much earlier stage, and that each piece of information which helps to control the initial conditions increases the chances of having success.

## 2.7. References

1. Folkers, G. *et al.* Computer-aided active-site-directed modeling of the herpes simplex virus 1 and human thymidine kinase. *J Comput Aided Mol Des* **5**, 385-404 (1991).
2. Black, M.E., Newcomb, T.G., Wilson, H.M. & Loeb, L.A. Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy. *Proc Natl Acad Sci U S A* **93**, 3525-9 (1996).
3. Encell, L.P., Landis, D.M. & Loeb, L.A. Improving enzymes for cancer gene therapy. *Nat Biotechnol* **17**, 143-7 (1999).
4. Kokoris, M., Sabo, P., Adman, E. & Black, M. Enhancement of tumor ablation by a selected HSV-1 thymidine kinase mutant. *Gene Ther* **6**, 1415-1426 (1999).
5. Bordignon, C. *et al.* Transfer of the HSV 1 TK gene into donor peripheral blood lymphocytes for in vivo modulation of donor anti-tumor immunity after allogeneic bone marrow transplantation. *Hum Gene Ther* **6**, 813-9 (1995).
6. Bonini, C. *et al.* HSV 1 TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia [see comments]. *Science* **276**, 1719-24 (1997).
7. Verzeletti, S. *et al.* Herpes simplex virus thymidine kinase gene transfer for controlled graft-versus-host disease and graft-versus-leukemia: clinical follow-up and improved new vectors. *Hum Gene Ther* **9**, 2243-51 (1998).

8. Sambrook, J., Fritsch, E.F. & Maniatis, T. *Molecular cloning : a laboratory manual*, 3 v. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y., 1989).
9. McPherson, A. Current approaches to macromolecular crystallization. *Eur J Biochem* **189**, 1-23 (1990).
10. Fetzer, J., Michael, M., Bohner, T., Hofbauer, R. & Folkers, G. A fast method for obtaining highly pure recombinant herpes simplex virus type 1 thymidine kinase. *Protein Expr Purif* **5**, 432-41 (1994).
11. Kauffman, M.G. & Kelly, T.J. Cell cycle regulation of thymidine kinase: residues near the carboxyl terminus are essential for the specific degradation of the enzyme at mitosis. *Mol Cell Biol* **11**, 2538-46 (1991).
12. Kauffman, M.G., Rose, P.A. & Kelly, T.J. Mutations in the thymidine kinase gene that allow expression of the enzyme in quiescent (G<sub>0</sub>) cells. *Oncogene* **6**, 1427-35 (1991).
13. Sutterluety, H., Bartl, S., Karlseder, J., Wintersberger, E. & Seiser, C. Carboxy-terminal residues of mouse thymidine kinase are essential for rapid degradation in quiescent cells. *J Mol Biol* **259**, 383-92 (1996).
14. Parsell, D.A., Silber, K.R. & Sauer, R.T. Carboxy-terminal determinants of intracellular protein degradation. *Genes Dev* **4**, 277-86 (1990).
15. Pilger, B. Swiss Federal Institute of Technology (1999).
16. Halpern, M.E. & Smiley, J.R. Effects of deletions on expression of the herpes simplex virus thymidine kinase gene from the intact viral genome: the amino terminus of the enzyme is dispensable for catalytic activity. *J Virol* **50**, 733-8 (1984).
17. Degreve, B., Johansson, M., De Clercq, E., Karlsson, A. & Balzarini, J. Differential intracellular compartmentalization of herpetic thymidine kinases (TKs) in TK gene-transfected tumor cells: molecular characterization of the nuclear localization signal of herpes simplex virus type 1 TK. *J Virol* **72**, 9535-43 (1998).
18. Silva, N.L., Haworth, R.S., Singh, D. & Fliegel, L. The carboxyl-terminal region of the Na<sup>+</sup>/H<sup>+</sup> exchanger interacts with mammalian heat shock protein. *Biochemistry* **34**, 10412-20 (1995).

19. QIAGEN. *The QIAexpressionist-A handbook for high level expression and purification of 6xHis-tagged proteins*, (, 1997).
20. Ducruix, A. & Giegé, R. *Crystallization of nucleic acids and proteins : a practical approach*, xxiv, 331 (IRL Press at Oxford University Press, Oxford [England] ; New York, 1992).
21. Keller, P.M. *et al.* Enzymatic phosphorylation of acyclic nucleoside analogs and correlations with antiherpetic activities. *Biochemical Pharmacology* **30**, 3071-7 (1981).
22. Masson, S., Desmoulin, F., Sciaky, M. & Cozzone, P.J. Catabolism of adenine nucleotides and its relation with intracellular phosphorylated metabolite concentration during ethanol oxidation in perfused rat liver. *Biochemistry* **32**, 1025-31 (1993).
23. Chen, M.S. & Prusoff, W.H. Association of thymidylate kinase activity with pyrimidine deoxyribonucleoside kinase induced by herpes simplex virus. *Journal of Biological Chemistry* **253**, 1325-7 (1978).
24. Elion, G.B. *et al.* Selectivity of action of an antiherpetic agent, 9-(2-hydroxyethoxymethyl) guanine. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5716-20 (1977).
25. Munch-Petersen, B., Tyrsted, G. & Cloos, L. Reversible ATP-dependent transition between two forms of human cytosolic thymidine kinase with different enzymatic properties. *J Biol Chem* **268**, 15621-5 (1993).
26. Lunato, A.J. *et al.* Synthesis of 5-(carboranylalkylmercapto)-2'-deoxyuridines and 3- (carboranylalkyl)thymidines and their evaluation as substrates for human thymidine kinases 1 and 2. *J Med Chem* **42**, 3378-89 (1999).
27. D'Arcy, A. Crystallizing proteins - a rational approach? *Acta Cryst D* **50**, 469-471 (1994).

Seite Leer /  
Blank leaf

## CHAPTER 3

### **The crystallization of thymidine kinases and enolase:**

Rational or black magic?

## Preface

The aim of this chapter is to give an overview of the experiences accumulated in macromolecular crystallization dealing with thymidine kinases and enolase. The reader should get an idea of the pitfalls and hits one may encounter and should learn that in most of the cases nothing works in the way he wants. Nevertheless, macromolecular crystallization still remains a good tool to train the way of thinking and looking at details in a rational way.

## Abstract

A first target for crystallization was the recombinant human cytosolic thymidine kinase and the results of different screenings that have been performed are reported. A major issue that we had to face during this experiments was protein precipitation linked to its instability. After several hits with salt crystals a protein modification and an optimized one step purification allowed finally to find conditions for the growth of small protein crystals. First X-ray measurements did not detect any diffraction and allowed to exclude the presence of inorganic crystals.

A second target was the crystallographic analysis of three active site mutants of HSV 1 TK. Two of them could successfully be crystallized and their structures are described in chapter 4 and 5. On this chapter only the crystallization aspects will be discussed and compared to the third mutant, which could not be crystallized even though the first two mutants had shown to crystallize under pretty similar condition as the wildtype enzyme, letting still open the question on the rationality of crystallization.

An additional target to crystallize was the enolase from spores of *Alternaria alternata*, a mould which is recognized as a major cause of fungal allergies of the respiratory tracts. This enzyme could be crystallized from similar conditions as published for a homologue protein, but could not reach suitable dimensions for X-ray measurements. The attempts to improve the crystal size are presented.

### 3.1 Introduction

At the beginning of crystallization experiments the typical question to be answered is: „What starting conditions should I choose?“ This point is of notable importance for the time and amounts of protein needed to succeed in crystallization. The more rational approach is to search for homologue proteins in a databank<sup>1,2</sup> and check if their structures are already deposited in the PDB Databank<sup>3</sup>. The information on the crystallization conditions of these proteins may give you good suggestion upon the starting point for an initial screening, because there are enough examples described in the literature, where homologue proteins crystallize under similar conditions, e.g. enolase<sup>4,5</sup>, phospholipase A<sup>6-8</sup> and thioredoxin<sup>9-11</sup>. Nonetheless, the experience also showed that even one mutation alone might change the behavior of the protein resulting in different crystal contacts and therefore completely different crystallization conditions. There are even cases where the identical protein prepared by different procedures or at different times may show significant variations<sup>12</sup>.

In the case that no homologue proteins are known the whole approach becomes more random, sometimes it is like searching the needle in the haystack. For such random approaches different screenings covering a wide range of crystallization conditions have been described<sup>13,14</sup> and are commercially available<sup>15,16</sup> (see Tab.3.1-3.2). These sets of conditions have virtually revolutionized the search for effective crystallization conditions. Although some important macromolecules continue to slip through these nets, the number of hits is impressive. Further classical screens with the two precipitant types ammonium sulfate or PEG 6K 12,<sup>17-19</sup> have successfully been used for the crystallization of the most macromolecules<sup>20</sup> (see Tab.3.3-3.4). The general aim of all these screenings is to find as fast as possible useful crystallization conditions or leads, independently of the quality and shape of the crystals. Once identified, these conditions can be optimized in a rational way using general schemes as described in chapter 1 to produce single crystals suitable for X-ray diffraction analysis. These single crystals should then be reproduced for X-ray data collections.

The human cytosolic TK1 and the active site mutants of HSV 1 TK, as well as the enolase from *Alternaria alternata* represented therefore ideal challenges to evaluate these approaches, and have also shown that even more factors might influence the procedure of finding the right crystallization conditions. Thousands of crystallization trials with the purified proteins were set up in hanging drop or sitting drop by varying different parameters like temperature, drop and reservoir volume, precipitating agents and additives.

In this chapter, with the aim of giving the impression of what somebody can expect, when starting in this field of research, the trials, which failed but added new knowledge and helped towards the achievement of some of the pursued aims, are reported. The successful attempts are separately described in chapter 4 and 5. The commonly used screenings are listed in the tables below.



**Tab. 3.1 Crystal Screen I reagent formulations**<sup>14,15</sup>

1. 30% MPD, 0.1 M Na Acetate pH 4.6, 0.02 M Calcium Chloride
2. 0.4 M K, Na Tartrate
3. 0.4 M Ammonium Phosphate
4. 2.0 M Ammonium Sulfate, 0.1 M Tris HCl pH 8.5
5. 30% MPD, 0.1 M Na Hepes pH 7.5, 0.2 M Na Citrate
6. 30% PEG 4000, 0.1 M Tris HCl pH 8.5, 0.2 M Mg Chloride
7. 1.4 M Na Acetate, 0.1 M Na Cacodylate pH 6.5
8. 30% 2-Propanol, 0.1 M Na Cacodylate pH 6.5, 0.2 M Na Citrate
9. 30% PEG 4000, 0.1 M Na Citrate pH 5.6, 0.2 M Ammonium Acetate
10. 30% PEG 4000, 0.1 M Na Acetate pH 4.6, 0.2 M Ammonium Acetate
11. 1.0 M Ammonium Phosphate, 0.1 M Na Citrate pH 5.6
12. 30% 2-Propanol, 0.1 M Na Hepes pH 7.5, 0.2 M Mg Chloride
13. 30% PEG 400, 0.1 M Tris HCl pH 8.5, 0.2 M Na Citrate
14. 28% PEG 400, 0.1 M Na Hepes pH 7.5, 0.2 M Ca Chloride
15. 30% PEG 8000, 0.1 M Na Cacodylate pH 6.5, 0.2M Ammonium Sulfate
16. 1.5 M Li Sulfate, 0.1 M Na Hepes pH 7.5
17. 30% PEG 4000, 0.1 M Tris HCl pH 8.5, 0.2 M Li Sulfate
18. 20% PEG 8000, 0.1 M Na Cacodylate pH 6.5, 0.2 M Mg Acetate
19. 30% 2-Propanol, 0.1 M Tris HCl pH 8.5, 0.2 M Ammonium Acetate
20. 25% PEG 4000, 0.1 M Na Acetate pH 4.6, 0.2 M Ammonium Sulfate
21. 30% MPD, 0.1 M Na Cacodylate pH 6.5, 0.2 M Mg Acetate
22. 30% PEG 4000, 0.1 M Tris HCl pH 8.5, 0.2 M Na Acetate
23. 30% PEG 400, 0.1 M Na Hepes pH 7.5, 0.2 M Mg Chloride
24. 20% 2-Propanol, 0.1 M Na Acetate pH 4.6, 0.2 M Ca Chloride
25. 1.0 M Na Acetate, 0.1 M Imidazole pH 6.5
26. 30% MPD, 0.1 M Na Citrate pH 5.6, 0.2 M Ammonium Acetate
27. 20% 2-Propanol, 0.1 M Na Hepes pH 7.5, 0.2 M Na Citrate
28. 30% PEG 8000, 0.1 M Na Cacodylate pH 6.5, 0.2 M Na Acetate
29. 0.8 M K, Na Tartrate, 0.1 M Na Hepes pH 7.5
30. 30% PEG 8000, 0.2 M Ammonium Sulfate
31. 30% PEG 4000, 0.2 M Ammonium Sulfate
32. 2.0 M Ammonium Sulfate
33. 4.0 M Na Formate
34. 2.0 M Na Formate, 0.1 M Na Acetate pH 4.6
35. 1.6 M Na, K Phosphate, 0.1 M Na Hepes pH 7.5
36. 8% PEG 8000, 0.1 M Tris HCl pH 8.5
37. 8% PEG 4000, 0.1 M Na Acetate pH 4.6
38. 1.4 M Na Citrate, 0.1 M Na Hepes pH 7.5
39. 2% PEG 400, 2.0 M Ammonium Sulfate, 0.1 M Na Hepes pH 7.5
40. 20% 2-Propanol, 20% PEG 4000, 0.1 M Na Citrate pH 5.6
41. 10% 2-Propanol, 20% PEG 4000, 0.1 M Na Hepes pH 7.5
42. 20% PEG 8000, 0.05 M K Phosphate
43. 30% PEG 1500
44. 0.2 M Mg Formate
45. 18% PEG 8000, 0.1 M Na Cacodylate pH 6.5, 0.2 M Zn Acetate
46. 18% PEG 8000, 0.1 M Na Cacodylate pH 6.5, 0.2 M Ca Acetate
47. 2.0 M Ammonium Sulfate, 0.1 M Na Acetate pH 4.6
48. 2.0 M Ammonium Phosphate, 0.1 M Tris HCl pH 8.5
49. 2% PEG 8000, 1.0 M Li Sulfate
50. 15% PEG 8000, 0.5 M Li Sulfate

**Tab. 3.2 Crystal Screen 2 reagent formulations**<sup>13,15</sup>

1. 10% PEG 6000, 2.0 M Na chloride
2. 0.5 M NaCl, 0.01 M CTAB, 0.01 M Mg chloride
3. 25% Ethylene glycol
4. 35% Dioxane
5. 5% Isopropanol, 2.0 M Ammonium sulfate
6. 1.0 M Imidazole pH 7.0
7. 10% PEG 1000, 10% PEG 8000
8. 10% Ethanol, 1.5 M Na chloride
9. 2.0 M Na chloride, 0.1 M Na acetate pH 4.6
10. 30% MPD, 0.1 M Na Acetate pH 4.6, 0.2 M NaCl
11. 1.0 M 1,6 Hexanediol, 0.1 M Na Acetate pH 4.6, 0.01 M Co chloride
12. 30% PEG 400, 0.1 M Na acetate pH 4.6, 0.1 M Cd chloride
13. 30% PEG MME 2000, 0.1 M Na Acetate pH 4.6, 0.2 M Ammonium sulfate
14. 2.0 M Ammonium sulfate, 0.1M Na Citrate pH 5.6 0.2 M K/Na Tartrate
15. 1.0 M Li sulfate, 0.1M Na Citrate pH 5.6, 0.5 M Ammonium sulfate
16. 2% Polyethyleneimine, 0.1 M Na Citrate pH 5.6, 0.5 M Na chloride
17. 35% tert-butanol, 0.1 M Na citrate pH 5.6
18. 10% Jeffamine M-600, 0.1 M Na citrate pH 5.6, 0.01M Ferric chloride
19. 2.5 M 1,6 Hexanediol, 0.1 M Na citrate pH 5.6
20. 1.6 M Mg sulfate, 0.1 M MES pH 6.5
21. 2.0 M Na chloride, 0.1 M MES pH 6.5, 0.2 M Na/K Phosphate
22. 12% PEG 20,000, 0.1 M MES pH 6.5
23. 10% Dioxane, 0.1 M MES pH 6.5, 1.6 M Ammonium sulfate
24. 30% Jeffamine M-600, 0.1 M MES pH 6.5, 0.05 M Cs chloride
25. 1.8 M Ammonium sulfate, 0.1 M MES pH 6.5, 0.01 M Co chloride
26. 30% PEG MME 5000, 0.1 M MES pH 6.5, 0.2 M Ammonium sulfate
27. 25% PEG MME 550, 0.1 M MES pH 6.5, 0.01 M Zn sulfate
28. 1.6 M Sodium citrate pH 6.5
29. 30% MPD, 0.1 M Hepes pH 7.5, 0.5 M Ammonium sulfate
30. 10% PEG 6000, 0.1 M Hepes pH 7.5, 5% MPD
31. 20% Jeffamine M-600, 0.1 M Hepes pH 7.5
32. 1.6 M Ammonium sulfate, 0.1 M Hepes pH 7.5, 0.1 M Na chloride
33. 2.0 M Ammonium formate, 0.1 M Hepes pH 7.5
34. 1.0 M Na acetate, 0.1 M Hepes pH 7.5, 0.05 M Cd sulfate
35. 70% MPD, 0.1 M Hepes pH 7.5
36. 4.3 M Na chloride, 0.1 M Hepes pH 7.5
37. 10% PEG 8000, 0.1 M Hepes pH 7.5, 8% Ethylene glycol
38. 20% PEG 10,000, 0.1 M Hepes pH 7.5
39. 3.4 M 1,6 Hexanediol, 0.1 M Tris pH 8.5, 0.2 M Mg chloride
40. 25% tert-butanol, 0.1 M Tris pH 8.5, 0.1 M Ca chloride
41. 1.0 M Li sulfate, 0.1 M Tris pH 8.5, 0.01 M Ni chloride
42. 12% Glycerol, 0.1 M Tris pH 8.5, 1.5 M Ammonium sulfate
43. 50% MPD, 0.1 M Tris pH 8.5, 0.2 M Ammonium phosphate
44. 20% Ethanol, 0.1 M Tris pH 8.5
45. 20% PEG MME 2000, 0.1 M Tris pH 8.5, 0.01 M Ni chloride
46. 30% PEG MME 550, 0.1 M Bicine pH 9.0, 0.1 M Na chloride
47. 2.0 M Mg chloride, 0.1 M Bicine pH 9.0
48. 10% PEG 20,000, 0.1 M Bicine pH 9.0, 2% Dioxane

**Tab.3.3 Grid Screen Ammonium Sulfate reagent formulations**<sup>12,16,19</sup>

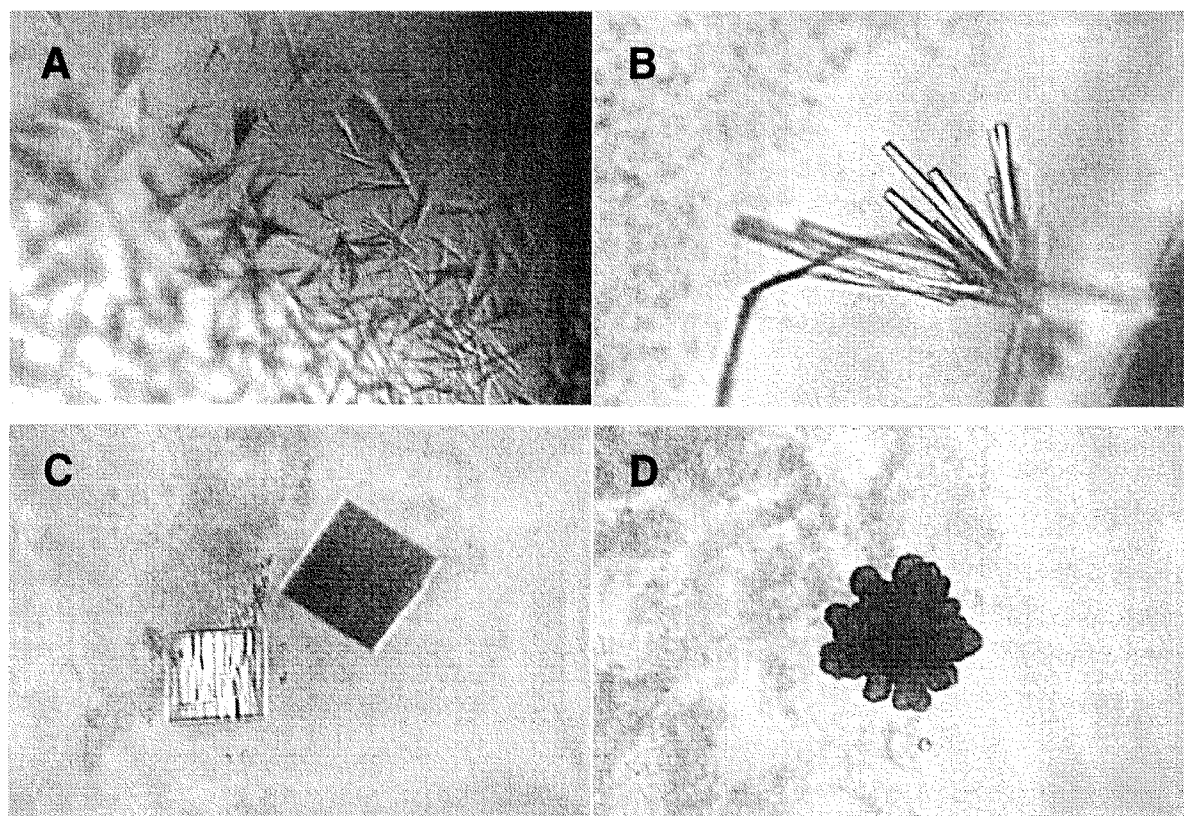
A1.	0.1 M Citric Acid pH 4.0, 0.8 M Ammonium Sulfate
B1.	0.1 M Citric Acid pH 4.0, 1.6 M Ammonium Sulfate
C1.	0.1 M Citric Acid pH 4.0, 2.4 M Ammonium Sulfate
D1.	0.1 M Citric Acid pH 4.0, 3.2 M Ammonium Sulfate
A2.	0.1 M Citric Acid pH 5.0, 0.8 M Ammonium Sulfate
B2.	0.1 M Citric Acid pH 5.0, 1.6 M Ammonium Sulfate
C2.	0.1 M Citric Acid pH 5.0, 2.4 M Ammonium Sulfate
D2.	0.1 M Citric Acid pH 5.0, 3.2 M Ammonium Sulfate
A3.	0.1 M MES pH 6.0, 0.8 M Ammonium Sulfate
B3.	0.1 M MES pH 6.0, 1.6 M Ammonium Sulfate
C3.	0.1 M MES pH 6.0, 2.4 M Ammonium Sulfate
D3.	0.1 M MES pH 6.0, 3.2 M Ammonium Sulfate
A4.	0.1 M HEPES pH 7.0, 0.8 M Ammonium Sulfate
B4.	0.1 M HEPES pH 7.0, 1.6 M Ammonium Sulfate
C4.	0.1 M HEPES pH 7.0, 2.4 M Ammonium Sulfate
D4.	0.1 M HEPES pH 7.0, 3.2 M Ammonium Sulfate
A5.	0.1 M Tris pH 8.0, 0.8 M Ammonium Sulfate
B5.	0.1 M Tris pH 8.0, 1.6 M Ammonium Sulfate
C5.	0.1 M Tris pH 8.0, 2.4 M Ammonium Sulfate
D5.	0.1 M Tris pH 8.0, 3.2 M Ammonium Sulfate
A6.	0.1 M Bicine pH 9.0, 0.8 M Ammonium Sulfate
B6.	0.1 M Bicine pH 9.0, 1.6 M Ammonium Sulfate
C6.	0.1 M Bicine pH 9.0, 2.4 M Ammonium Sulfate
D6.	0.1 M Bicine pH 9.0, 3.2 M Ammonium Sulfate

**Tab. 3.4 Grid Screen PEG 6000 reagent formulations**<sup>12,16,19</sup>

A1.	0.1 M Citric Acid pH 4.0, 5% Polyethylene Glycol 6000
B1.	0.1 M Citric Acid pH 4.0, 10% Polyethylene Glycol 6000
C1.	0.1 M Citric Acid pH 4.0, 20% Polyethylene Glycol 6000
D1.	0.1 M Citric Acid pH 4.0, 30% Polyethylene Glycol 6000
A2.	0.1 M Citric Acid pH 5.0, 5% Polyethylene Glycol 6000
B2.	0.1 M Citric Acid pH 5.0, 10% Polyethylene Glycol 6000
C2.	0.1 M Citric Acid pH 5.0, 20% Polyethylene Glycol 6000
D2.	0.1 M Citric Acid pH 5.0, 30% Polyethylene Glycol 6000
A3.	0.1 M MES pH 6.0, 5% Polyethylene Glycol 6000
B3.	0.1 M MES pH 6.0, 10% Polyethylene Glycol 6000
C3.	0.1 M MES pH 6.0, 20% Polyethylene Glycol 6000
D3.	0.1 M MES pH 6.0, 30% Polyethylene Glycol 6000
A4.	0.1 M HEPES pH 7.0, 5% Polyethylene Glycol 6000
B4.	0.1 M HEPES pH 7.0, 10% Polyethylene Glycol 6000
C4.	0.1 M HEPES pH 7.0, 20% Polyethylene Glycol 6000
D4.	0.1 M HEPES pH 7.0, 30% Polyethylene Glycol 6000
A5.	0.1 M Tris pH 8.0, 5% Polyethylene Glycol 6000
B5.	0.1 M Tris pH 8.0, 10% Polyethylene Glycol 6000
C5.	0.1 M Tris pH 8.0, 20% Polyethylene Glycol 6000
D5.	0.1 M Tris pH 8.0, 30% Polyethylene Glycol 6000
A6.	0.1 M Bicine pH 9.0, 5% Polyethylene Glycol 6000
B6.	0.1 M Bicine pH 9.0, 10% Polyethylene Glycol 6000
C6.	0.1 M Bicine pH 9.0, 20% Polyethylene Glycol 6000
D6.	0.1 M Bicine pH 9.0, 30% Polyethylene Glycol 6000

### 3.2 The rational approach

A good example for the rational approach following the starting conditions of a known molecule was the mutant Q125N. Starting from the known crystallization conditions for the wild type enzyme (50 mM  $\text{KH}_2\text{PO}_4$ , 20% PEG 8K)<sup>21</sup> first experiments were performed with different concentrations of protein at 4°C. The same experiments were carried out in parallel with the elution buffer alone as blank. After 4 days needles in the range of 100-500  $\mu\text{m}$  length were grown from the protein samples.(fig. 3.1A) Further conditions were screened by varying the salt and PEG concentrations around the starting set up in a range of  $\pm 20\%$ . With higher PEG concentration only the length of the needles could significantly be improved. The thickness did not pass the 20  $\mu\text{m}$  range.(fig.3.1B) This preferred growth in one dimension was tried to be inhibited by the addition of small amounts of glycerol, since it is known that the presence of glycerol up to 10% may influence crystal growth by slowing it down<sup>22</sup>; without any success.



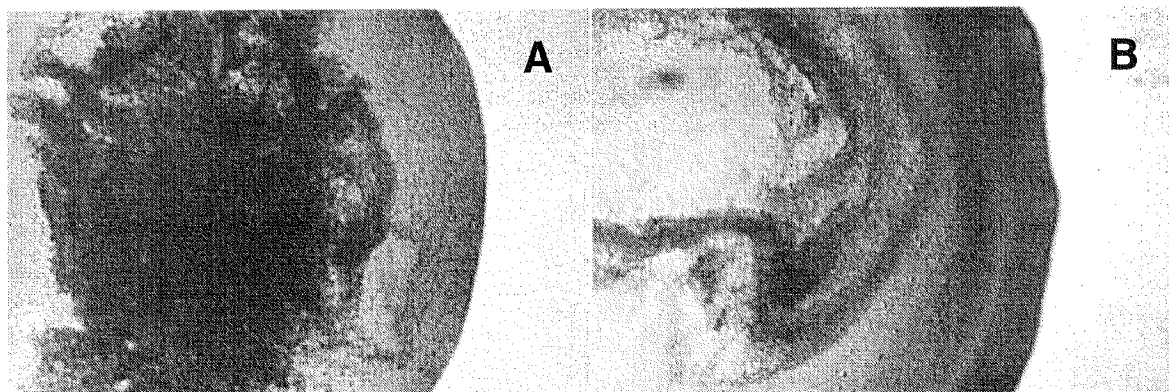
**Fig. 3.1 First attempts to crystallize the active site mutant Q125N with the known conditions from the wild type HSV TK.** A: First needles of Q125N obtained at 4°C with 50 mM  $\text{KH}_2\text{PO}_4$ , 20% PEG 8K. B: Crystals obtained with similar conditions at higher PEG concentrations. C: Plates obtained by varying the volume and the mixing proportion of the drop. D: Reproduction trials of the conditions in C with a new protein batch.

Further optimization involving the change of the volume and the mixing proportion of the drops resulted in first thin plates with dimensions of max. 100 x 200  $\mu\text{m}$ . (fig.3.1C) First X-ray measurement detected diffraction to 7.5  $\text{\AA}$  resolution with some minor reflections at 4.1  $\text{\AA}$ . The crystals were also characterized by high mosaicity, so that the quality had to be improved in respect to a better resolution, and lower mosaicity. Even though a lot of efforts were done also by checking different pH varying from 4.5 to 6.5, no suitable crystals could be reproduced using the same precipitant (fig 3.1D).

### 3.3 The crystallization screenings

The number of variables that influence crystal growth is too large to enable exhaustive screening to be carried out. If the amount of protein is limited, the task is even harder. To restrict the number of experiments, it may be necessary therefore to make a choice between a „knowledge-based“ approach, where only the physical properties of the protein in question are considered, e.g. limited screen conditions in a pH range around pI value of the protein (see Tab.3.3-3.4), and sparse matrix sampling, by which a range of trial conditions is pre-selected. Using a database of successful protein crystallization conditions Jancarik and Kim<sup>14</sup> followed by Cudney et coworkers<sup>13</sup> have determined sets of different conditions for random sampling. In this way a wide range of pHs, counterions and precipitating agents can be rapidly screened with no more than few milligrams of protein or nucleic acid. Although these screening protocols cover a wide range of successful crystallization conditions, some important macromolecules still slip through this net.

The human thymidine kinase has shown to be one of those enzymes, because no crystals, beside salt (see 3.4) could be grown from these random sampling procedures, which were set up with different protein concentrations at 4°C, 16°C and 23°C.

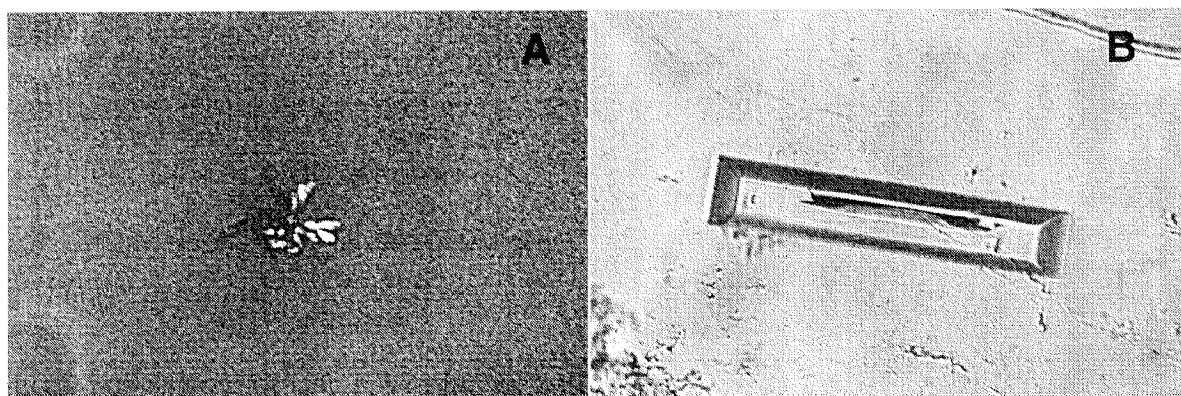


**Fig. 3.2 Typical view of drops containing protein precipitate.** A and B shows two examples obtained using Jancarik screen with human TK1 at high concentrations.

The same experiments were set up with the wild type HSV 1 TK (in collaboration with R. Perozzo) and the other active site mutants. In general a lot of precipitate was observed. (fig. 3.2)

### 3.4 Crystallization of artefacts

The first attempts performed with the human TK were set up with the same conditions of the wildtype TK<sup>21</sup>. Here the first crystals could be grown very fast, but only to a small size. The variation of precipitant did not improve crystal growth. Changes in salt can sometimes produce crystals of varied quality, morphology, and in some cases diffraction properties<sup>12</sup>. Taking this into consideration different attempts were performed by replacing potassium ion by ammonium and phosphate by sulfate. The crystal size could be slightly improved, but only to a size of 50-70  $\mu\text{m}$  (Fig. 3.3A).



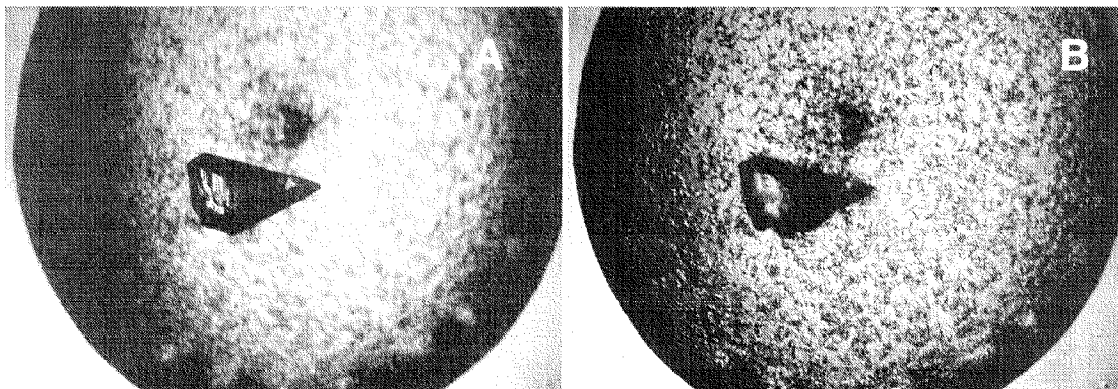
**Fig. 3.3 Crystallization of salts - A normal experience in crystallographer's life.** A. Phosphate salt crystals produced attempting to crystallize the human TK1. B. Imidazole crystals grown from 25% saturated ammonium sulfate containing 100 mM MES pH 6.5.

The crystals showed beautiful colors and polarized light very strongly. They could be reproduced in size and shape, but no further growth could be achieved. First X-ray measurements showed single reflections at very high resolution, revealing clearly the presence of salt crystals: “A normal experience in crystallographer’s life?” Yes. It is not nice, but emphasizes the importance of the X-ray measurement, which should be done as soon as possible to avoid to work in the wrong direction and save time.

A further interesting example of crystallizing an artefact was the experience made with the His-tag-TK1. A crystallization screening performed with protein samples containing imidazole from the purification (see 2.3.2.) lead to beautiful crystals with dimensions around  $300 \times 100 \times 100 \mu\text{m}^3$  (Fig.3.3B). These crystals could also be reproduced with blanks set up with the elution buffer alone, showing here the importance of setting up blank experiments. Nevertheless, the X-ray experiment was performed to confirm the presence of a small organic molecule crystals.

These two experiments show the role of contaminants, such as mono- or bivalent cations or imidazole, present in the protein solution, which may be responsible for false positive results.

The presence of contaminants alone is not always the cause of false positive results. Another important factor is high precipitate concentration. This could be shown in a crystallization experiment with the active site mutant Y101F, where after few weeks large crystals grew from the microcrystalline precipitate in 46% saturated ammonium sulfate containing 100 mM HEPES pH 8.0. (Fig.3.4).



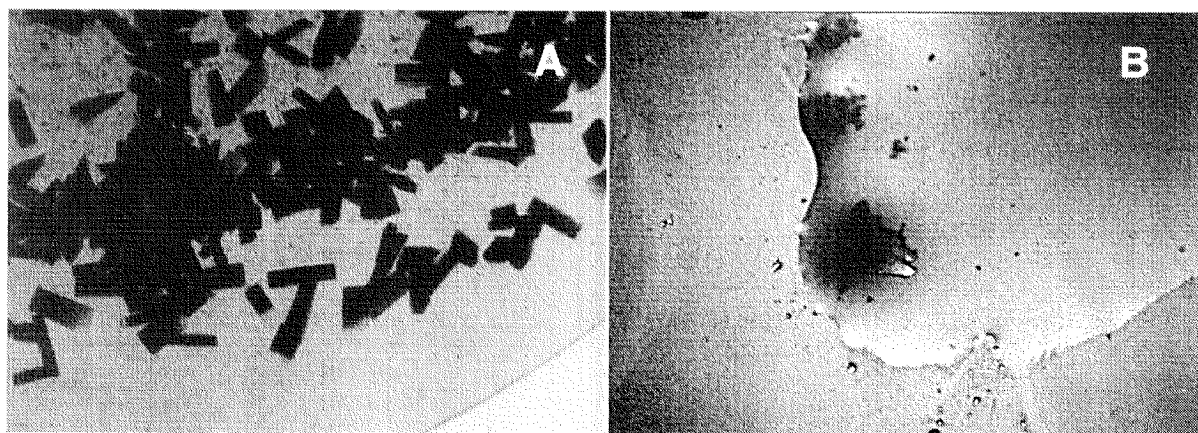
**Fig. 3.4 Ammonium sulfate crystal grown within two weeks from a drop containing micro-crystals of Y101F. A. Focus on the salt crystal, B. Focus on the micro-crystals**



From the same conditions at lower salt concentrations (22-26% saturated ammonium sulfate) protein crystals could also be grown (see 3.6) In contrast to the salt crystal, the protein crystals grew much faster. Even though the growth rate can not be used to distinguish between salt and protein.

### 3.4.1 Izit<sup>®</sup> protein or salt?

A crystal! Is it protein or is it salt? This is the first question one asks himself when the first crystal appears from a crystallization drop. If there is an X-ray facility in house it is not such a problem. One could mount the crystal in question and take a quick look at the diffraction pattern. If such a facility is not available next the door, it is always a question that one would like to answer before covering long distances to the available X-ray source. Well, there is always the crush test. One simply takes a micro needle or small probe and crushes the suspect crystal. A click, or solid crunch is indicative of salt while a powder or silent destruction might indicate one just destroyed a perfect protein crystal. If the crystals are large enough one can put a crystal between two fingers and rub. Salt crystals feels like sand while protein crystals are not palpable.[Personal communication E. Meyer] This procedures are applicable if more than one crystal were grown. In the other case one should prefer to cover a long distance instead of crushing the only crystal on-hand.



**Fig. 3.5 Is it protein or salt?** A. Crystals of the active site mutant Q125N colored with Izit<sup>®</sup>. B. Negative test with magnesium phosphate crystals grown in a crystallization condition of crystal screen I.



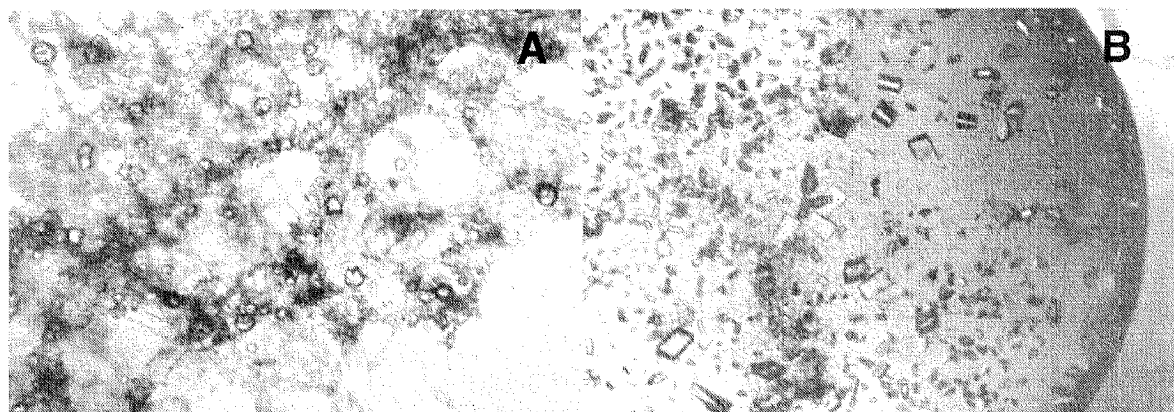
A solution to this problem might be Izit<sup>®</sup> [Hampton Research<sup>®</sup>]. It is a small dye which penetrates the solvent channels in protein crystals coloring the crystals blue. Small molecules or inorganic crystals do not possess these large solvent channels and can not absorb the dye, leaving behind a clear crystal and a blue drop.

The procedure is very simple: One just need to place one microliter of Izit<sup>®</sup> in the sample drop and wait for about an hour or even more, depending on the crystal packing. With the appropriate dilution, Izit<sup>®</sup> will leave a clear drop with blue crystals.

A typical coloring of protein crystals is shown in fig.3.5. This procedure was usually used to test the crystals before bringing them to the X-ray facility.

### 3.5 The effect of detergents

At the beginning experiments were always performed in the presence of Triton-X-100 to stabilize the enzyme. After the approaches performed with the active site mutant Q125N at 4°C other known crystallization conditions<sup>23</sup> for the wild type TK were set up at room temperature. In the first ammonium sulfate screening in the presence of Triton-X-100 no clear crystal growth could be achieved. Only some small clusters diverging from the precipitate could be observed in the drop. At the same time, experiments performed with the wild type full length TK (in collaboration with R. Perozzo) indicated a negative effect of the detergent. A Jancarik screen performed in the absence of Triton-X-100 showed crystal growth in one condition. Consequently, a new ammonium sulfate screen was set up in the absence of Triton-X-100 and the results were impressive. Over night a shower of beautiful crystals of different size was grown. The difference between the presence and absence of detergent is shown in Fig.3.6. The drops without detergent were much clearer and contained less precipitate.



**Fig. 3.6 Difference between presence and absence of detergent.** A. Crystallization trial in ammonium sulfate with Q125N containing Triton-X-100. B. Same crystallization trial of a protein sample without Triton-X-100.

This observation was the basis for the optimization of all further conditions, which lead to the successful crystallization and structure solution of the two mutants Q125N and Y101F presented in chapter 4 and 5.

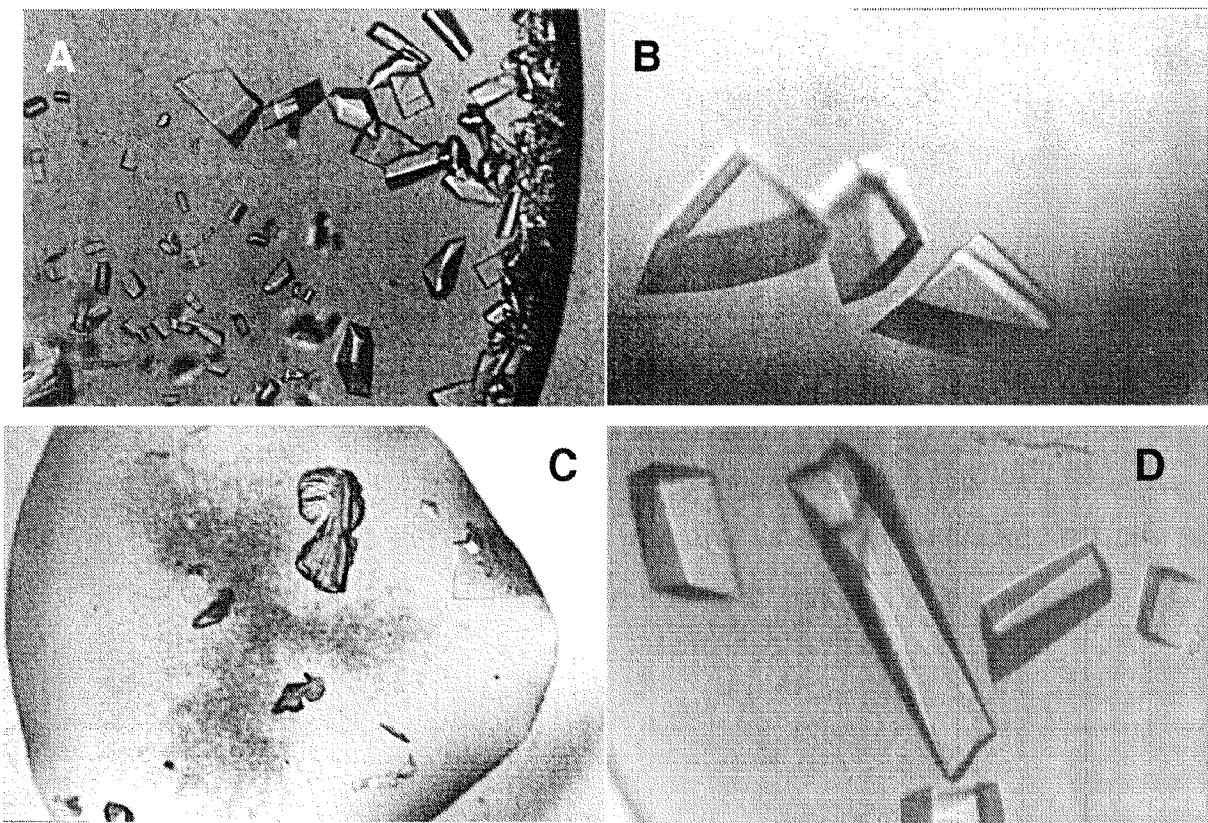
### 3.6 Precipitating agents and their concentrations

Protein precipitants fall into four broad categories: salts(a), organic solvents(b), long-chain polymers(c) and low-molecular-mass polymers together with non-volatile organic compounds (d). The first two classes are represented by ammonium sulfate and ethanol respectively, while the third class contains higher polymers like PEG 4000. In the fourth category MPD and PEG 400 are the common compounds. The latter group is often used as additive to improve crystal growth<sup>12</sup>.

Here I would like to focus on the first category because of the positive experiences made with it. Precipitants and therefore also salts exert their effect by dehydrating proteins through competition with water molecules. Their ability to do this is proportional to the square of the valences of the ionic components<sup>24</sup> Thus, multivalent ions, particularly anions are the most efficient precipitants. Sulfates, phosphates and citrates have traditionally been employed with success.

One might think there would be little variation between different salts so long as their ionic valences were the same, or that there would be variation with two different sulfates such as  $\text{Li}_2\text{SO}_4$  and  $(\text{NH}_4)\text{SO}_4$ . This was the case as shown with the experiments performed with the active site mutants Q125N and Y101F.

The first crystals of both mutants were obtained from ammonium sulfate screens. Although a lot of efforts were done to improve the crystal size by fine tuning the conditions, no suitable crystals for X-ray measurements could be obtained. By changing the conditions to lithium sulfate, differences in crystal formation were observed (Fig.3.7A&B). In contrast to the ammonium sulfate conditions, in which many small crystals were grown, the lithium sulfate conditions allowed to slowly grow more ordered and larger crystals.



**Fig. 3.7 The effect of different salts and concentrations in crystal growth.** A. Crystals of the active site mutant Q125N grown from ammonium sulfate pH 8.0. B. Crystals of the active site mutant Q125N grown from lithium sulfate pH 8.0. C. Crystals grown at lower ammonium sulfate concentrations. D. Crystals grown at higher ammonium sulfate concentrations.

This is in agreement with the general theoretical aspect linked to lithium solvation which are described in the literature. In addition to the salting out effect, there are also specific protein-ion interactions that must be considered, because of the unique polyvalent character of individual proteins, their structural complexity, and the intimate dependence of their physical properties on environmental conditions and interacting molecules<sup>12</sup>. Therefore, it is necessary to examine a broader range of salts.

Not only changes in salt alone can sometimes produce crystals of different quality, morphology, and in some cases diffraction properties. Changes in salt concentration may cause these differences too. In the case of the crystallization of the active site mutants the concentration effect was notable. At lower precipitant concentration the crystals grew more disordered and had the strong tendency to form twins. In contrast this effect was not observed at higher precipitant concentration. The addition of MPD or PEG 400 to inhibit the twin formation was not successful. The typical differences, which were observed independently on the salt (ammonium or lithium sulfate) are shown in fig. 3.7C&D.

### **3.7 The effect of light or vibration**

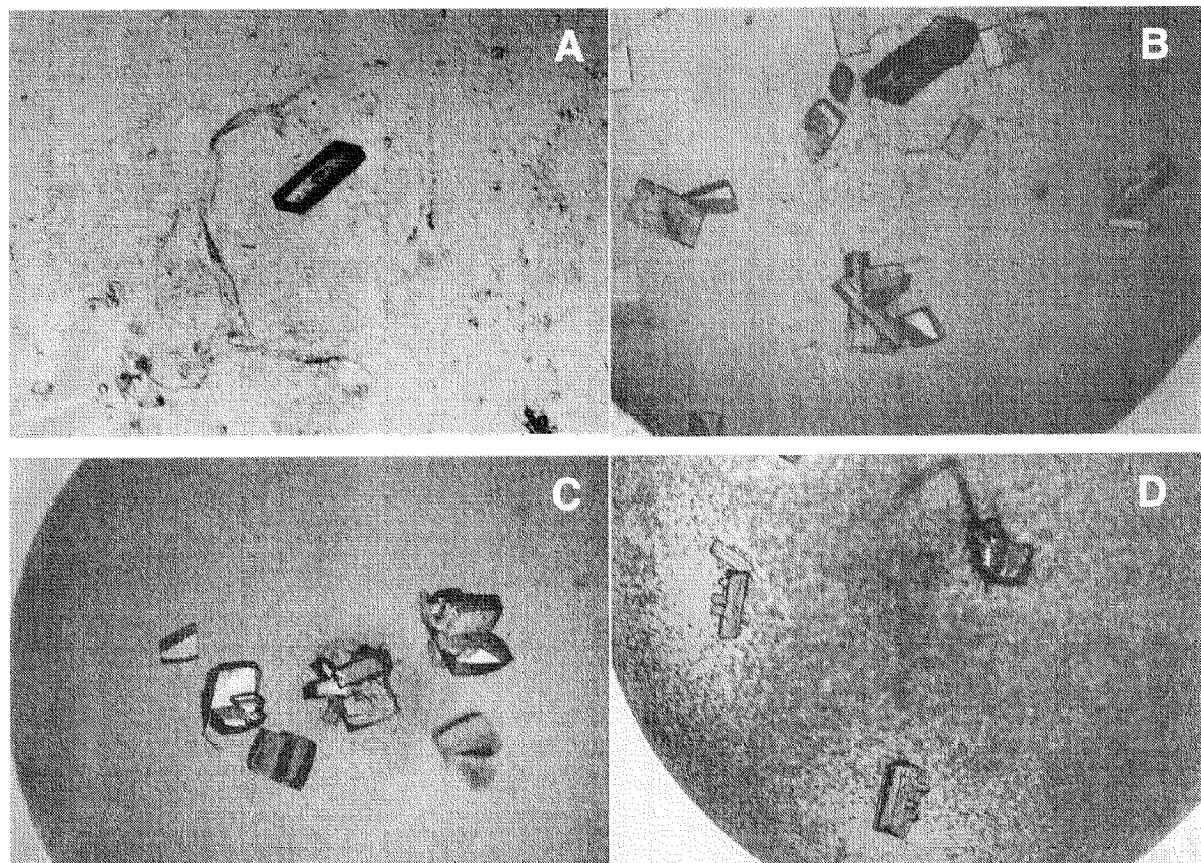
Macromolecular crystallization, like any crystallization, is a multi-parametric process, composed by the three steps: nucleation, growth and cessation of growth. What makes crystal growth of proteins different is firstly the existence of a larger number of influencing factors compared to the small molecule crystal growth and, secondly, the peculiar physico-chemical properties of these compounds<sup>25</sup>. Table 3.5 lists some physico-chemical and biological variables that influence to a greater or lesser extent the crystallization of proteins. The assignment of the appropriate weight to each of the factors is a prerequisite for successful crystallization, but it is also a difficult task because of the nature of proteins that can change by every single mutation (see 3.10). Therefore each factor may differ considerably in importance for each individual protein<sup>12</sup>.

**Tab.3.5 Factors affecting the crystallization of macromolecules.**

1. pH and buffer
2. ionic strength
3. temperature
4. concentration of macromolecule and precipitant
5. purity of macromolecules
6. additives, effectors and ligands
7. source of macromolecule
8. gravity, convection and sedimentation
9. vibrations and sound
10. electric and magnetic fields
11. contamination by microbes
12. volume of crystallization sample
13. Surface of crystallization vessel

Furthermore the various parameters are not independent from each other and their interrelations may be difficult to discern. Starting from these considerations it is clear that it is not easy to elaborate rational guidelines able to increase the probability of success in crystallizing a protein.

However, for a rational design of growth conditions, physical and biological parameters have to be known and controlled. After setting up a crystallization experiment one should observe accurately the samples in order to be able to set up the final crystallization protocol. At the beginning it's worth to have a look at the experiments once a day, especially to get an impression of what is going on in the sample. Once the right crystallization conditions are found crystal growth should be disturbed as less as possible. Fig. 3.8 summarizes the experience linked to the effect of light.

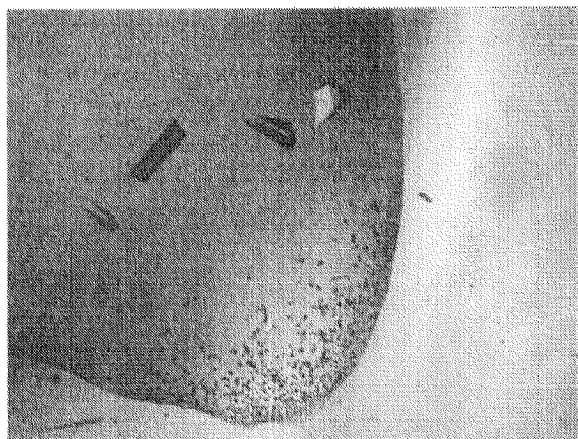


**Fig. 3.8** Effect of light on the crystallization of two active site mutants of HSV 1 TK. A: Single crystal of Q125N grown over night at 23°C. B: Same drop 2h after inspection under microscope. C: Single crystal of Y101F grown in lithium sulfate at pH 8.0. D: Effect of light inspection on Y101F crystals.

Here the influence of light on the crystal growth of the active site mutants Q125N and Y101F is presented. Looking too early at the experiments lead to an induced nucleation resulting in growth of a lot of additional small protein crystals. Once the crystal growth was arrested, no further nucleation could be induced. This effect was observed in both mutants but not in the wildtype enzyme. Thus, to grow suitable crystals for X-ray measurements, the samples were set up and not inspected for at least one week.

Not only light is able to induce secondary nucleation, but also any type of waves. One may remember the first experiment in organic chemistry holding in one hand a flask filled with an ice cold solution of an organic compound, in the other hand a glass rod to rub against the wall of the flask to induce crystallization. The same effect can happen in the drop, if one does not move the boxes carefully enough causing vibrations which may tear up the drop. The friction against the wall may

induce crystallization following the same principles. A typical example of such effects is shown in fig. 3.9. At the border of the drop lots of small crystals appeared and inhibited further crystal growth.



**Fig. 3.9 Vibrations may cause secondary nucleation.** Secondary nucleation at the border of a drop, resulting from vibrations.

### 3.8 Seeding

Seeding experiments have been performed with the two mutants Q125N and Y101F according to Ducruix<sup>25</sup>.

With macroseeding a single crystal is introduced into a suitably pre-equilibrated solution aiming the growth of a larger qualitatively ameliorated crystal. For a successful application of this technique it is necessary to maintain constant conditions and especially to avoid dehydration of the drop keeping unaltered the state of supersaturation. The effect of an incorrect manipulation is reflected in unwanted nucleation. Therefore the experiment should be performed in a very humid environment by using large drops to reduce dehydration.

In our case crystals were picked up from the drop with a microcapillary and transferred repeatedly in three drops containing stabilizing solution (mother liquor) as shown in fig. 3.10.



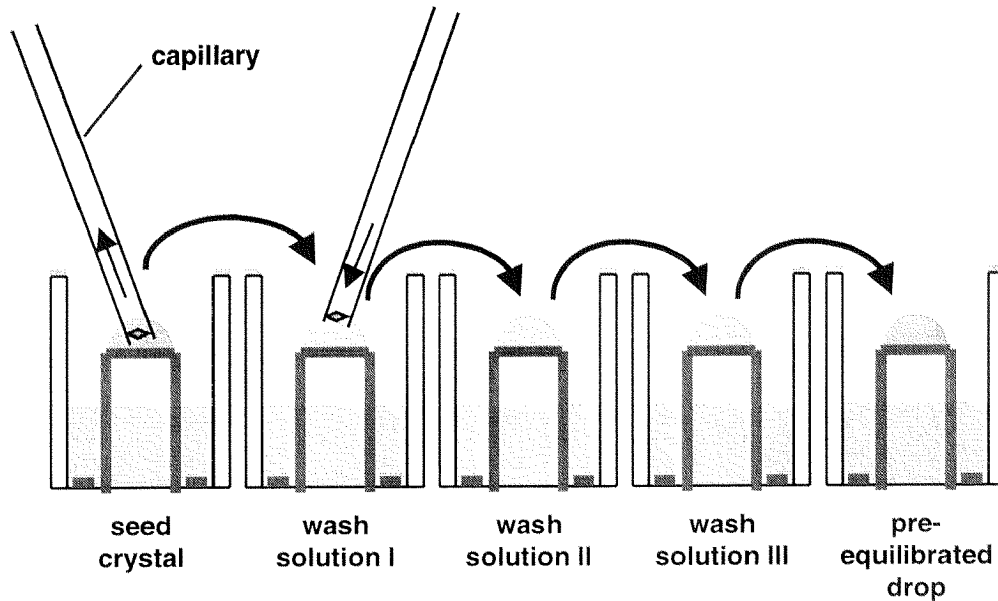


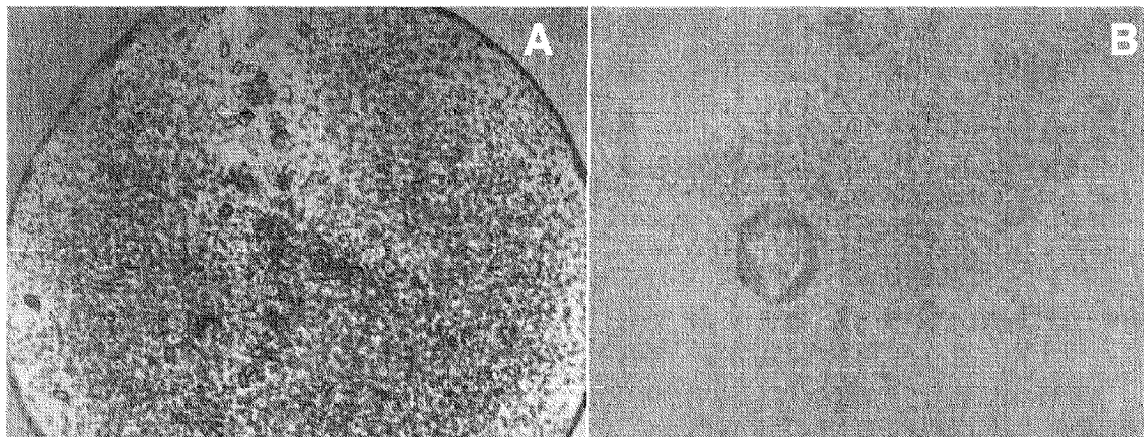
Fig. 3.10 Illustration of the steps involved in the macroseeding technique.

The washed crystals were then transferred to a pre-equilibrated drop containing protein and precipitant solution.

The washing steps are also called etching and are necessary to remove the top layer of protein from the surface of the seed, because it contains possible defects. Besides macroseeding one can try also microseeding with a needle or streak seeding with a whisker. The principle and aim of microseeding are equal to those of macroseeding. The thinner the needle or the whisker, the less seeds are transferred into the new protein solution. Some typical results from seeding experiments are shown in fig. 3.11.

As depicted in fig. 3.11 this procedure did not lead to the desired results. Even though the crystals were washed by transferring them in three drops, too many seeds were still present and transferred into the new drop. In most of the cases the crystal cracked in the equilibrated drop instead of growing. Additionally thousands of small crystals grew from those solutions.





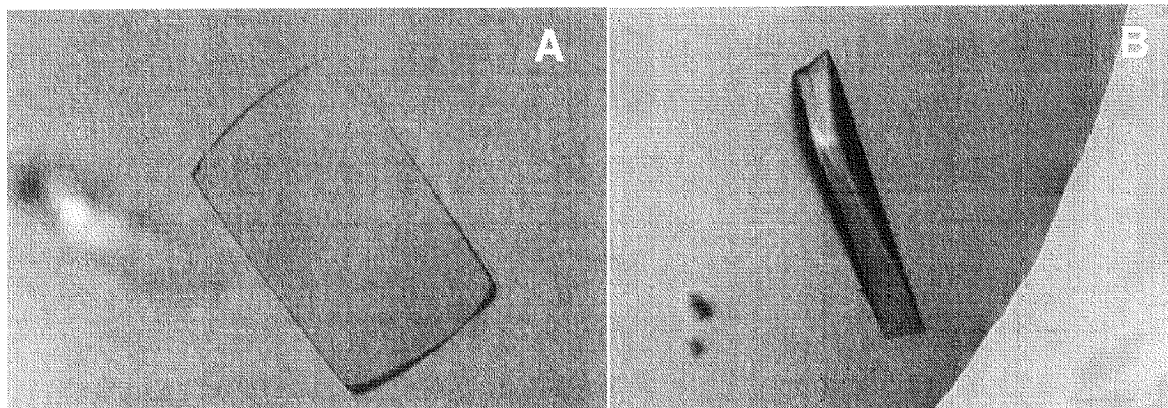
**Fig 3.11 Seeding experiments performed with the active site mutant Y101F and enolase A:** Streak seeding with a whisker: Bigger crystals grew in the direction of the streak. The presence of too many seeds is indicated by the immense growth of very small crystals; B: Makro seeding with an enolase crystal enveloped into protein precipitate (see 3.10.1)

### 3.9 N-/ C-terminus or not?

#### 3.9.1 The active site mutant Y101F

The wild type HSV 1 TK has shown to crystallize in both the full length and the truncated form<sup>21,23,26-28</sup>. A further structure of the active site mutant Q125N, which was N-terminal truncated could also be solved the same way. (see chapter 4) Starting with the second active site mutant Y101F one would imagine a straight forward procedure, since a simple mutation from tyrosine to phenylalanine into the active site is not such a big change. However, the experiences made with this mutant evidence the opposite. Although the same screenings were set up, Y101F did not crystallize under the same conditions as truncated enzyme. The question that arose was whether the cleavage of the N-terminus would influence in such a way the crystallization conditions. Looking at the theoretical pI values of both full length (pI 8.6) and truncated (pI 6.3) enzyme one would expect some differences. Consequently, an ammonium sulfate screen in a range between pH 6.5 and 8.5 was set up for both mutants to see whether a difference between full length and truncated enzyme in the crystallization behavior could be observed. With the “full length” mutant Y101F very small crystals were grown at the border of the drop in 40% saturated ammonium sulfate at pH 8.0. In contrast the truncated enzyme did not show any sign of crystallization.

This pH was then further screened in a range of 20-42% saturated ammonium sulfate with two different protein concentrations. Between 26 and 30%, crystals were grown over night, but the quality was not satisfying. Following the experiences made with the active site mutant Q125N a further screen with lithium sulfate was set up to try to improve crystal quality. With sitting drop, the difference was not so drastic. The same screen set up with hanging drop lead finally to the ideal conditions, from which crystals for X-ray measurements could be grown. The crystal of the active site mutant Y101F in complex with (N)-MCT, which lead to the solution of the structure at 2.4 Å resolution is shown in fig. 3.12.

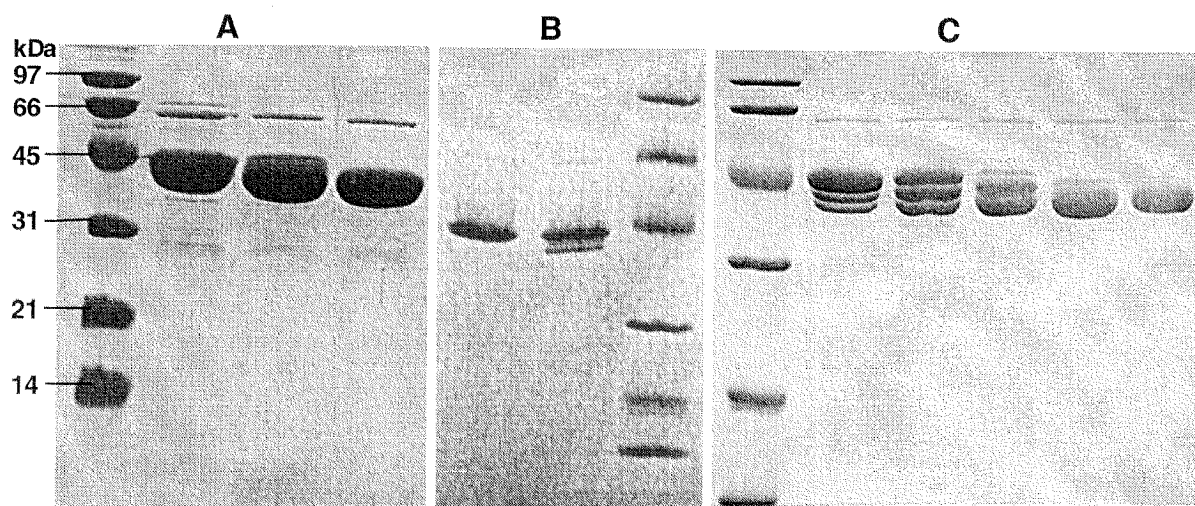


**Fig. 3.12 Crystals of Y101F in complex with (N)-MCT grown within two weeks in 1.18M lithium sulfate at pH 8.0.** The crystals reached maximal dimensions of  $500 \times 350 \times 50 \mu\text{m}^3$  and diffracted X-rays isotropically to beyond  $2.4 \text{ \AA}$  under cryo conditions.

### 3.9.2 The triple mutant H58LM128FY172F

The exchange of the H58/M128/Y172 to the L58/F128/F172 into the active site of HSV 1 TK aiming the mimic of the putative sandwich-like complex of TK with narrow substrate diversity show the role of the amino acids in substrate binding involved in this particular feature. The enzyme shows catalytic activity only towards the natural substrate dT with a 600-fold decrease in affinity. This lost in affinity has been postulated to lie in a different orientation of the base as all the amino acids that form hydrogen bonds to dT are still present. Thermodynamic data indicated a restored flexibility of the enzyme compared to other mutants and let assume the triad to be responsible for a better hydrophobic fit to natural substrates allowing the occurrence of the successive movement for completing the catalytic cycle<sup>29</sup>.

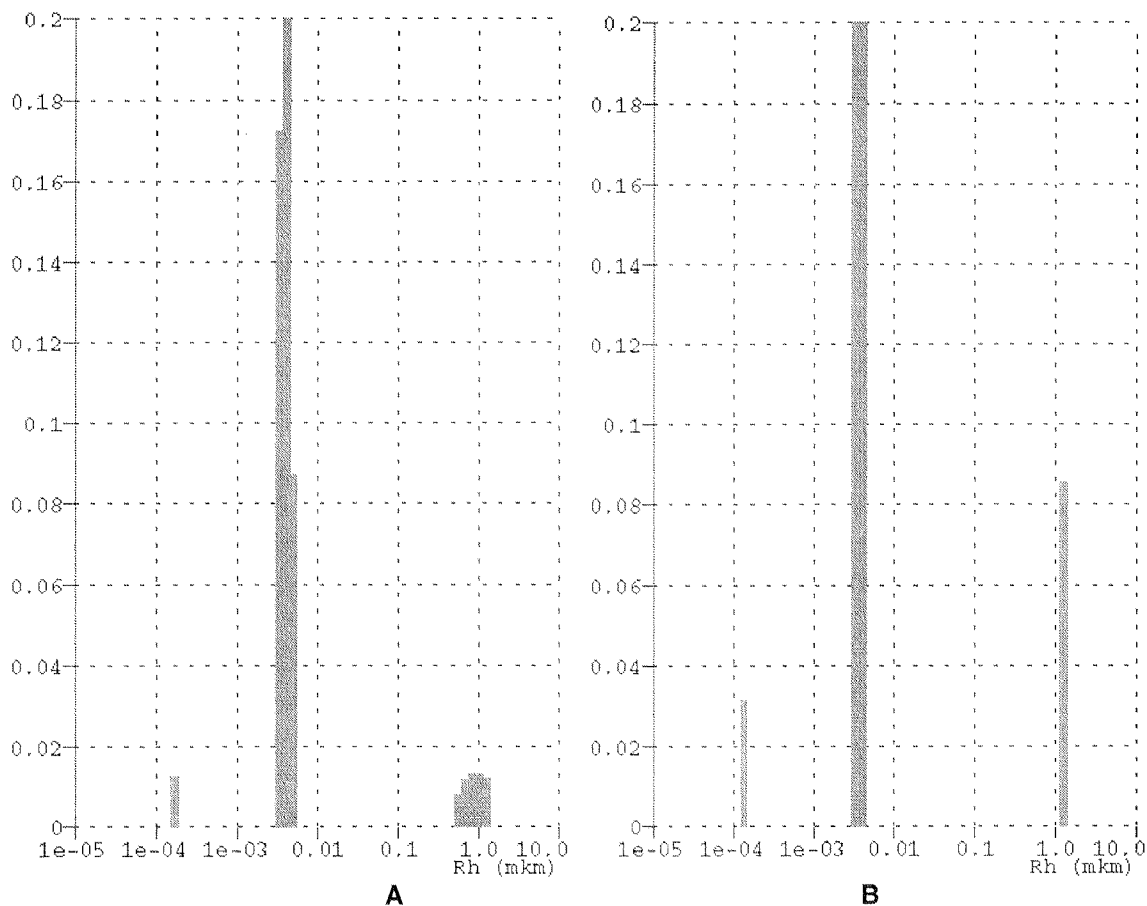
Crystallization experiments were carried on with the aim to see, whether a modified conformation could be observed in the crystal structure. The previously made observations let therefore expect some possible troubles. The first experiences with the purification of this mutant corroborated these hypothesis, since the proteolytic cleavage of the enzyme revealed to be more difficult as compared to the previous mutants. The complete cleavage of the enzyme could be achieved only after 84 h at room temperature (fig. 3.13). After this time the catalytic activity of the enzyme was still present.



**Fig. 3.13** A: Y101F after cleavage over night at room temperature: The cleaved enzyme could not be crystallized. B: Lane 1: Full length wild type HSV 1 TK after PreScission cleavage, Lane 2: Cleavage of Y101F stopped with 5 mM EDTA just after collecting the fractions from the column: this enzyme could be crystallized, Lane 3: SDS-Marker low range C: Lane 1: SDS-Marker low range, Lane 2: Cleavage of H58LM128FY172F after elution from the column, Lane 3: after cleavage over night in solution, Lane 4: after 39 h, Lane 5: after 63 h, Lane 6: after 84 h

The question came up whether the mutation may influence the N-terminus in such a way to obstruct the proteolytic cleavage by thrombin. The residues are quite hidden and far away from the cleavage sites and should not influence the cleavage. Even though the wild type structure does not allow to make some proof statement, since the N-terminus has only been defined starting from residue 46 in all available crystal structures<sup>21,23,26-28</sup>.

Dynamic light scattering measurements were performed and showed for both mutants the same dispersity pattern. In the measured samples the mutant Y101F showed a slight higher particle radius ( $3.94 \pm 0.4$  nm) compared to the radius of the triple mutant ( $3.61 \pm 0.4$  nm). Higher aggregates were also present in both samples (fig. 3.14).



**Fig. 3.14** Size distribution of a sample Y101F(1.0 mg/ml) (A) and H58LM128FY172F(1.0 mg/ml) (B) measured after filtration. Both samples show the same monodispersity with some minor higher aggregates. The non filtered samples could not be measured because of to high fluctuations during the measurements.

Even though both mutants showed the same dispersity pattern, Y101F could be crystallized and H58LM128FY172F not. The same crystallization screenings with ammonium and lithium sulfate used for the active site mutant Y101F were set up with different concentrations of the enzyme. Neither these conditions, nor additional Jancarik screen or further attempts with the standard approaches<sup>21</sup> lead to the growth of crystals.

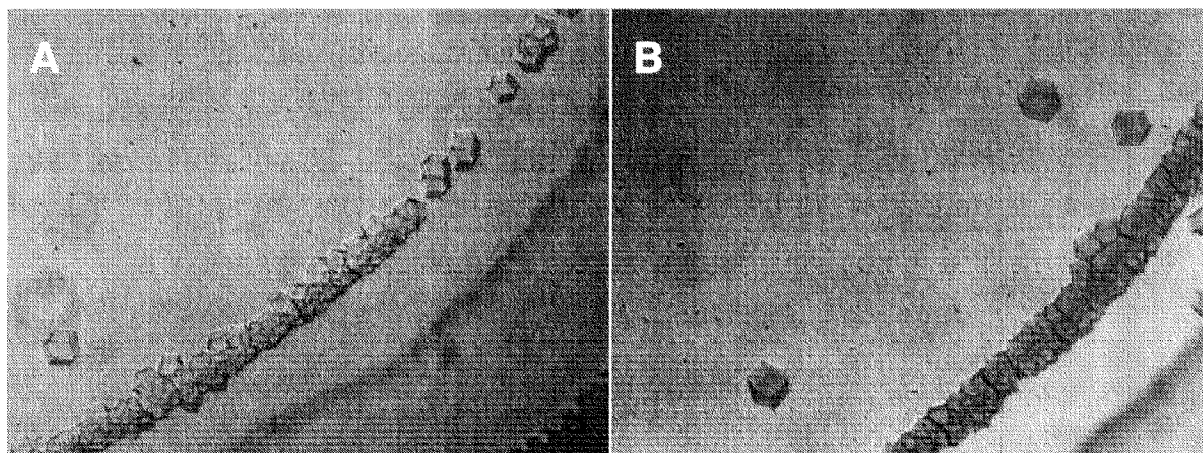
These results support the assumptions of a modified conformation<sup>29</sup>.

### 3.9.3 The deletion mutant $\Delta 40TK1$

The deletion mutant of the human cytosolic TK was constructed based on the observations described in chapter 2, after a lot of failures encountered with the inhomogeneous full length and the His-tag form of the enzyme.

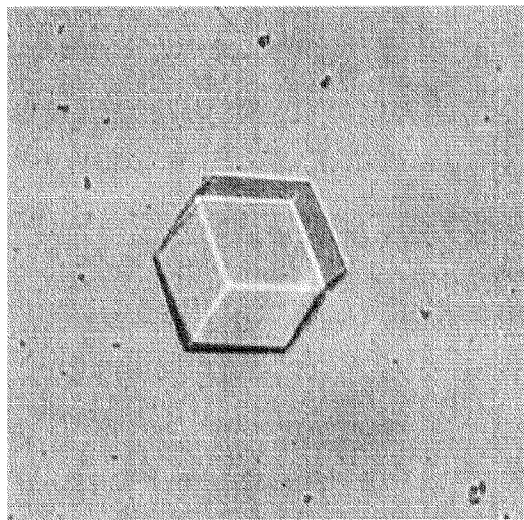
The purified enzyme at a concentration of 2.5 mg/ml was first screened with the Crystal screen I, Crystal screen II (Tab.3.1 &3.2) and ammonium sulfate. For the ammonium sulfate screening the pH ranged from 6.5-9.0 and the concentration of the precipitant was lying between 15 and 80% ammonium sulfate saturation<sup>12</sup>. The screening were performed at room temperature using the sitting drop technique. In both crystal screens I and II a lot of precipitate could be observed. In contrast, the ammonium sulfate screen showed the more promising results at pH 8.0. The precipitate was similar to the one observed in the Q125N trials at high precipitant concentrations, where at the edge of the drop a veil of very fine glittering points could be observed. These conditions were further analyzed by setting up trials with the double protein concentration in a range of 15-40% saturated ammonium sulfate at pH 8.0 with two different buffers Tris and HEPES using sitting drop and hanging drop technique.

With 40% ammonium sulfate small crystals could be grown in the hanging drop conditions containing HEPES buffer. These crystals tested with Izit<sup>®</sup> colored blue, only after 5 hours (Fig 3.15). This may be due to poor crystal packing or simply to the small size of the crystals.



**Fig.3.15** First crystals of the deletion mutant  $\Delta 40TK1$ , which crystallized from a hanging drop equilibrated against 40% saturated ammonium sulfate solution containing 100 mM HEPES pH 8.0.

The concentration of the protein was still relatively low (3 mg/ml). Following the general trend that crystals grow better and slower at low supersaturation<sup>30,31</sup> the protein concentration was increased up to 7.5 and 10.5 mg/ml. For better fine-tuning and reproducibility a new ammonium sulfate screen in HEPES pH 8.0 was made in a range of 0.1 – 2.2 M with increments of 0.1M. This screen showed the trend of the protein to give the best crystals just in a small range between 0.4 and 0.5 M ammonium sulfate. The biggest crystals reached dimensions of 50x50x50  $\mu\text{m}^3$  (fig.3.16) and were analyzed on a rotating anode X-ray facility. No diffraction pattern could be observed.



**Fig. 3.16** Largest crystal of the deletion mutant  $\Delta 40\text{TK1}$  grown from 0.5M ammonium sulfate/HEPES pH 8.0. The crystal grew within two days to a dimension of 50x50x50  $\mu\text{m}^3$  but did not diffract.

To explain the absence of diffraction the following can be hypothesized: Bad diffraction salt, poorly packed protein crystals or the crystal was simply too small. The IZIT<sup>®</sup>-test was so definite that the presence of a salt crystal is rather improbable. Experiences made with the mutant crystals let assume the second suggestion to be the right one. Even three times bigger crystals than these ones had shown no reflections.

So far no further improvement of the size could be achieved.

### 3.10 Enolase from *Alternaria alternata*

Over the past two decades most allergologists have recognized a need for more information on fungal allergy because of increased awareness of the problem and the greater number of patients suffering from asthma and rhinitis due to fungi. In industrialized countries the prevalence of allergic inhalant diseases is some 15-20%. More than 10% of these individuals are sensitized to fungal allergens. Seasonal variation in regard to fungi in the air or to those in homes, and their detection and monitoring but also the biological relevance of the allergenic protein in the fungal organism are new fields of study.

Investigations of fungal air spores have demonstrated that conidium-forming fungi usually predominate over other fungal groups. *Cladosporium herbarum*, *Alternaria alternata*, *Penicillium sp.* and *Aspergillus sp.*, listed in decreasing frequency of occurrence, are most consistently associated with the highest mean percentages of total fungal spore catches.

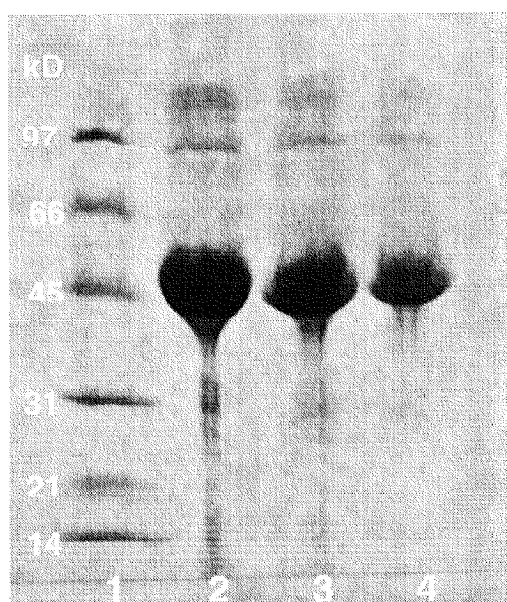
*Alternaria alternata* and *Cladosporium herbarum* are known as major causes of mould allergies. Known habitats of *Alternaria alternata* are soils, corn silage, rotten wood, composts and various forest plants. It is frequently found on wet window frames. It is considered an outdoor mould and appears when the weather is warm. Their enolases play an important role as cross-reactive allergens in fungal species. Even other enzymes like aldehyde dehydrogenase (ALDH), YCP4 (previously found as a *Saccharomyces cerevisiae* protein of unknown function) and the acidic ribosomal protein P2 were found to be allergens in both fungi too. All of these allergens are cytoplasmatic proteins and are highly conserved in evolution of distant species. Lack of knowledge of the identity of these fungal allergens still is a major obstacle for improvement of diagnosis and therapy of allergies to moulds<sup>32</sup>. The structure determination of these enzymes will be an important step in the understanding of fungal allergies and in the development of new and improved methods of diagnosis and therapy<sup>33,34</sup>.



### 3.10.1 Crystallization of Enolase from *Alternaria alternata*

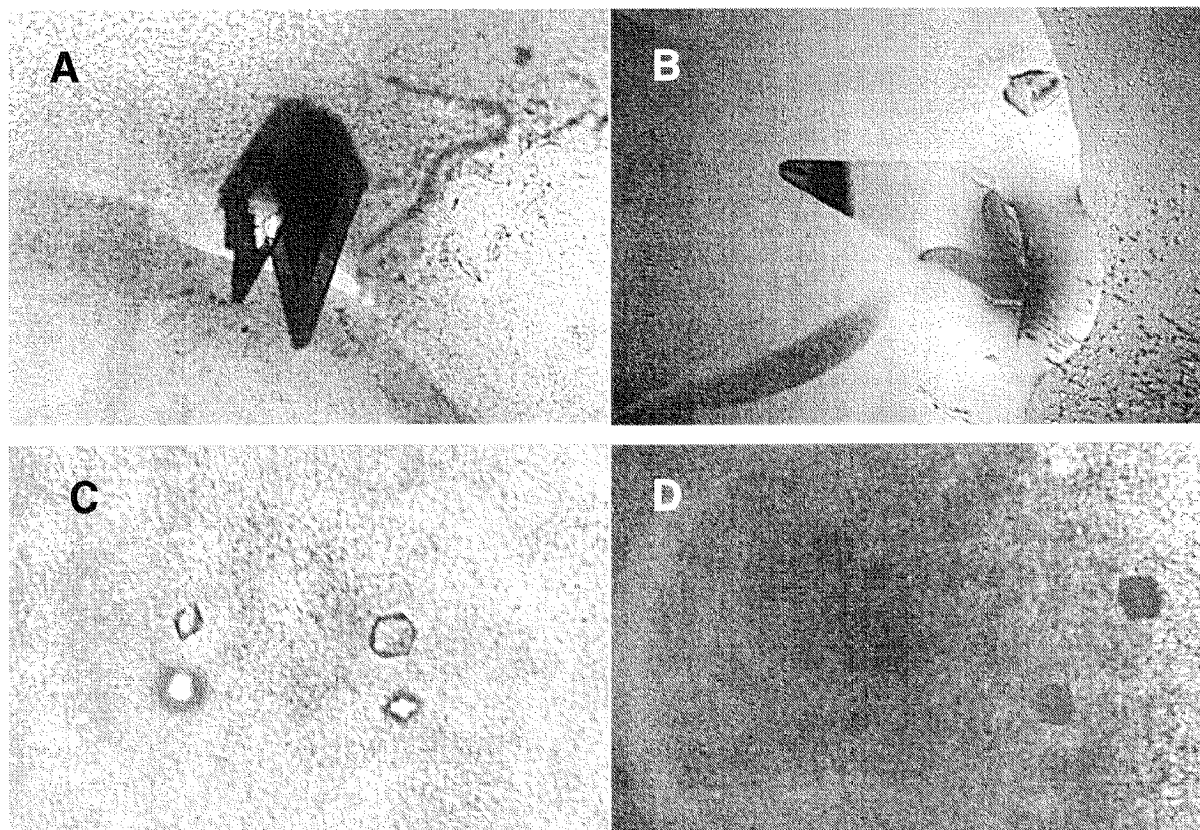
Enolase was produced as recombinant protein in *E.coli* using the coding sequence of the *Alternaria alternata* gene, purified by ion exchange chromatography to a purity of approximately 98% (Fig. 3.17) and lyophilized in diluted phosphate buffer. The purified enzyme binds specific to human IgE.

Before crystallization the lyophilized enzyme was dissolved to a concentration of 1 mg/ml in water and the phosphate buffer was quickly removed from the protein solution by several dilution and ultrafiltration steps. The obtained “phosphate-free” enzyme was finally concentrated to 7.5-15 mg/ml.



**Fig. 3.17 SDS-PAGE of different concentrations of enolase from *Alternaria alternata*:** Lane 1: SDS low-range protein standardsMarker ;Lane 2: Enolase 20 mg/ml in H<sub>2</sub>O ; Lane 3 Enolase 10 mg/ml in H<sub>2</sub>O ;Lane 4: Enolase 5 mg/ml in H<sub>2</sub>O

The first crystallization screening (Jancarik, Tab. 3.1) was carried out at 4°C and 20°C by vapor diffusion using the hanging drop and sitting drop technique in order to get suitable crystals for x-ray diffraction studies. First crystals were grown from condition Nr.6 containing PEG 4K and magnesium chloride. Blanks performed with the same buffer and a phosphate buffer sample prepared in the same way as the protein sample lead to beautiful phosphate crystal growth within one day.(Fig. 3.18)



**Fig. 3.18 Crystallization of enolase of *Alternaria alternata*.** A: First crystals obtained from Jancarik screen condition 6, B: Reproduction of the crystals with diluted sodium phosphate buffer, C: Crystals of enolase grown from 28% PEG, Tris pH 8.0, D: Positive Izit test of the enolase crystals.

In contrast to the previously described crystallization conditions of two homologue enolases from yeast and lobster<sup>4,5</sup> an ammonium sulfate screen set up at room temperature with 20-70% saturated ammonium sulfate in a pH range from 6.5 to 9.0 did not lead to the same results. More promising conditions could be found with a PEG 6K screen (Tab.3.4) performed at 4°C following further published conditions<sup>35</sup>. Crystals grew between 27 and 30% PEG 6K at pH 8.0 reaching maximal dimensions of 60x60x60  $\mu\text{m}^3$  (Fig.3.18 C&D). Different attempts of macroseeding and microdialysis which were performed according to Ducruix<sup>25</sup> to improve the size of the crystals failed. The crystals were enveloped by a gelatinous precipitate in such a way that etching was impossible (Fig. 3.11 B).

In the first x-ray measurements the crystals exhibited not enough reflections because of their small size, but could confirm at least the presence of protein crystals and exclude salt crystals.

### 3.11 Discussion

The advances in recombinant DNA technology in recent years have caused a dramatic effect on the area of protein crystallization. Large amounts of pure protein produced in various expression systems have made it possible to conduct experiments which would have been impossible with material from natural sources. With many more laboratories becoming involved in crystallizing proteins a great deal of new information has been generated on techniques to eliminate the so called "bottleneck of crystallization" in determining the three-dimensional structure of a macromolecule. Major emphasis has been placed in the past decade on kits for screening crystallization conditions of macromolecules<sup>12-14,36</sup>. Such approaches, that assemble the experience of years of crystallographic work, have undoubtedly speeded up the initial screening and, to a certain extent, helped in reducing the amount of protein required for the initial survey.

In spite of all these efforts, the rationalization of crystal growth of biological macromolecules, in terms of having precise recipes, still remains an empirical and frequently exhausting process.

#### *Grid screening*

An effective screening technique often termed "grid screening" evaluates the variance of pH and precipitant concentration for different kinds of precipitants to promote crystal growth<sup>12,18</sup>. Typically, such a crystallization trial utilizes a four by six grid to screen pH (example 4, 5, 6, 7, 8, 9) versus a concentration range of a single precipitant which may be a salt, polymer or organic solvent. If the amount of protein is limited, the range of variables can further be restricted by setting up a "knowledge based" approach in a pH range around the pI value of the analyzed protein. In a similar way one can set up a screen around the known crystallization conditions of a homologue protein. The effectiveness of such a screening technique has been evaluated on this study with the different HSV 1 TK mutants, the human TK and the enolase. The results have demonstrated the possibility to perform crystallization in a rational way, but have also shown that not in all cases this approach leads to the desired hit.

### *The importance of purity*

Purity is certainly one, if not the major, parameter to control in crystallogenesi and lack of its control explains many non-reproducible results. Although crystallization can be used to purify molecules, lack of purity hampers growth of single crystals, especially when the impurities share a structural resemblance with the molecules to be crystallized<sup>37</sup>. Minute amounts of contaminants may interfere with basic phenomena of macromolecular crystal growth. Our observations support fully this statement. In our case, we have shown imidazole or bivalent cations to take in the place of such contaminants which may be responsible for false positive results indicated by the growth of beautiful salt crystals. False positive results may also appear at high precipitate concentrations despite high purity of the enzyme. These experiments showed clearly the crystallization of salts to be a normal experience in crystallographer's life.

### *Detergents*

Many detergents have been widely and successfully used as crystallization additive tool to produce and improve the quality of crystals of macromolecules, particularly where aggregation was a serious problem<sup>13,38,39</sup>. Nevertheless, in our study we have seen the opposite, where the removal of detergent lead to crystal growth. It is therefore important to keep in mind that one should investigate a variety of detergents, as one would screen classes of salts, polymers or organic solvents. Evaluating the role of a single detergent in preventing aggregation would be analogous to accepting the role of ammonium sulfate as indicative of the expected results for all salts. It is important that the quantity of detergents as well as precipitants has to be under control.

### *Crystal packing*

Mechanical stability, internal order and diffraction power of protein crystals rely on the nature and strength of packing contacts within crystalline lattices<sup>40</sup>. These contacts involve molecular recognition of complementary regions of protein by hydrogen bonds, van der Waals contacts and salt bridges, which is reminiscent of the interactions occurring at subunit interfaces in oligomeric proteins. Changing solvent conditions can affect the nature and number of these contacts and yield other crystal shape, as it was shown by the variation of cations and precipitant concentrations in the crystallization experiments with the active site mutants, in which growth rate and quality of the crystals could be influenced. Contacts are often located along lattice symmetry axes, thus implying an anisotropy of intermolecular contacts within crystals<sup>40</sup>. Following these lines, it becomes understandable that mutations affecting domain movement may impair crystallizability of proteins by packing contact modification<sup>41</sup>. The encountered difficulties in crystallizing the triple mutant could be explained by such changes in crystal packing. Chemical heterogeneity and conformational flexibility in the overhang termini of the protein, known not to be essential for catalysis, may also cause enough variability in crystal contacts. Why did the active site mutant Y101F crystallize only in the full length and not in the N-terminal truncated form? Why did the deletion mutant of the human TK crystallize and the full length not? Our results clearly demonstrate the influence of such overhang termini on macromolecular crystallization.

### **3.12 Concluding remarks**

The purpose of this chapter was to try answering the question if there are some magics in protein crystallization. Jan Drenth described it in 1988 as follows: "At the beginning, protein crystallization was at the same level as chemistry in mediaeval times. In that period the alchemists tried to make gold. They did not know what they were doing and they never succeeded in making gold. The protein crystallographers were already one step further because they often succeed in growing crystals, but the science behind it is as mystic as the science of the alchemists"<sup>42</sup>.

On this chapter we have set up different approaches to circumvent many of the difficulties encountered during the crystallization experiments. We have learned not to hesitate in setting up a wide range of screening conditions in searching the needle in the haystack and how to prevent unexpected outcome by checking for impurities or avoiding to disturb crystal growth by light or vibrations. The rationality of all this stuff is evident. Nevertheless we are not able to exclude the presence of some kind of magic, because the range of possibilities to be analyzed is extremely wide. Even if the methodology is as rational as possible, the impression that one gets by looking at those beautiful colors and shapes on a polarizing microscope still remains something magic.

### 3.13 References

1. Gilliland, G.L. & Bickham, D.M. The biological macromolecule crystallization database: A tool for developing crystallization strategies. *METHODS: A Companion to Methods in Enzymology* **1**, 6-11 (1990).
2. Gilliland, G.L., Tung, M., Blakeslee, D.M. & Ladner, J. The Biological Macromolecule Crystallization Database, Version 3.0: New Features, Data, and the NASA Archive for Protein Crystal Growth Data. *Acta Crystallogr.* **D50**, 408-413. (1994).
3. Bernstein, F.C. *et al.* The Protein Data Bank. A computer-based archival file for macromolecular structures. *Eur J Biochem* **80**, 319-24 (1977).
4. Duquerroy, S., Le Bras, G. & Janin, J. Lobster enolase crystallized by serendipity. *Proteins* **18**, 390-3 (1994).
5. Lebioda, L. & Brewer, J.M. Crystallization and preliminary crystallographic data for a tetragonal form of yeast enolase. *J Mol Biol* **180**, 213-5 (1984).
6. Dijkstra, B.W. *et al.* Role of the N-terminus in the interaction of pancreatic phospholipase A2 with aggregated substrates. Properties and crystal structure of transaminated phospholipase A2. *Biochemistry* **23**, 2759-66 (1984).

7. Thunnissen, M.M. *et al.* Crystal structure of a porcine pancreatic phospholipase A2 mutant. A large conformational change caused by the F63V point mutation. *J Mol Biol* **232**, 839-55 (1993).
8. Scott, D.L. *et al.* Structures of free and inhibited human secretory phospholipase A2 from inflammatory exudate. *Science* **254**, 1007-10 (1991).
9. Genovesio-Taverne, J.C. *et al.* Crystallization and preliminary X-ray diffraction studies of the spinach-chloroplast thioredoxin f. *J Mol Biol* **222**, 459-61 (1991).
10. Katti, S.K., LeMaster, D.M. & Eklund, H. Crystal structure of thioredoxin from *Escherichia coli* at 1.68 Å resolution. *J Mol Biol* **212**, 167-84 (1990).
11. Weichsel, A., Gasdaska, J.R., Powis, G. & Montfort, W.R. Crystal structures of reduced, oxidized, and mutated human thioredoxins: evidence for a regulatory homodimer. *Structure* **4**, 735-51 (1996).
12. McPherson, A. Current approaches to macromolecular crystallization. *Eur J Biochem* **189**, 1-23 (1990).
13. Cudney, R., Patel, S., Weisgraber, K., Newhouse, Y. & McPherson, A. Screening and Optimization Strategies for Macromolecular Crystal Growth. *Acta Cryst.* **D50**, 414-423 (1994).
14. Jancarik, J. & Kim, S.H. Sparse matrix sampling: a screening method for crystallization of proteins. *J Appl Cryst* **24**, 409-411 (1991).
15. Hampton-Research. Crystal Screen Kit series. (, 1999).
16. Hampton-Research. Grid Screens. (, 1999).
17. Samudzi, C.T. *J. Crystal Growth* **123**, 47-58 (1992).
18. Weber, P. 31-37 (Academic Press, New York, 1990).
19. Weber, P.C. Physical principles of protein crystallization. *Adv Protein Chem* **41**, 1-36 (1991).
20. McPherson, A., Malkin, A.J. & Kuznetsov, Y.G. The science of macromolecular crystallization. *Structure* **3**, 759-68 (1995).
21. Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett* **368**, 289-92 (1995).

22. Sousa, R. Use of glycerol, polyols and other protein structure stabilizing agents in protein crystallization. *Acta Cryst* **D51**, 271-277 (1995).
23. Brown, D.G. *et al.* Crystal structures of the thymidine kinase from herpes simplex virus type-1 in complex with deoxythymidine and ganciclovir. *Nat Struct Biol* **2**, 876-81 (1995).
24. Carbonnaux, C., Ries-Kautt, M. & Ducruix, A. Relative effectiveness of various anions on the solubility of acidic *Hypoderma lineatum* collagenase at pH 7.2. *Protein Sci* **4**, 2123-8 (1995).
25. Ducruix, A. & Giegé, R. *Crystallization of nucleic acids and proteins : a practical approach*, xxiv, 331 (IRL Press at Oxford University Press, Oxford [England] ; New York, 1992).
26. Champness, J.N. *et al.* Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands. *Proteins* **32**, 350-61 (1998).
27. Bennett, M.S. *et al.* Structure to 1.9 Å resolution of a complex with herpes simplex virus type-1 thymidine kinase of a novel, non-substrate inhibitor: X-ray crystallographic comparison with binding of aciclovir. *FEBS Lett* **443**, 121-5 (1999).
28. Wild, K., Bohner, T., Folkers, G. & Schulz, G.E. The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue. *Protein Sci* **6**, 2097-106 (1997).
29. Pilger, B. *et al.* Substrate Diversity of Herpes Simplex Virus Thymidine Kinase - Impact of the kinematics of the enzyme. *J Biol Chem* **274**, 31967-31973 (1999).
30. Boistelle, R. & Astier, J.P. Crystallization mechanisms in solution. *J Cryst Growth* **90**, 14-30 (1988).
31. Durbin, S.D. & Feher, G. Protein crystallization. *Annu Rev Phys Chem* **47**, 171-204 (1996).
32. Achatz, G. *et al.* Molecular characterization of *Alternaria alternata* and *Cladosporium herbarum* allergens. *Adv Exp Med Biol* **409**, 157-61 (1996).



33. Achatz, G. *et al.* Molecular cloning of major and minor allergens of *Alternaria alternata* and *Cladosporium herbarum*. *Mol Immunol* **32**, 213-27 (1995).
34. Breitenbach, M. *et al.* Enolases are highly conserved fungal allergens. *Int Arch Allergy Immunol* **113**, 114-7 (1997).
35. Wedekind, J.E., Poyner, R.R., Reed, G.H. & Rayment, I. Chelation of serine 39 to Mg<sup>2+</sup> latches a gate at the active site of enolase: structure of the bis(Mg<sup>2+</sup>) complex of yeast enolase and the intermediate analog phosphonoacetohydroxamate at 2.1-Å resolution. *Biochemistry* **33**, 9333-42 (1994).
36. Stura, E.A., Satterthwait, A.C., Calvo, J.C., Kaslow, D.C. & Wilson, I.A. Reverse screening. *Acta Cryst.* **D50**, 448-455 (1994).
37. Giegé, R., Lorber, B. & Théobald-Dietrich, A. Crystallogenesi s of Biological Macromolecules: Facts and Perspectives. *Acta Cryst.* **D50**, 339-350 (1994).
38. Reiss-Husson, F. Crystallization of membrane proteins. in *Crystallization of nucleic acids and proteins* (eds. Ducruix, A. & Giegé, R.) 175-193 (IRL Press, Oxford New York Tokyo, 1992).
39. McPherson, A. *et al.* An experiment regarding crystallization of soluble proteins in the presence of beta-octyl glucoside. *J Biol Chem* **261**, 1969-75 (1986).
40. Salemme, F.R., Genieser, L., Finzel, B.C., Hilmer, R.M. & Wendoloski, J.J. Molecular factors stabilizing protein crystals. *J Cryst Growth* **90**(1988).
41. Mittl, P.R.E., Berry, A., Scrutton, N.S., Perham, R.N. & Schulz, G.E. A designed mutant of the enzyme glutathione reductase shortens the crystallization time by a factor of forty. *Acta Crystallogr.* **D50**, 228-231 (1994).
42. Drenth, J. Concluding remarks. *J Cryst Growth* **90**, 368-370 (1988).

Seite Leer /  
Blank leaf

## CHAPTER 4

### **The nucleoside binding site of *Herpes simplex* type 1 TK studied by X-ray crystallography**

Joachim Vogt<sup>#†</sup>, Remo Perozzo<sup>\*†</sup>, Alex Pautsch<sup>#</sup>, Andrea Prota<sup>\*</sup>, Pierre Schelling<sup>\*</sup>,  
Beatrice D. Pilger<sup>\*\$</sup>, Gerd Folkers<sup>\*</sup>, Leonardo Scapozza<sup>\*</sup>, Georg E. Schulz<sup>#</sup>

Institut für Organische Chemie und Biochemie  
Albert-Ludwigs-Universität, Albertstraße 21  
D-79104 Freiburg im Breisgau, Germany

<sup>\*</sup>Department of Applied Biosciences  
Swiss Federal Institute of Technology (ETH)  
Winterthurerstr. 190  
CH-8057 Zürich, Switzerland

<sup>†</sup>These authors contributed equally to this work.

<sup>\$</sup>Present address: Harvard Medical School, Dept. of Biological Chemistry and  
Molecular Pharmacology, 250 Longwood Ave , Boston, MA 02115

Submitted to PROTEINS, Structure, function and genetics

**Running title:** Structure of herpal thymidine kinase

**Keywords:** enzyme-prodrug gene therapy, nucleoside binding, thymidine kinase, X-ray crystallography

**Abbreviations:** ACV, Aciclovir; dT, (2-deoxy)-thymidine; dTMP, thymidine monophosphate; 9-HPA, 9-(2-*rac*-hydroxypropyl)adenine; Lid, a peptide segment at the active center of all nucleoside monophosphate kinases; PMEA, 9-(2-phosphonylmethoxyethyl)adenine; TK<sub>HSV1</sub>, thymidine kinase from *Herpes simplex* virus type 1; TK<sub>HSV1</sub>(apo), unligated full length enzyme; TK<sub>HSV1</sub>(Q125N):dT, the mutant of enzyme TK<sub>HSV1</sub> in complex with dT.

## ABSTRACT

The crystal structures of the ligand-free *Herpes simplex virus* type 1 thymidine kinase, of the enzyme ligated with a 9-substituted adenine, and of the thymidine-ligated enzyme mutant Q125N have been determined at resolutions between 1.9 Å and 2.5 Å. The structure of the ligand-free enzyme reveals that on substrate binding an induced fit seems to be likely and water molecules are displaced from the active center. It could be shown that 9-(2-*rac*-hydroxypropyl)adenine binds in a thymidine-like orientation indicating that the substrate acceptance of the enzyme is broader than expected. The structure analysis of the thymidine-ligated mutant Q125N showed that the mutation altered neither the binding mode of thymidine nor the polypeptide conformation, but that the two major hydrogen bonds to thymidine in the wild-type structure were replaced by a single water-mediated hydrogen bond. These findings provide new insights for structure-based design of new drugs for antiviral therapy and for gene-therapy of cancer.

## INTRODUCTION

Thymidine kinase (TK, EC 2.7.1.21) is a key enzyme in the pyrimidine salvage pathway catalyzing the phosphorylation of thymidine to thymidine monophosphate in the presence of  $Mg^{2+}$  and ATP<sup>1</sup>. In the cell dTMP is subsequently di- and triphosphorylated and finally used as DNA building block.

Thymidine kinase of *Herpes simplex virus type 1* (TK<sub>HSV1</sub>) exhibits a broad substrate diversity for nucleosides in contrast to the cellular enzyme, that accepts only thymidine and structurally close analogs<sup>2-4</sup>. This feature is the molecular basis for the classical prodrug-based antiviral therapy and the use of TK<sub>HSV1</sub> in virus-directed enzyme-prodrug gene therapy of cancer, where the TK<sub>HSV1</sub> gene is introduced into tumor cells by retroviral or adenoviral vectors. Dividing cells expressing TK<sub>HSV1</sub> can then phosphorylate nontoxic nucleoside analogs like ganciclovir to triphosphates, which in turn inhibit cellular polymerases causing cell death and tumor ablation<sup>5-9</sup>.

Clinical trials with ganciclovir are presently running, but up to now the efficacy of this strategy has been limited, because of the immunosuppressive side-effects of the dosages required for tumor regression<sup>10-12</sup>. Other limiting factors like target specificity, poor vector distribution within solid tumors or low transfection efficacy have also become subjects of intense research efforts<sup>13,14</sup>. An improvement of the efficacy of this therapeutic approach is expected by engineering TK mutants with increased substrate specificity towards the prodrug and decreased thymidine utilization<sup>11</sup>. For this purpose the mechanism of binding and the binding structures of substrate to the enzyme have to be understood. Several active site mutants had been constructed to clarify the contribution of each residue to the nucleoside binding<sup>15</sup>. In order to broaden the knowledge on the range of substrate acceptance of TK<sub>HSV1</sub>, the synthesis of new compounds with varying base moieties like *e.g.* 9-Hydroxypropyl-substituted adenine-derivatives was pursued<sup>16</sup>.

These experiments now have to be supported by the structure analyses of the respective enzyme-substrate complexes. Here we present the structure of the full-length TK<sub>HSV1</sub> in its unligated form, following structures of several complexes between N-terminal truncated TK<sub>HSV1</sub> and different substrates<sup>17-21</sup>. Moreover, we

report the TK<sub>HSV1</sub> structure in complex with *rac*-9-hydroxypropyl-adenine (9-HPA). The structure of a mutant affecting the thymidine-binding pocket mutant in complex with thymidine is presented. The water replacement on substrate-binding is discussed as well as the exceptionally low substrate specificity.

## Materials and Methods

### Materials

Restriction endonucleases, T4 DNA ligase and thrombin were from Promega. The plasmid pGEX-6P-2, the PreScission<sup>TM</sup>-protease and the strain BL21 (ompT<sup>-</sup>, F<sup>-</sup>, hsdS(rB<sup>-</sup>,mB<sup>-</sup>), gal) were purchased from Pharmacia. *E. coli* strain 71/18 (F<sup>+</sup>, *lacI*<sup>q</sup>, *lacZ*ΔM15), which produces high levels of *lac* repressor to suppress uncontrolled expression from the inducible *ptac* promoter, was used for the cloning procedures. Reagents for enzyme assays and crystallization were obtained by Sigma. Linbro plates and cover slides were purchased from Hampton Research, the sitting drop bridges were from DROP. 9-HPA was synthesized and generously supplied by Dr. M. Mintas of Zagreb University.

### Preparation of the expression vectors

The gene for the full length TK<sub>HSV1</sub> was taken from plasmid pGEX-2T-TK<sup>22</sup> by cleavage with *Bam*HI and *Eco*RI. It was then inserted into the *Bam*HI- and *Eco*RI-digested plasmid pGEX-6P-2 to yield plasmid pGEX-6P-2-TK. The production of mutant Q125N has been described earlier<sup>15</sup>.

For sequence verification competent *E. coli* 71/18 were transfected with the pGEX-6P-2-TK DNA. After DNA isolation of several clones, both strands of the entire TK gene were sequenced, using the dye terminator method (ABI PRISM<sup>TM</sup> 310) to confirm the absence of further mutations.

### Protein expression and purification

The expression of the proteins was performed according to a modified version of a previously published protocol<sup>22</sup> using *E. coli* strain BL21. Recombinant full length TK<sub>HSV1</sub> was isolated from *E. coli* as C-terminal part of a PreScission™-cleavable GST-fusion protein, after expression using the pGEX-6P-2-TK plasmid. The active site mutant Q125N was expressed using the pGEX-2T-Q125NTK plasmid and isolated as thrombin-cleavable GST-fusion protein. Both proteins were purified using glutathione affinity chromatography. The crude extract derived from 1 liter bacterial culture was applied five times to the glutathione sepharose column. On column washing was performed with buffer A (1M NaCl in 50 mM Tris pH 7.5, 10% glycerol and, 1 mM DTT, 0.1% Triton X-100), buffer B (250 mM KH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, 150 mM NaCl, 0.1% Triton X-100, pH 7.0) and buffer C (250 mM KH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, 150 mM NaCl, 0.1% Triton X-100, 10 mM MgATP, pH 7.4). After these washing steps, the proteolytic cleavage of the column-bound fusion proteins with PreScission™ and with thrombin yielded pure wild-type TK<sub>HSV1</sub> and Q125N mutant TK<sub>HSV1</sub> protein, respectively. Expression and purification were monitored by SDS-PAGE.

### Crystallization

The purified enzymes were concentrated by ultrafiltration (Centricon 30) to 25 mg/ml. Crystallization screenings were carried out at 4, 16, 20 and 23 °C by sitting and hanging drop vapor diffusion. Crystals grew in 0.9 – 1.2 M Li<sub>2</sub>SO<sub>4</sub>, 1 mM DTT and 0.1 M HEPES buffer at pH 7.5 - 8.0. With mutant Q125N we added 1 mM dT to the crystallization solution. Best crystals appeared in 4 μL drops within 1 to 2 days at 23°C and grew over one week up to sizes of 350 x 200 x 200 μm<sup>3</sup>. In a soaking experiment wild-type crystals were transferred into crystallization solution containing 5 mM 9-HPT kept there for 30 min. Subsequently, the crystals were bathed in crystallization buffer containing 30 % (v/v) glycerol for 1 minute and flash-frozen to 100 K.



### **X-ray data collection**

Data were collected using a wire-frame detector (model X1000, Siemens) with Cu K<sub>α</sub> radiation from a rotating anode generator (Rigaku, model RU200) run at 45 kV and 120 mA. For two data sets we used synchrotron radiation at beamlines X11 and BW7B of the EMBL-outstation (Hamburg) with MAR345 image plates. The image plate data were integrated with program MOSFLM<sup>23</sup>, and the wire-frame data with program XDS<sup>24</sup>. All data were further processed using the programs SCALA and TRUNCATE<sup>23</sup>. Solvent accessible surface areas were calculated with the Molecular Surface Package<sup>31</sup>. Geometric analysis of the protein structures was carried out using program PROCHECK<sup>42</sup>.

### **Refinement**

The raw starting model was the structure of the complex TK<sub>HSV1</sub>:dTMP:ADP<sup>21</sup> without ligands and water molecules was used as a raw starting model. The refinement and difference-Fourier electron density map calculations were done using program REFMAC<sup>23</sup>. For model manipulations we used program O<sup>25</sup>. The programs MOLSCRIPT<sup>26</sup> and RASTER3D<sup>27</sup> were used for figure production. The coordinate data-sets have been sent to the Protein Data for deposition.

**Table 1** Data collection and refinement statistics

Data set <sup>1</sup>	substrate-free TK <sub>HSV1</sub>	TK <sub>HSV1</sub> - (Q125N):dT	TK <sub>HSV1</sub> :9-HPA
Diffraction data			
X-ray source	X11	Cu K <sub>α</sub>	BW7B
Unit cell [Å]	a = 113.9	a = 113.4	a = 113.0
	b = 117.3	b = 117.5	b = 116.9
	c = 107.9	c = 108.6	c = 108.0
Resolution range [Å]	20 - 1.9	30 - 2.5	20 - 1.9
Completeness [%]	99	92	96
Multiplicity	3.8	2.4	3.3
Unique Reflections	56856	23273	54250
R <sub>sym</sub> <sup>2</sup> (last shell) [%]	3.9 (20)	6.3 (17)	5.0 (42)
I/σ (last shell)	26.5 (4.8)	11.5 (3.1)	12.5 (2.7)
Refinement and final model			
R-factor (R <sub>free</sub> ) [%]	21.5 (24.9)	21.2 (27.3)	21.7 (26.9)
Average B-factor [Å <sup>2</sup> ]	33	28	42
Number of protein atoms	4714	4731	4684
substrate atoms	-	34	28
Water molecules	277	111	221
Sulfate molecules	2	2	2

<sup>1</sup> Crystals were of space group *C222*<sub>1</sub>. All data sets were collected at 100 K.

<sup>2</sup>  $R_{\text{sym}} = \frac{\sum_h \sum_i |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i I_{hi}}$  where h stands for the unique reflections and i counts through symmetry-related reflections.

## RESULTS AND DISCUSSION

### High resolution structure of substrate-free TK

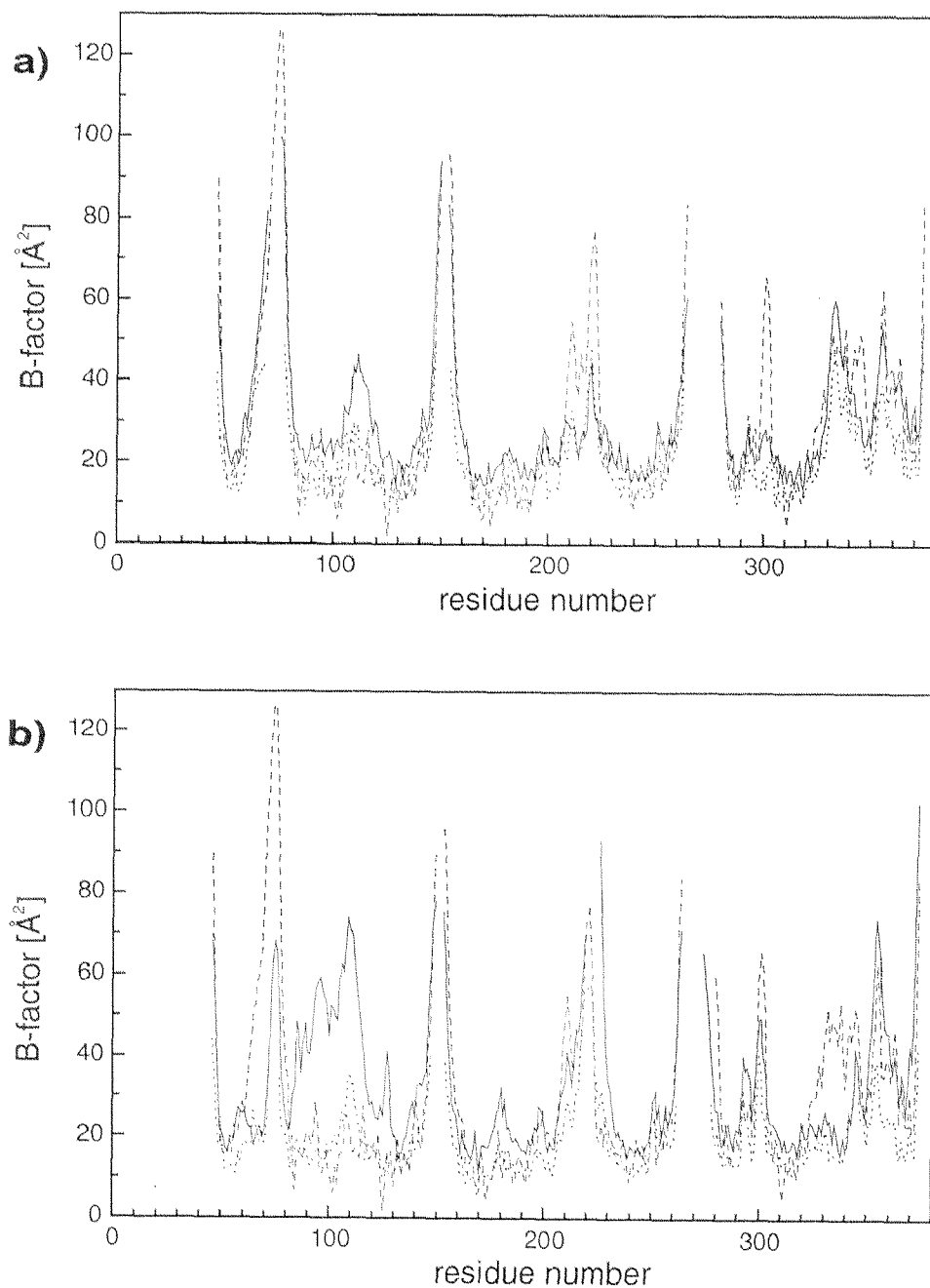
TK<sub>HSV1</sub>(apo) crystallized in space group  $C222_1$  with two subunits of 41 kDa per asymmetric unit. The structure was refined to an R-factor of 21.5 % ( $R_{\text{free}} = 24.9\%$ ) at 1.9 Å resolution with data collected from one crystal at 100 K (**Table 1**). The model is of good quality since 93 % of the residues are in most favored and 6 % in additionally allowed regions of the Ramachandran plot. Only the catalytically competent residue Arg163 is in a disallowed region, in agreement with other TK<sub>HSV1</sub> structures<sup>17-21</sup>. The model contains 276 water molecules and 2 sulfate molecules that are bound to the two P-loops in the beta-phosphate position of ATP as usual<sup>28</sup>. The N-terminal 45 residues have no electron density in the crystal structure, indicating that they are mobile. Other missing parts of the model are 23 residues of subunit A (positions 70-74, 150-152 and 265-279) and 18 residues of subunit B (positions 150-153, 220-225 and 265-273). In both subunits there are several mobile parts with *B*-factors higher than 60 Å<sup>2</sup> (**Figure 1**). Residues 65-80, 245-255 and 263-280 are mobile independently from the crystal packing when compared to the structure of TK<sub>HSV1</sub>:dT:ADP in space group  $I4_1$ <sup>21</sup> or TK<sub>HSV1</sub>:dT in space group  $C222_1$ <sup>19</sup>.

Four water molecules in the active center occupy positions of the hydrogen bond-forming polar groups of the substrate dT. The absence of dT that could have been introduced by the preparation from *E. coli* cells was checked by omitting the water molecules in the binding pocket and modeling dT molecules in both subunits and refining them. Even at occupancies as low as 0.3 no dT could be fitted as judged from the  $s_A$ -weighted ( $2F_o - F_c$ )-map and the ( $F_o - F_c$ )-difference maps.

### LID-domains of substrate-free TK<sub>HSV1</sub>

The two subunits are in different crystallographic environments and are therefore involved in different crystal contacts. The overall conformation of both subunits is similar to the one observed for TK in complex with the substrate. The LID domain (residues 219-226) of subunit A (LID<sub>A</sub>) is fixed by crystal contacts. This

conformational state can be described as "closed" in analogy to the nucleoside monophosphate kinases<sup>29</sup>. In contrast, the more solvent exposed LID domain of subunit B (LID<sub>B</sub>) shows high mobility reflected by poorly defined electron density.



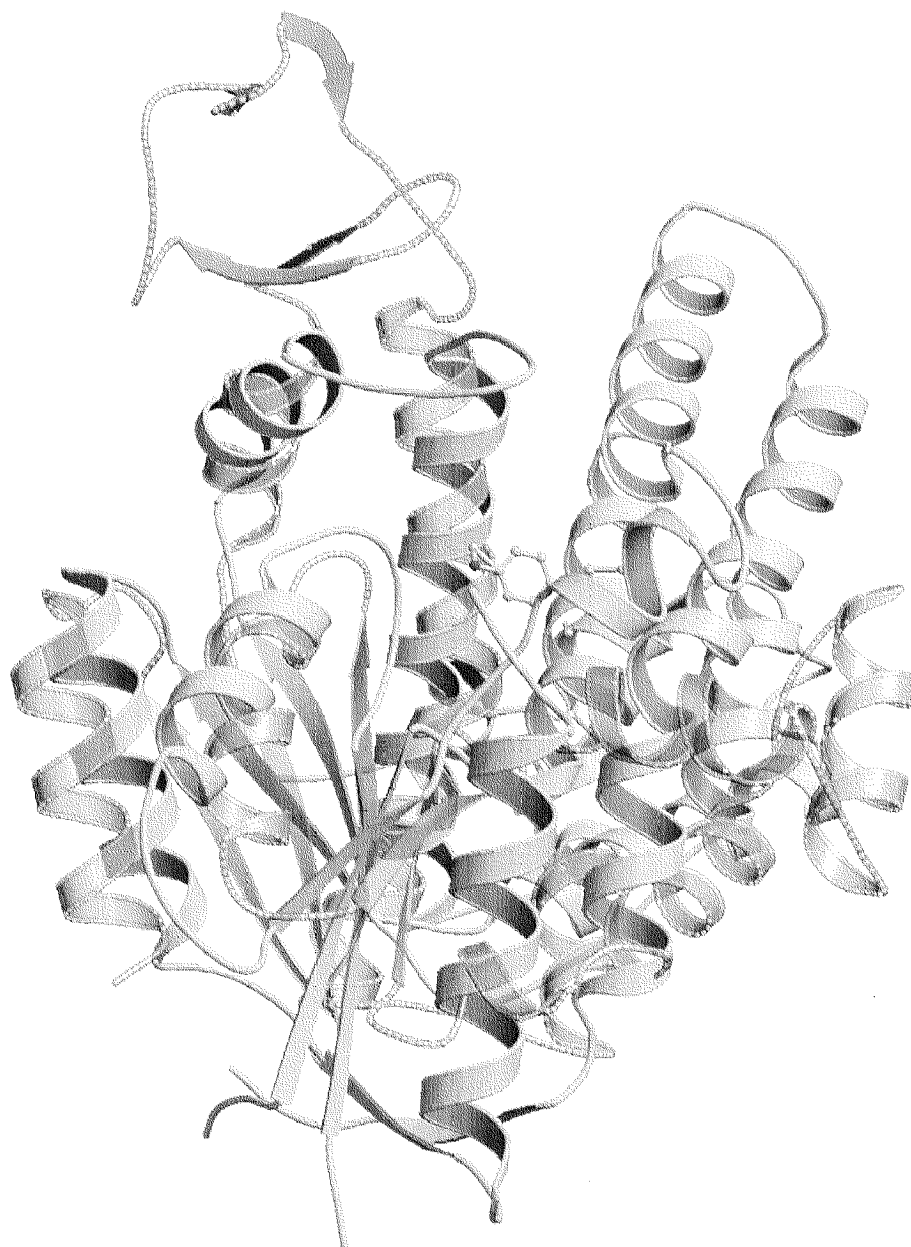
**Figure 1:** B-Factor-plots of subunits A (a) and B (b) of substrate-free TK<sub>HSV1</sub> (straight line), TK<sub>HSV1</sub>:dT **19** (dotted line) in spacegroup C222<sub>1</sub> and TK<sub>HSV1</sub>:dT:ADP **21** (dashed line) in spacegroup I4<sub>1</sub>.

The comparison of substrate-free  $\text{TK}_{\text{HSV1}}$  with the substrate complex  $\text{TK}_{\text{HSV1}}:\text{dT}$  solved in the same space group<sup>19</sup> and the structure of  $\text{TK}_{\text{HSV1}}:\text{dT}:\text{ADP}^{21}$  in spacegroup  $I4_1$  show rms- $\text{C}\alpha$ -deviations that are within the coordinate error. The lack of substrate molecules affects the structure in several regions mainly in subunit B as can be visualized in the  $B$ -factor plot (Figure 1). The mobilities of residues in the structures of Substrate-free  $\text{TK}_{\text{HSV1}}$  and  $\text{TK}_{\text{HSV1}}:\text{dT}^{19}$  in space group  $C222_1$  and the structure of  $\text{TK}_{\text{HSV1}}:\text{dT}:\text{ADP}^{21}$  are compared. Residues forming the substrate binding pocket (Tyr101, Tyr172, Met128, Gln125) have significantly higher mobility in the ligand-free structure but keep their positions in the range of the coordinate errors. By binding of dT to the protein  $\text{LID}_B$  is fixed whereas residues 220-225 are too mobile to be observed in substrate-free  $\text{TK}_{\text{HSV1}}$ .  $\text{LID}_B$  is therefore in the "open" state and most likely represents the *in vivo* state of the whole dimer.

### Superposition TK-AK (Figure 2)

Important domain movement are further suggested by the structural homology to  $\text{AK}_{\text{eco}}$  and thermodynamic studies<sup>30</sup>. In all presented structures a sulfate ion occurring from the crystallization solution is bound to the P-loop. This is in agreement with thermodynamic data of phosphate buffer experiments measured by ITC, which clearly indicate the influence of phosphate ions to the binding properties of the enzyme ( $\text{KD}(\text{dT}) = 5.22 \pm 0.04 \mu\text{M}$  vs.  $\text{KD}(\text{dT}+\text{PO}_4) = 0.59 \pm 0.04 \mu\text{M}$ ) and let assume a bigger 3D-rearrangement of the enzyme in analogy to  $\text{AK}_{\text{eco}}$ <sup>30</sup>.

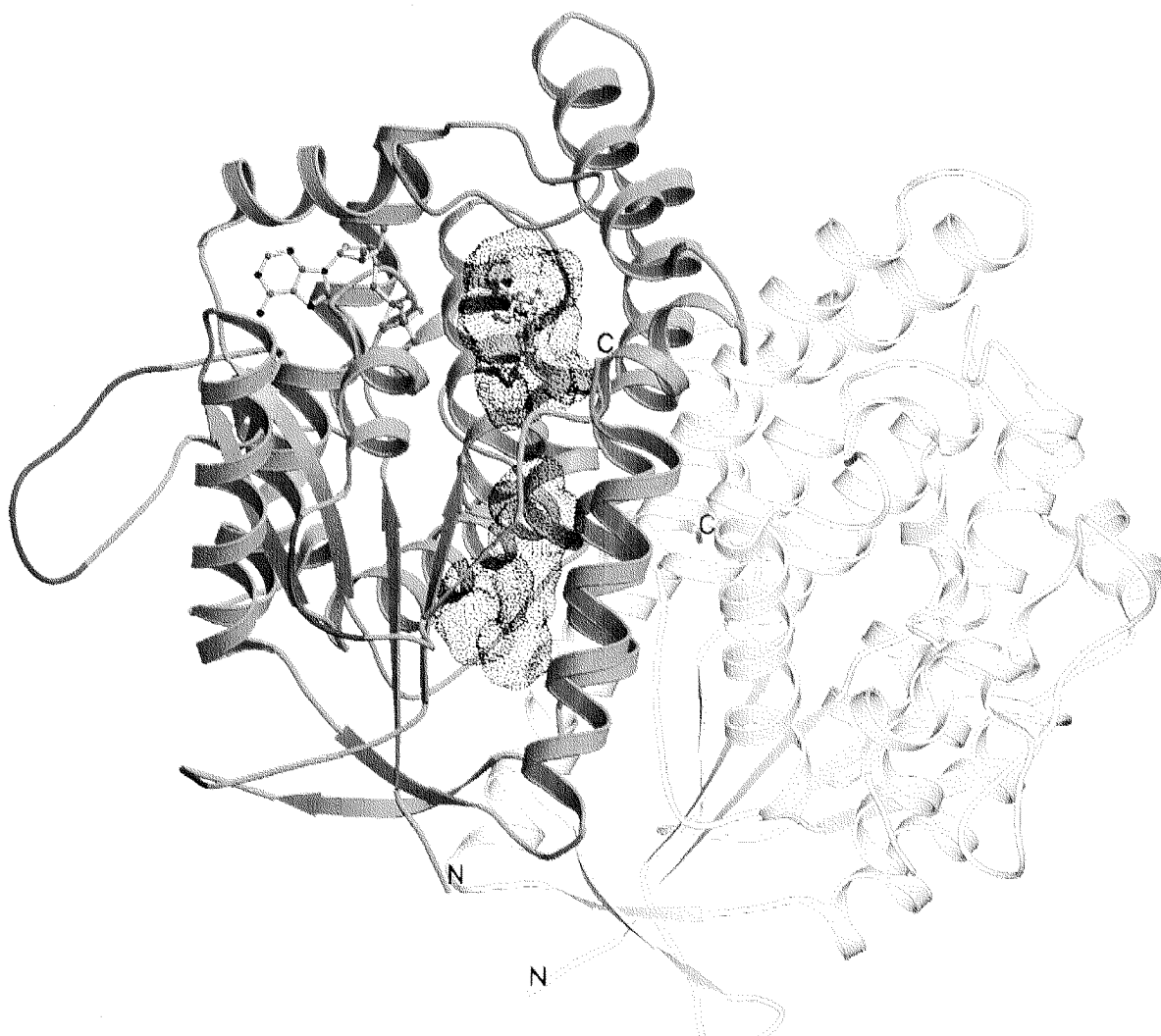
Under these premise the logical consequence is that the presence of sulfate does not allow to see the domain movement. The importance of the high non physiological concentrations of sulfate must be considered, since experiments where the displacement of sulfate by ATP was attempted failed.



**Figure 2** Superposition of protomer B of substrate-free TK<sub>HSV1</sub> (yellow) and AK<sub>eco</sub> (grey) fitted by P-loop and the first core  $\beta$ -strand. The LID<sub>B</sub> corresponds to the flexible domain of the AK<sub>eco</sub>.

### The role of water for binding of the substrate

Two chains of water molecules in the regions of residues 125, 168 and 169 at one hand and 132 and 163 at the other could indicate how the water molecules could move during the substrate binding process. The solvent accessible surface area in these regions analyzed using the MS program<sup>31</sup> with a probe radius of 1.4 Å shows a putative water channel that is restricted only by Tyr132 (Figure 3).



**Figure 3.** Ribbon diagram of  $TK_{HSV1}(apo)$  with dT and ADP molecules from the  $TK_{HSV1}:dT:ADP$ -structure (Wild *et al.*, 1997) for illustration. Subunit A is shown in orange with  $Lid_A$  colored in green. To clarify the connections unmodeled residues are depicted in blue. Subunit B is shown in grey. The putative water channel (Connolly surface area) is depicted in blue.

It is known from the complex structure of  $TK_{HSV1}$  with BVDU (5-bomovinyloxy-deoxyuridine) that this tyrosine ring can be turned away<sup>19</sup> so that a continuous channel linking the buried active site with the solvent could form. The arrangement of the water within this channel with its hydrogen-bond network, its relatively low B-factors ( $25 \text{ \AA}^2$ ) is similar to this of other water channels reported in the literature<sup>32</sup>. Furthermore, the presence of Tyr172 that interrupts the channels and must be dynamically moved to permit the displacement of individual water molecules make this channels resembling the ones described for porcine pancreatic elastase<sup>33</sup> and human leucocyte elastase<sup>34,35</sup>. The crystallographic results suggest that the channel is a sort of "rear exit" that opens up to permit trapped water to escape during substrate (dT) binding and water to flow back during product (dTMP)

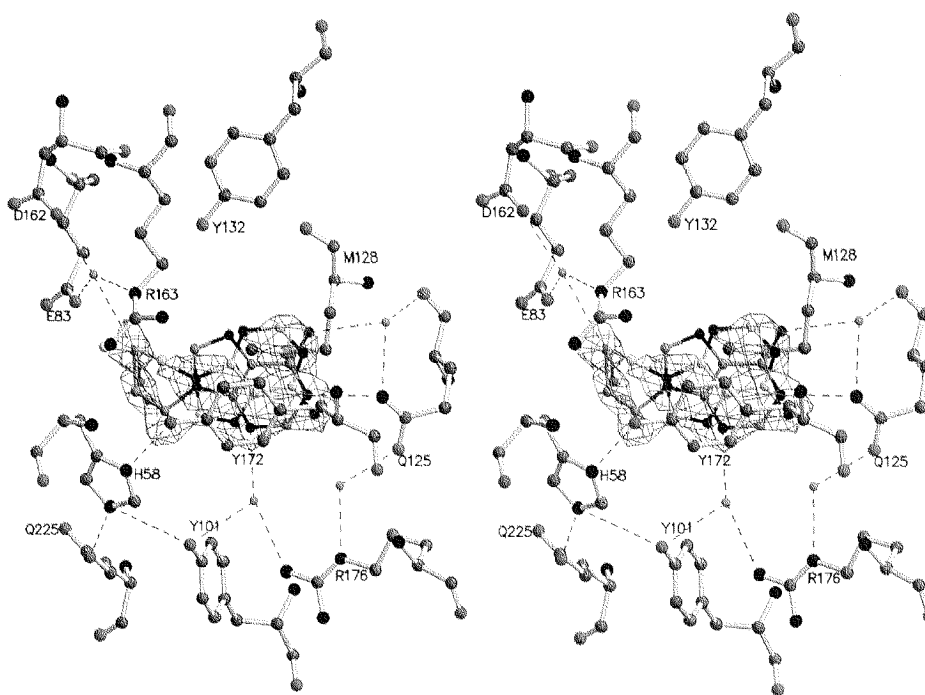
release. Thus, the water molecules in the channel undergo the so called ligand-induced hydrodynamic displacement<sup>32</sup>.

The role of water is also supported by the fact that dT has a relatively low affinity to TK<sub>HSV1</sub> ( $K_M=0.2 \mu\text{M}$ , calculated free energy of binding using  $\Delta G=-RT\ln K_i$  is  $-9.14 \text{ kcal/mol}$ ) despite the strong hydrogen-bond network with 5 direct and 2 water-mediated H-bonds<sup>20</sup> and the strength of the sandwich like complex<sup>36</sup>. Recently a survey of experimental data has been reported showing that the maximal free energy of binding contribution per non-hydrogen atom is  $\cong -1.5 \text{ kcal/mol}$ <sup>37</sup>. Following this empirical calculation dT with 17 non-hydrogens would have a theoretical maximal free energy of binding of  $\cong -25 \text{ kcal/mol}$  that is high above the value calculated from the  $K_M$  where  $K_M$  can be set equal  $K_i$ <sup>38</sup>. Comparing dT with other systems of the same size reported by Kuntz and co-workers we can see that the affinity of dT for TK<sub>HSV1</sub> is more than order of magnitude smaller than this of the ligand of the same size with its correspondent target. The reduction of free energy of binding may be explained by strong entropic effect part of which is due to desolvation including displacement of water from the binding site.

### The structure of TK<sub>HSV1</sub>:9-HPA

The complex structure of TK<sub>HSV1</sub>:9-HPA is the first structure of an adenine derivative. It was first refined applying strict NCS-restraints to the dimeric model. After several cycles the restraints were removed. As for TK<sub>HSV1</sub>(apo) several parts of the protein could not be modeled. These are in subunit A residues 64-75, 149-152 and 266-278, in subunit B residues 148-150, 221-224 and 265-273. The N-terminal 45 residues are missing in both subunits. The  $(2F_o-F_c)$ -map of the resulting model indicated clear densities for the substrate molecule 9-HPA in both subunits. In subunit A two different positions for 9-HPA were possible (Figure 4). The alternate binding corresponds to the poor binding affinity of 9-HPA as compared to the guanine derivative aciclovir (ACV) ( $K_{i(9\text{-HPA})}= 5.3 \text{ mM}$  vs.  $K_{i(\text{ACV})}= 0.2 \text{ mM}$ ). For simplicity we denoted the two positions I and II where the 6-amino group points towards Y132 and R176, respectively.



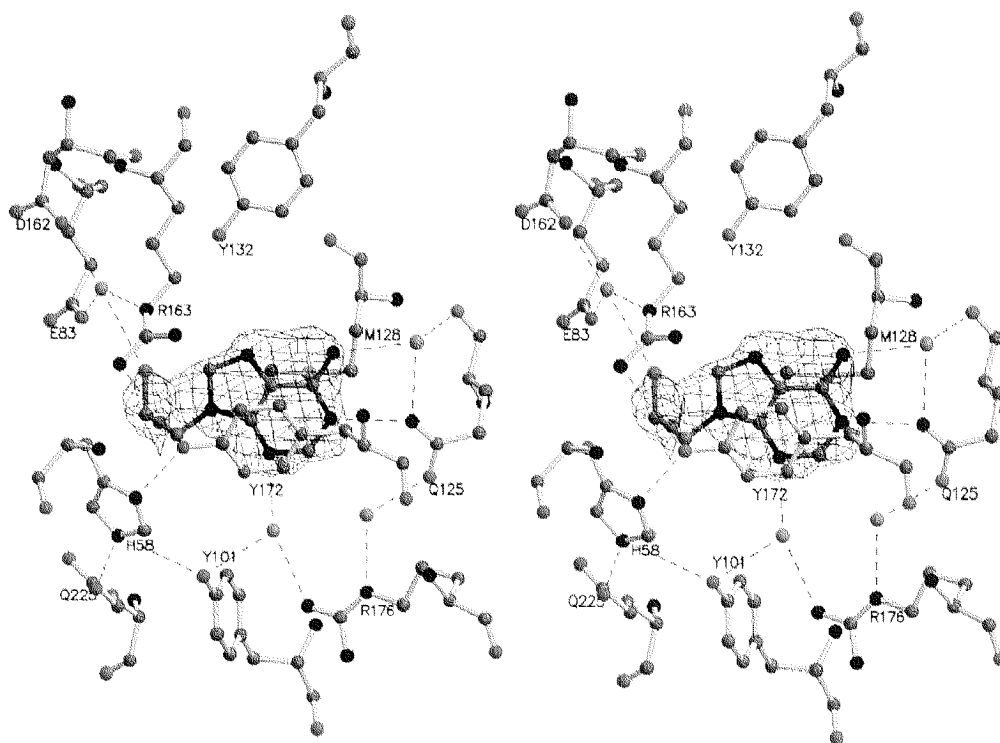


**Figure 4.** Active site ( $F_o-F_c$ )-electron density map (green) at 1.9 Å and  $2.5 \sigma$  in protomer A before introduction of water molecules, sulfate ions and 9-HPA showing both plausible compound orientations of the (S)-isomer. The substrate molecule in position I is shown in pink, position II in yellow. Water molecules and H-bonds are only drawn for position I.

The (S)-isomer in position I was refined first (Figure 5a), 221 water molecules and 2 sulfate ions in both *P*-loops were added. The complex of 9-HPA with TK<sub>HSV1</sub> was finally refined to 1.9 Å resolution with a R-factor of 21.7 % ( $R_{\text{free}}=26.9$  %) (Table 1). Again, only Arg163 was in a disallowed region. Adenine was stacked between Met128 and Tyr172 as observed in other TK<sub>HSV1</sub> complex structures. Both the purine ring and the  $C_\alpha$  of the side chain fit well into the electron density at a contour level of 1.8. In subunit A the substrate molecule was refined with an occupancy of 0.8 resulting in an *B*-factor around  $40 \text{ \AA}^2$  which is in the same range as its environment. The 2'-OH group is oriented towards Glu83 allowing the phosphorylation like the 5'OH of thymidine in TK<sub>HSV1</sub>:dT. Two water mediated hydrogen bonds with adenine are formed. One water molecule is fixed by the main-chain carbonyl of Gln125 and interacts with the 6-amino-group. The other water molecule was located between the guanidinium group of Arg176, the hydroxy group of Tyr101 and N3 of adenine. The amide group of Gln125 forms a hydrogen bond to N5 of adenine. One more water mediated hydrogen bond is formed between OH of the hydroxy-propyl part and the carboxy group of Glu83 at one hand and the  $N_\epsilon$  of Arg163 (Figure 5). The methyl-moiety of 9-(S)-HPA forms

hydrophobic interactions with Ile97 and the hydrophobic part of Arg222. For the (R)-isomer such interactions not possible.

The complex structure was then refined for position II with the same occupancy. The  $(2F_o - F_c)$ -map is also well defined and has the shape of 9-HPA. The 6-amino group looks towards Arg176. All water molecules are arranged at the same locations and interact with the same residues as in position I. Residue Gln125 is rotated  $180^\circ$  with the carbonyl as H-bond acceptor for the 6-amino-group of 9-HPA. Its N1 interacts with another water molecule that is fixed by Gln125-NH<sub>2</sub>. The water molecules that are attached to Arg176 interact with the 6-amino group and N7 of 9-HPA, respectively. The hydroxy-group of the (S)-isomer cannot form a hydrogen bond as observed in position I but the (R)-isomer can. Still the H-bond pattern is not as favorable as for position I, judged from analyzing the H-bond acceptor-donor relations of the adenine moiety. Though position I with the (S)-stereoisomer seems to be the favored binding mode of *rac*-9-HPA.



**Figure 5.** Stereo view of the refined 9-(S)-HPA structure in subunit A, position I. The final  $(2F_o - F_c)$ -maps are shown at  $1.2 \sigma$  contour level, water molecules are shown in green.

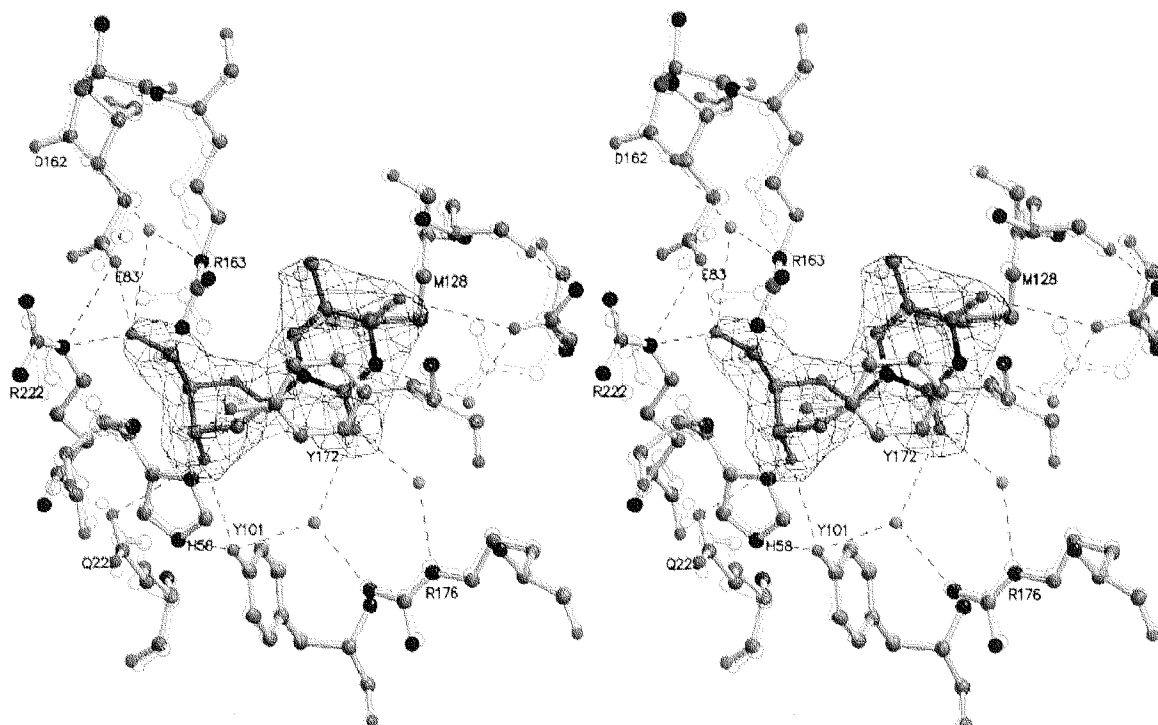
For subunit B no detailed analysis of the binding mode could be made because of the poor electron density map in the active center. The substrate molecule was finally fitted into the  $(2F_o - F_c)$ -map in position I with an occupancy of 0.4 resulting in  $B$ -factors around  $40 \text{ \AA}^2$  whereas in position II the  $B$ -factors arose up to  $90 \text{ \AA}^2$ . This seems to make plausible that position I is in favor for subunit B. When comparing the binding mode of dT and 9-(S)-HPA in position I it is observed that the stacking interaction of both rings is the same, but the strong Watson-Crick like interaction of Gln125 with dT is not established with 9-HPA. Instead only one direct interaction and one water mediated hydrogen bond via the nitrogen of the carboxamide group of Gln125 is formed. The positions of the water molecules bound to Arg176 are identical. They interact either with O2 of dT or with N1 and N3 of the purin ring of 9-HPA. The position of the hydroxy-group of 9-HPA corresponds to the 5'-OH of dT, interacting with Glu83. Additionally, dT forms a hydrogen bond to Tyr101 via its 3'-OH group. There is no equivalent for this interaction for 9-HPA. Altogether, this might explain the poor binding at one hand and the broad substrate acceptance of TK<sub>HSV1</sub> at the other.

### **The structure of TK<sub>HSV1-Q125N</sub>:dT**

The structure of the mutant TK<sub>HSV1</sub> in complex with dT was solved to  $2.5 \text{ \AA}$  resolution at 100 K. Crystals belong to space group  $C222_1$  with two subunits of 39 kDa per asymmetric unit. The smaller size compared to TK<sub>HSV1</sub>(apo) and TK<sub>HSV1</sub>:9-HPA results from the purification protocol using thrombin instead of the PreScission protease, cleaving off the first 33 N-terminal residues. The R-factor of the refined model is 21.2 % ( $R_{\text{free}} = 27.3 \%$ ) (Table 1), 111 water molecules, 2 sulfate ions in both P-loops and two dT molecules were added to the model. Only Arg163 of both subunits is in a disallowed region in the Ramachandran diagram as observed in TK<sub>HSV1</sub>(apo) and in TK<sub>HSV1</sub>:9-HPA. Missing parts in the model of TK<sub>HSV1</sub>(Q125N):dT are besides the N-terminal part 1-45 in both subunits residues 70-74, 150-152 and 265-279 in subunit A and residues 150-152, 220-226 and 268-273 in subunit B.

The nucleoside binding site in the active center has been described by Wild *et al.* (1997). The binding of dT and other substrate analogs is mainly formed by two

different types of interactions. The thymine ring is stacked between Tyr172 and Met128. The hydroxyl group of Tyr172 is fixed by hydrogen bonding to Arg163-NH2 and His58-ND1. Watson-Crick-like hydrogen bonding of atoms N3 and O4 $\alpha$  of the thymine ring with the carboxamide group of Gln125 positions the substrate<sup>19,21</sup>. To clarify the contribution of Gln125 in substrate binding and the hydrogen bonding pattern of thymidine in the active site mutant Q125N in complex with dT was analyzed. The structure of TK<sub>HSV1</sub>(Q125N):dT now revealed that the overall structure is the same as for the wild-type, no domain movements or rearrangements of residues could be observed. All differences are within the limits of error. The superposition of the active centers of both mutant and wild-type shows no reorientation or shift of the substrate in the binding site (Figure 6).



**Figure 6.** Superposition of the active sites of both wildtype (grey) and mutant Q125N. The final ( $2F_o - F_c$ )-map is shown at 1.2  $\sigma$  contour level. Water molecules are depicted in green. No reorientation of dT in the binding site could be observed.

The thymine ring is fixed by the geometric complementarity of Met128 and Tyr172. The distance of the carboxamide group of Asn125 to the O4 $\alpha$  of the thymine ring is increased by 1.6 Å to 3.9 Å which is too far for a strong hydrogen bond. Thus, the two major hydrogen bonds present in the wild-type structure are replaced by the insertion of a water molecule between the carboxamide group of Asn125 and the water next to thymidine and Arg176.

The side chain of Gln125 is observed to be rotated 180° when TK<sub>HSV1</sub> structures with bound thymidine and bound aciclovir are compared. The active site mutant Q125N showed a 50fold decrease in binding affinity for dT ( $K_{M,Q125N} = 10 \mu\text{M}$ ;  $K_{M,WT} = 0.2 \mu\text{M}$ ;  $k_{cat,Q125N} = 0.437 \text{ s}^{-1}$ ;  $k_{cat,WT} = 0.437 \text{ s}^{-1}$ ) compared to the wild type and only a three times decrease for ACV ( $K_{i,Q125N} = 555 \mu\text{M}$ ;  $K_{i,WT} = 170 \mu\text{M}$ ;  $k_{cat,Q125N} = 0.062 \text{ s}^{-1}$ ;  $k_{cat,WT} = 0.125 \text{ s}^{-1}$ ), respectively<sup>15</sup>. The enzymatic activity for dT was retained while for ACV it was reduced two-times compared with wild-type TK<sub>HSV1</sub>. The comparison of the specificity terms calculated using the equation  $(k_{cat}/K_M[\text{dT}]) / (k_{cat}/K_M[\text{prodrug}]$ )<sup>39</sup> indicates clearly a kinetic advantage for the mutant. The decrease of the value of the specificity terms from 2971 for wild-type TK<sub>HSV1</sub> to 391 for TK<sub>HSV1</sub>(Q125N) indicates that this mutant has an increased specificity for ACV. This is compatible with the results of Kokoris *et al.* (1999), which showed that an increased specificity for the prodrug corresponds to an enhancement in the efficacy of the enzyme-prodrug gene therapy approach *in vivo*. Therefore the active site mutant was not only taken to clarify the contribution to the substrate binding mechanism at structural level but could also be interesting for *in vivo* approaches. The structural information will be helpful in the development of new prodrugs.

## CONCLUSIONS

In this study we present the structures of the TK<sub>HSV1</sub> with ligand free active site and in complex with *rac*-9-hydroxypropyl-adenine and the active site mutant Q125N in complex with thymidine. We gained insights into the role of water during catalysis and achieved a better understanding of the broad substrate acceptance of the enzyme. The TK<sub>HSV1</sub> substrate acceptance is even larger than stated up to now shown by the complex structure with the adenine derivative 9-HPA. Although 9-HPA cannot be used directly in therapy because of its poor binding to TK<sub>HSV1</sub>, it represents a lead compound for gene therapy, because it is known that a phosphorylated analog (PMEA) is a potent inhibitor of cellular DNA-polymerase<sup>41</sup>. Mutant Q125N shows an increased specificity for aciclovir than dT compared to the wild type. This increase could be exploited in gene therapy, using TK<sub>HSV1</sub>(Q125N) in combination with ACV or other substrates..

## ACKNOWLEDGEMENT

We thank the team of the EMBL-outstation Hamburg for help with synchrotron data collection. We also thank Dr. M. Mintas for the gift of compound 9-HPA. Pierre Schelling was supported by the Stipendienfonds der Basler Chemischen Industrie.

## REFERENCES

1. Okazaki, R. & Kornberg, A. Deoxythymidine Kinase of Escherichia coli. *Journal of Biological Chemistry* **239**, 275-284 (1964).
2. Keller, P.M. *et al.* Enzymatic phosphorylation of acyclic nucleoside analogs and correlations with antiherpetic activities. *Biochem Pharmacol* **30**, 3071-7 (1981).

3. Elion, G.B. *et al.* Selectivity of action of an antiherpetic agent, 9-(2-hydroxyethoxymethyl) guanine. *Proc Natl Acad Sci U S A* **74**, 5716-20 (1977).
4. Fyfe, J.A., Keller, P.M., Furman, P.A., Miller, R.L. & Elion, G.B. Thymidine kinase from herpes simplex virus phosphorylates the new antiviral compound, 9-(2-hydroxyethoxymethyl)guanine. *J Biol Chem* **253**, 8721-7 (1978).
5. Tiberghien, P. "Suicide" gene for the control of graft-versus-host disease. *Curr Opin Hematol* **5**, 478-82 (1998).
6. Panis, Y. & Houssin, D. [Gene therapy. A new prospect in the treatment of liver tumors (editorial)]. *Presse Med* **24**, 1681-3 (1995).
7. Sturtz, F.G. *et al.* Parameters influencing the efficiency of the thymidine kinase/ganciclovir strategy in human glioblastoma cell lines. *Stereotact Funct Neurosurg* **68**, 252-7 (1997).
8. Nagy, H. *et al.* Are hepatomas a good target for suicide gene therapy? An experimental study in rats using retroviral-mediated transfer of thymidine kinase gene. *Surgery* **123**, 19-24 (1998).
9. Klatzmann, D. Gene therapy for metastatic malignant melanoma: evaluation of tolerance to intratumoral injection of cells producing recombinant retroviruses carrying the herpes simplex virus type 1 thymidine kinase gene, to be followed by ganciclovir administration. *Hum Gene Ther* **7**, 255-67 (1996).
10. Black, M.E., Newcomb, T.G., Wilson, H.M. & Loeb, L.A. Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy. *Proc Natl Acad Sci U S A* **93**, 3525-9 (1996).
11. Encell, L.P., Landis, D.M. & Loeb, L.A. Improving enzymes for cancer gene therapy. *Nat Biotechnol* **17**, 143-7 (1999).
12. Kokoris, M., Sabo, P., Adman, E. & Black, M. Enhancement of tumor ablation by a selected HSV-1 thymidine kinase mutant. *Gene Ther* **6**, 1415-1426 (1999).

13. Wildner, O. *et al.* Adenoviral vectors capable of replication improve the efficacy of HSVtk/GCV suicide gene therapy of cancer. *Gene Ther* **6**, 57-62 (1999).
14. Beltinger, C. *et al.* Herpes simplex virus thymidine kinase/ganciclovir-induced apoptosis involves ligand-independent death receptor aggregation and activation of caspases. *Proc Natl Acad Sci U S A* **96**, 8699-704 (1999).
15. Kussmann-Gerber, S., Kuonen, O., Folkers, G., Pilger, B.D. & Scapozza, L. Drug resistance of herpes simplex virus type 1--structural considerations at the molecular level of the thymidine kinase. *Eur J Biochem* **255**, 472-81 (1998).
16. Pilger, B. *et al.* Synthesis and structure-function studies of novel classes of herpesviral thymidine kinase ligands. *in press* (1999).
17. Bennett, M.S. *et al.* Structure to 1.9 Å resolution of a complex with herpes simplex virus type-1 thymidine kinase of a novel, non-substrate inhibitor: X-ray crystallographic comparison with binding of aciclovir. *FEBS Lett* **443**, 121-5 (1999).
18. Brown, D.G. *et al.* Crystal structures of the thymidine kinase from herpes simplex virus type-1 in complex with deoxythymidine and ganciclovir. *Nat Struct Biol* **2**, 876-81 (1995).
19. Champness, J.N. *et al.* Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands. *Proteins* **32**, 350-61 (1998).
20. Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett* **368**, 289-92 (1995).
21. Wild, K., Bohner, T., Folkers, G. & Schulz, G.E. The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue. *Protein Sci* **6**, 2097-106 (1997).
22. Fetzer, J., Michael, M., Bohner, T., Hofbauer, R. & Folkers, G. A fast method for obtaining highly pure recombinant herpes simplex virus type 1 thymidine kinase. *Protein Expr Purif* **5**, 432-41 (1994).



23. CCP4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763 (1994).
24. Kabsch, W. Evaluation of single-crystal X-ray diffraction data from a position sensitive detector. *J. Appl. Cryst.* **21**, 916-924 (1988).
25. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47**, 110-9 (1991).
26. Kraulis, P.J. MOLSCRIPT - a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* **24**, 946-950 (1991).
27. Merritt, E.A. & Bacon, D.J. Raster3D: Photorealistic Molecular Graphics. *Methods in Enzymology* **277**, 505-524. (1997).
28. Dreusicke, D. & Schulz, G.E. The glycine-rich loop of adenylate kinase forms a giant anion hole. *FEBS Lett* **208**, 301-4 (1986).
29. Vonrhein, C., Schlauderer, G.J. & Schulz, G.E. Movie of the structural changes during a catalytic cycle of nucleoside monophosphate kinases. *Structure* **3**, 483-90 (1995).
30. Perozzo, R., Folkers, G. & Scapozza, L. manuscript in preparation. (1999).
31. Connolly, M.L. Analytical molecular surface calculation. *Journal Appl. Crystallogr.* **16**, 548-558 (1983).
32. Meyer, E. Internal water molecules and H-bonding in biological macromolecules: a review of structural features with functional implications. *Protein Sci* **1**, 1543-62 (1992).
33. Meyer, E., Cole, G., Radhakrishnan, R. & Epp, O. Structure of native porcine pancreatic elastase at 1.65 Å resolutions. *Acta Crystallogr B* **44**, 26-38 (1988).
34. Bode, W. *et al.* X-ray crystal structure of the complex of human leukocyte elastase (PMN elastase) and the third domain of the turkey ovomucoid inhibitor. *Embo J* **5**, 2453-8 (1986).

35. Bode, W., Meyer, E., Jr. & Powers, J.C. Human leukocyte and porcine pancreatic elastase: X-ray crystal structures, mechanism, substrate specificity, and mechanism-based inhibitors. *Biochemistry* **28**, 1951-63 (1989).
36. Alber, F., Kuonen, O., Scapozza, L., Folkers, G. & Carloni, P. Density functional studies on herpes simplex virus type 1 thymidine kinase-substrate interactions: the role of Tyr-172 and Met-128 in thymine fixation. *Proteins* **31**, 453-9 (1998).
37. Kuntz, I.D., Chen, K., Sharp, K.A. & Kollman, P.A. The maximal affinity of ligands. *Proc Natl Acad Sci U S A* **96**, 9997-10002 (1999).
38. Kussmann-Gerber, S. *et al.* Interaction of the recombinant herpes simplex virus type 1 thymidine kinase with thymidine and aciclovir: a kinetic study. *Nucleosides Nucleotides* **18**, 311-30 (1999).
39. Christians, F.C., Scapozza, L., Cramer, A., Folkers, G. & Stemmer, W.P. Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling. *Nat Biotechnol* **17**, 259-64 (1999).
40. Pilger, B. *et al.* Substrate Diversity of Herpes Simplex Virus Thymidine Kinase - Impact of the kinematics of the enzyme. *J Biol Chem* **274**, 31967-31973 (1999).
41. Hatse, S., De Clercq, E. & Balzarini, J. Impact of 9-(2-phosphonylmethoxyethyl)adenine on (deoxy)ribonucleotide metabolism and nucleic acid synthesis in tumor cells. *FEBS Lett* **445**, 92-7 (1999).
42. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291 (1993).

## CHAPTER 5

### **Kinetics and high resolution structure of HSV1 TK and the engineered Y101F mutant in complex with an antiviral drug with conformationally restricted sugar ring pucker**

Andrea Prota<sup>†</sup>, Joachim Vogt<sup>#†</sup>, Beatrice D. Pilger<sup>\*§</sup>, Remo Perozzo<sup>\*</sup>,  
Christine Wurth<sup>\*</sup>, Victor E. Marquez<sup>++</sup>, Pamela Russ<sup>++</sup>, Georg E. Schulz<sup>#</sup>,  
Gerd Folkers<sup>\*</sup> and Leonardo Scapozza<sup>\*</sup>

<sup>\*</sup>Department of Applied Biosciences, Swiss Federal institute of technology (ETH)  
Winterthurerstr. 190, CH-8057 Zürich, Switzerland

<sup>#</sup>Institut für Organische Chemie und Biochemie, Albert-Ludwigs-Universität,  
Albertstraße 21, D-79104 Freiburg im Breisgau, Germany

<sup>++</sup>Laboratory of Medicinal Chemistry, Division of Basic Sciences,  
National Cancer Institute, National Institute of Health, Bethesda, Maryland 20892

<sup>†</sup>These authors contributed equally to this work.

<sup>§</sup>Present address: Harvard Medical School, Dept. of Biological Chemistry and  
Molecular Pharmacology, 250 Longwood Ave , Boston, MA 02115

Submitted to BIOCHEMISTRY

The coordinate data-sets have been submitted for deposition to the Protein Data Bank at Brookhaven, MA.

Running title: Crystal structure of wild type and engineered mutant HSV1 TK in complex with (N)-methanocarpa-thymidine

**Keywords:** TK, antiviral drug, prodrug-enzyme gene therapy, X-ray structure, sugar ring pucker

**Abbreviations:**

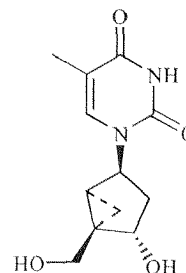
TK, thymidine kinase

HSV1, *Herpes simplex virus* type 1;

dT, thymidine

dTMP, thymidine monophosphate

N-MCT, 2'-*exo*-methanocarpa-thymidine; (1*R*,2*S*,4*S*,5*S*)-1-(Hydroxymethyl)-2-hydroxy-4-(5-methyl-2,4(1*H*,3*H*)-dioxypyrimidin-1-yl)bicyclo[3.1.0]hexan:



HCMV, human cytomegalovirus

ACV, aciclovir

EBV, Epstein-Barr virus

CPE, cytopathogenic effect

## ABSTRACT

Conformational preferences of nucleoside analogs play a role in their antiviral activity. Here we describe the kinetic and crystallographic analysis of (N)-methanocarpa-thymidine ((N)-MCT), a potent antiviral compound, with wild-type *Herpes simplex* virus type 1 thymidine kinase (HSV1 TK) and with a mutant kinase (Y101F) designed for gene therapy. The kinetic study reveals a twenty fold increase in  $K_M$  of thymidine for the Y101F compared to the wild-type HSV1 TK. Activity studies indicate that (N)-MCT is a substrate of both wild-type and mutant HSV1 TK. Inhibition experiments show that (N)-MCT binds to HSV1 TK and Y101F HSV1 TK with a  $K_i$  value of 11.3  $\mu$ M and 51.4  $\mu$ M, respectively. The affinity for human cytosolic TK is very low with a  $K_i$  value of 1.6 mM indicating that HSV1 TK is the selectivity filter for the antiviral activity of (N)-MCT. Furthermore, Y101F HSV1 TK shows an increased specificity for (N)-MCT when compared to wild-type HSV1 TK.

Crystal structures of wild-type and mutant HSV1 TK in complex with (N)-MCT have been determined to resolutions of 1.7 Å and 2.4 Å, respectively. The base of (N)-MCT binds like the base of thymidine and the conformationally restricted sugar-like bicyclo[3.1.0]hexan assumes a 2-*exo* envelope conformation (nucleoside numbering) that is more flat than the one observed for the free compound in the crystal and in solution. Furthermore, the hydrogen bonding pattern around the sugar-like moiety differs from that of the natural 2'-deoxy moiety and shows the importance of the rigid conformation of (N)-MCT with respect to hydrogen bonds. With these results (N)-MCT becomes a lead compound for the design of resistance-repellent drugs for antiviral therapy, and mutant Y101F in combination with (N)-MCT provides new insights into the possibility of tailor-made HSV1 TK for gene therapy.

## INTRODUCTION

*Herpes simplex* virus type 1 (HSV1) is a well-characterized widespread infectious agent in human populations<sup>1,2</sup>. HSV1 infections are associated with oral-facial infections (e.g. gingivostomatitis, pharyngitis, and recurrent herpes labialis), skin infections (e.g. eczema herpeticum, and erythema multiforme), and genital infections. HSV1 infection can have a severe outcome, including blindness and central nervous system damage<sup>3</sup>.

During the past decade, potent agents against herpes virus infections have been found. The action of these agents is based on the difference between viral thymidine kinase (TK) and human TK. Thymidine kinase (TK, EC 2.7.1.21) is a key enzyme in the pyrimidine salvage pathway that catalyzes the phosphorylation of thymidine (dT) to thymidine monophosphate (dTMP) in the presence of  $Mg^{2+}$  and ATP<sup>4</sup>. HSV1 TK exhibits a broad substrate diversity in contrast to the cellular isoenzyme, accepting pyrimidine as well as purine nucleoside analogues<sup>5-7</sup>. These compounds are selectively activated through phosphorylation by HSV1 TK to finally act in the triphosphorylated form as inhibitors of polymerases as well as DNA chain terminators<sup>6-9</sup>. The increasing clinical use of aciclovir (ACV) and ganciclovir has been associated with the emergence of drug-resistance, mainly linked to TK mutations decreasing the activation efficiency of the prodrug into the active drug and fewer linked to mutations of viral DNA polymerase<sup>10</sup>. The development of resistance against aciclovir and the other nucleoside analogues is not yet a problem in the treatment of immunocompetent individuals. However, in immunocompromised patients, ACV-resistant HSV strains pose increasing problems<sup>11</sup>. Several new prodrugs have been developed. All of them are ACV analogs with improved oral bioavailability. While the efficacy of valaciclovir (prodrug of ACV) and famciclovir (prodrug of penciclovir) in the treatment of *Herpes genitalis* and acute *Herpes zoster* has been well documented in large clinical trials<sup>11-15</sup> the problem of cross-resistance has not been solved. The compounds are chemically similar and they bind similarly to HSV1 TK<sup>16-18</sup>.

The search for resistance repellent prodrugs has led to several types of compounds. Monophosphate analogs<sup>19,20</sup> such as adefovir and cidofovir which were synthesized to circumvent the obligate HSV1 TK in vivo activation of nucleoside analogs have shown promising in vitro activity against HSV1 and HSV2. In vivo however, they were ten-fold less potent than ACV **21**. Another type of resistance repellent drugs is represented by conformationally rigid compounds that differ as little as possible from the natural substrate<sup>22-25</sup>.

Several rigid nucleosides containing bicyclo[3.1.0]hexane have been synthesized<sup>23,26-28</sup> and tested for antiviral activity against HSV1 and HSV2, HCMV and EBV. Very potent antiherpetic activity against HSV1 and HSV2 was discovered for (N)-MCT<sup>23</sup>.

The enzyme-prodrug gene therapy of cancer and TK-controlled graft-versus-host disease in allogeneic bone marrow transplantation<sup>29-31</sup> resembles classical antiviral therapy. In enzyme-prodrug gene therapy, the gene of HSV1 TK is introduced into tumor cells by retroviral or adenoviral vectors. Dividing cells that express HSV1 TK are then capable of converting nontoxic nucleoside analogs such as ganciclovir into their nucleoside triphosphates which inhibit cellular polymerases causing cell death and consequently tumor ablation<sup>32-36</sup>.

Clinical trials with ganciclovir showed a limited efficacy of this strategy, due to unfavorable side-effects such as immunosuppression that were caused by the large dosages required for tumor regression. Protein engineering aiming at the creation of drug-specific HSV1 TK-mutants with improved specificity towards the prodrug and concomitantly a decreased thymidine utilization is therefore a useful undertaking<sup>37-39</sup>. In this study we present the engineering of a HSV1 TK mutant dedicated to gene therapy, the kinetic characterization and the crystal structures of (N)-MCT in complex with wild-type and the mutant HSV1 TK, showing two new approaches for antiviral and suicide gene therapy.

## EXPERIMENTAL PROCEDURES

### Materials

(Methyl-1',2'-<sup>3</sup>H)thymidine (3 TBq/mmol) was obtained from Amersham. Nucleotides and AmpliTaq Gold<sup>TM</sup> polymerase were from Perkin Elmer. Restriction endonucleases, T4 DNA ligase and thrombin were from Promega. Reagents for enzyme assays were obtained from Sigma.

Strain DH5 $\alpha$  (deo<sup>R</sup>, endA1, recA1, rel A1, gyrA96, thi-1, hsdR17, supE44, lacZ  $\Delta$  M15 F'  $\lambda$ ) (Clontech) was used for the cloning steps of the mutant. *E.coli* strain 71/18 (F', lac<sup>R</sup>, lacZ $\Delta$ M15), which produces high levels of *lac* repressor to suppress uncontrolled expression from the inducible *ptac* promoter, was used for the cloning procedures of TK into pGEX-6P-2. Strain BL21 (ompT<sup>-</sup>, F<sup>-</sup>, hsdS (rB<sup>-</sup>, mB<sup>-</sup>), gal) (Pharmacia) served as host for the expression. The plasmids pGEX-2T, pGEX-6P-2 and the PreScission<sup>®</sup>-protease were purchased from Pharmacia. The plasmid pBR322-TK containing the TK gene of HSV1 strain F was a gift from S. McKnight. Linbro plates and cover slides were purchased from Hampton Research, the sitting drop bridges were from DROP.

### Protein engineering

The active site mutant Y101F was constructed using the polymerase chain reaction based on the three-primer-method<sup>40,41</sup>. The mutagenic primer (Y101F: 5'-GCG AAC ATC TTC ACC ACA CAA C-3') and the M13 universal primer (5'-GCT ATG ACC ATG TTA CG-3') were from Microsynth (Balgach, Switzerland). The DNA fragment containing the desired mutation was cloned into the expression vector pGEX-2T-TK by digestion with the restriction enzymes *Bam*HI and *Kpn*I followed by gel purification and ligation. This resulted in the expression vector named pGEX-2T-Y101FTK that encodes for the thrombin cleavable glutathione-S-transferase-Y101F-HSV1 TK fusion protein.



### Sequence verification

Competent *E. coli* DH5 $\alpha$  cells were transfected with the mutated pGEX-2T-TK DNA. After DNA isolation of several clones, both strands of the entire gene of the mutant were sequenced using the dye terminator method (ABI PRISM™ 310) to confirm that there were no additional mutations.

### Protein expression and purification

For the wild-type HSV1 TK, the expression vector pGEX-6P-2-TK coding for a PreScission™ Protease cleavable glutathione-S-transferase-fusion protein was used. The expression vector pGEX-2T-Y101FTK resulting from protein engineering and pGEX-2T-hTK1 were utilized for the Y101F HSV1 TK mutant and human cytosolic TK expression, respectively. Wild-type HSV1 TK, human TK1 as well as the Y101F HSV1 TK mutant were expressed in *E. coli* strain BL21 as glutathione-S-transferase-fusion proteins using a modified version of the previously published protocol<sup>42</sup>. For purification we used glutathione affinity chromatography, but refrained from an additional ATP-affinity chromatography step. The crude extract derived from 1 liter bacterial culture was applied five times to the glutathione sepharose column. The following three washing steps were carried out before proteolytic cleavage with buffer A (1M NaCl in 50 mM Tris pH 7.5, 10% glycerol and, 1 mM DTT, 0.1% Triton X-100), buffer B (250 mM KH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, 150 mM NaCl, 0.1% Triton X-100, pH 7.0) and buffer C (250 mM KH<sub>2</sub>PO<sub>4</sub>/Na<sub>2</sub>HPO<sub>4</sub>, 150 mM NaCl, 0.1% Triton X-100, 10 mM MgATP, pH 7.4). Buffer C allowed the ATP-dependent elution of a co-purified 70 kDa protein (dnaK)<sup>43</sup>. The column was then equilibrated with cleavage buffer without Triton X-100. For the mutant Y101F the thrombin cleavage was carried out at room temperature over two hours. For the wild-type HSV1 TK the cleavage with PreScission®-protease was performed overnight at 4°C. The eluted proteins were monitored by SDS-PAGE and showed a purity of higher than 98%. Both proteins were then directly used for kinetic studies and crystallization trials.

### Assessment of phosphorylation pattern

Substrate phosphorylation (dT, (N)-MCT) was monitored qualitatively by high performance liquid chromatography using a previously published protocol based on reverse-phase ion-pair chromatography<sup>44</sup>. Blank reactions without enzyme or without substrate were run concurrently to account for background ATP hydrolysis. The detection limit for phosphorylated substrates was 20 nmol<sup>44</sup>. The reactions were performed for wild-type HSV1 TK, Y101FHSV1 TK and hTK1 with an enzyme concentration of 70 µg/ml in the presence of 2 mM substrate, 5mM ATP and 5mM of Mg<sup>2+</sup>.

### Inhibition kinetics

Kinetic studies measuring the conversion of labeled thymidine to thymidine monophosphate in the presence of various concentrations of (N)-MCT were performed using the DEAE-cellulose method<sup>45,46</sup>. Reactions were carried out in 30 µl 50 mM Tris-HCl pH 7.2, 5 mM ATP, 5 mM MgCl<sub>2</sub>, 1.5 mg/ml BSA. The amount of enzyme and concentrations of (<sup>3</sup>H)-thymidine were chosen in consideration of Michaelis Menten conditions for initial velocity measurements. The concentration of the compound was varied according to their affinity towards thymidine kinase. The K<sub>i</sub> values were determined by non-linear fit of the raw data to the equation  $V = V_{max} * [dT] / (K_M * (1 + [compd.] / K_i) + [dT])$  for competitive inhibition using the Systat 5.02 software (Systat Inc. Evanston, IL USA). These values were measured based on at least three independent assays.

### Crystallization

The purified enzymes were concentrated by ultrafiltration using Centricon 30 concentrators to 25 mg/ml. Crystals of the mutant protein were obtained by adding (N)-MCT to the solution to a final concentration of 1 mM. Crystallization screenings were carried out at 23°C by sitting and hanging drop vapor diffusion by mixing equal volumes of reservoir and purified protein solution. Best crystals grew from a crystallization buffer containing 0.9 – 1.2 M Li<sub>2</sub>SO<sub>4</sub>, 1 mM DTT and 0.1 M HEPES at pH 8.0 to maximal sizes of 500 x 400 x 100 µm<sup>3</sup>. For wild-type TK in complex

with (N)-MCT, crystals of apo-TK were soaked 30 min into crystallization buffer containing 5 mM compound.

The complex TK:(N)-MCT diffracted X-rays isotropically to beyond 1.7 Å at a synchrotron source and the Y101F:(N)-MCT complex to 2.4 Å on a rotating anode. Both crystals belong to space group  $C222_1$  and contained two monomers of  $M_r$  41 kDa per asymmetric unit.

### **X-ray data collection**

Crystals were transferred into crystallization buffer containing 30 % (v/v) glycerol and flash-frozen to 100 K<sup>47</sup>. Data were collected on a MAR300 image plate detector using a 0.3 mm diameter collimated Cu  $K_{\alpha}$  radiation from a rotating anode generator (Rigaku, model RU2HC) at 45 kV and 120 mA. Further data collection was performed at 100K on EMBL beamline BW7B at DESY (Hamburg) with a MAR345 image plate. All diffraction data were processed using the programs MOSFLM, SCALA and TRUNCATE<sup>48</sup>.

### **Refinement**

The structure of TK:dTMP:ADP<sup>16</sup> without ligand and water molecules was used as the starting model. Refinement and difference Fourier electron density map calculations were done with program REFMAC<sup>48</sup> using the same protocol for both structures. The initial steps included non-crystallographic restraints of the dimeric model. Subsequently individual B-factors were applied and water molecules were added automatically using program ARPP<sup>48</sup>. Sulfate ions and substrate molecules were introduced manually. In the last refinement cycles the NCS-restraints were completely omitted. Modeling was done using the program O<sup>49</sup>. The programs MOLSCRIPT<sup>50</sup> and RASTER3D<sup>51</sup> were used for figure production.

**Table 1** Data Collection and Refinement Statistics

data set <sup>1</sup>	TK <sub>HSV1</sub> :(N)-MCT	TK <sub>HSV1</sub> -(Y101F):(N)-MCT
diffraction data		
X-ray source	BW7B	Cu K <sub>α</sub>
unit cell [Å]	a = 114.0	a = 114.0
	b = 117.7	b = 118.3
	c = 108.2	c = 108.6
resolution range [Å]	20 - 1.7	20 - 2.4
completeness [%]	99	97
multiplicity	4.1	3.4
unique Reflections	79549	27908
R <sub>sym</sub> <sup>2</sup> (last shell) [%]	5.0 (58)	8.4 (42)
I/σ (last shell)	10.1 (2.1)	11.3 (3.1)
refinement and final model		
R-factor (R <sub>free</sub> ) [%]	20.7 (25.4)	22.2 (27.9)
average B-factor [Å <sup>2</sup> ]	33	38
number protein non-H atoms	4771	4708
substrate atoms	36	36
water molecules	369	155
sulfate molecules	2	2

<sup>1</sup> Crystals were of space group *C222*<sub>1</sub>. All data were collected at 100 K.

<sup>2</sup>  $R_{\text{sym}} = \frac{\sum_h \sum_i |I_{hi} - \langle I_h \rangle|}{\sum_h \sum_i I_{hi}}$  where h stands for the unique reflections and i counts through symmetry-related reflections.

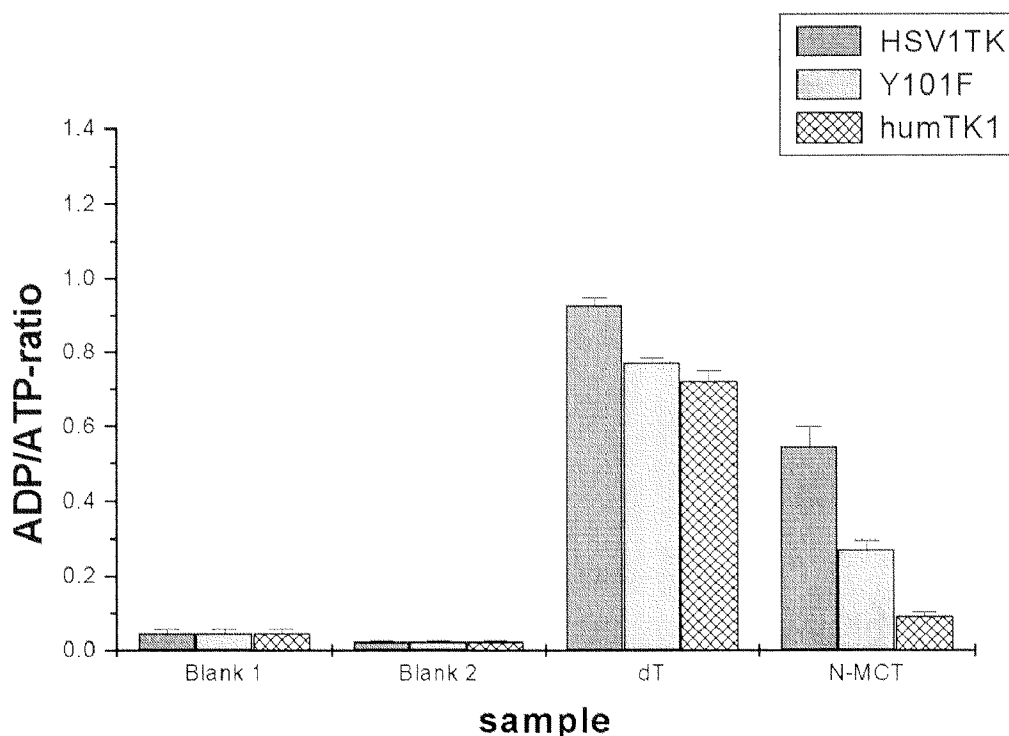
## RESULTS

### Protein engineering

The Y101F mutant was constructed by oligonucleotide-directed PCR mutagenesis. The mutated gene was sequenced entirely, confirming that only the designed mutation is present. Wild-type HSV1 TK and human TK1 as well as the Y101F HSV1 TK mutant were expressed in *E. coli* as glutathione-S-transferase-fusion proteins and purified to homogeneity and high purity by one-step affinity chromatography as described.

### Assessment of phosphorylation pattern

The phosphorylation was monitored by observing the formation of the corresponding phosphorylated compounds, ADP, the disappearance of (N)-MCT and the decrease of ATP.



**Figure 1: Histogram of the ratio between ADP and ATP resulting from the HPLC experiments with (N)-MCT using wild-type HSV1 TK, the mutant Y101F and human TK.** Control experiments were performed to monitor spontaneous hydrolysis of ATP. Blank 1 contained no substrate. Blank 2 contained no enzyme. The positive control was performed with dT at 2 mM concentration. The relative standard deviation is given as vertical bars. The experiments have been performed in triplicates using 2 mM (N)-MCT. All reaction mixes contained 5mM ATP and 5mM  $Mg^{2+}$ .

The ADP/ATP-ratios summarized in Figure 1 show an ADP formation linked to the reaction of (N)-MCT with ATP catalyzed by wild-type and mutant HSV1 TK that is clearly above simple ATP hydrolysis.

Less significant change in the ADP/ATP-ratio, compared to the negative controls, was detected for human cytosolic thymidine kinase with (N)-MCT. The ratios suggest that (N)-MCT is phosphorylated quite efficiently by wild-type HSV1 TK and by mutant Y101F, but not by the human cytosolic thymidine kinase.

### Kinetics

Competition studies were performed with wild-type HSV1 TK, mutant Y101F and human cytosolic TK1 using radioactively labeled thymidine and (N)-MCT as competitive “inhibitor”. Initial velocities ( $V_i$ ) were determined by monitoring the increase of radioactively labeled dTMP at various (N)-MCT concentrations. Table 2 summarizes the kinetic characterization of wild-type TK, Y101F and human cytosolic TK1 in presence or absence of (N)-MCT. The wild type TK showed a  $K_M$  of 0.2  $\mu\text{M}$  for dT and  $K_i$  of 11.4  $\mu\text{M}$  for (N)-MCT. Mutant Y101F showed a 20 fold increase in  $K_M$  for dT ( $K_M = 4.2 \mu\text{M}$ ), but only a five fold increase of  $K_i$  for the prodrug (N)-MCT ( $K_i = 51.4 \mu\text{M}$ ). The human cytosolic TK1 accepts (N)-MCT only with an affinity in the mM range ( $K_i = 1.6 \text{ mM}$ ) in agreement with the results obtained during the assessment of the phosphorylation pattern.

The comparison of the specificity terms, reflected in the ratio of  $K_i$  (N)-MCT to  $K_M$  thymidine, where in the case of HSV1 TK  $K_M$  equals  $K_i$ <sup>52</sup> shows clearly that wild-type HSV1 TK has a higher specificity (ratio = 57) for (N)-MCT compared to the human TK (ratio = 3197). Interestingly, mutant Y101F proves to be even more specific for the new prodrug (ratio = 12.3) than the wild-type enzyme.

**Table 2.** Inhibition constants ( $K_i$  values) of (N)-MCT in presence of dT with wild-type *Herpes simplex* virus type 1 thymidine kinase (HSV1 TK), the mutant Y101F or human cytosolic thymidine kinase (hTK1) compared to the  $K_M$  values of dT.

Enzyme	Thymidine (dT)		(N)-MCT	ratio
	$K_M$ [ $\mu\text{M}$ ]	$k_{\text{cat}}$ [ $\text{s}^{-1}$ ]	$K_i$ [ $\mu\text{M}$ ]	$\frac{K_i((\text{N})-\text{MCT})}{K_M(\text{dT})}$
HSV1 TK	$0.2 \pm 0.05$	$0.35 \pm 0.014$	$11.4 \pm 1.6$	57
Y101F	$2.5 \pm 0.48$	$0.19 \pm 0.05$	$51.5 \pm 0.1$	12.3
hTK1	$0.49 \pm 0.08$	$0.71 \pm 0.18$	$1566^* \pm 120$	3200

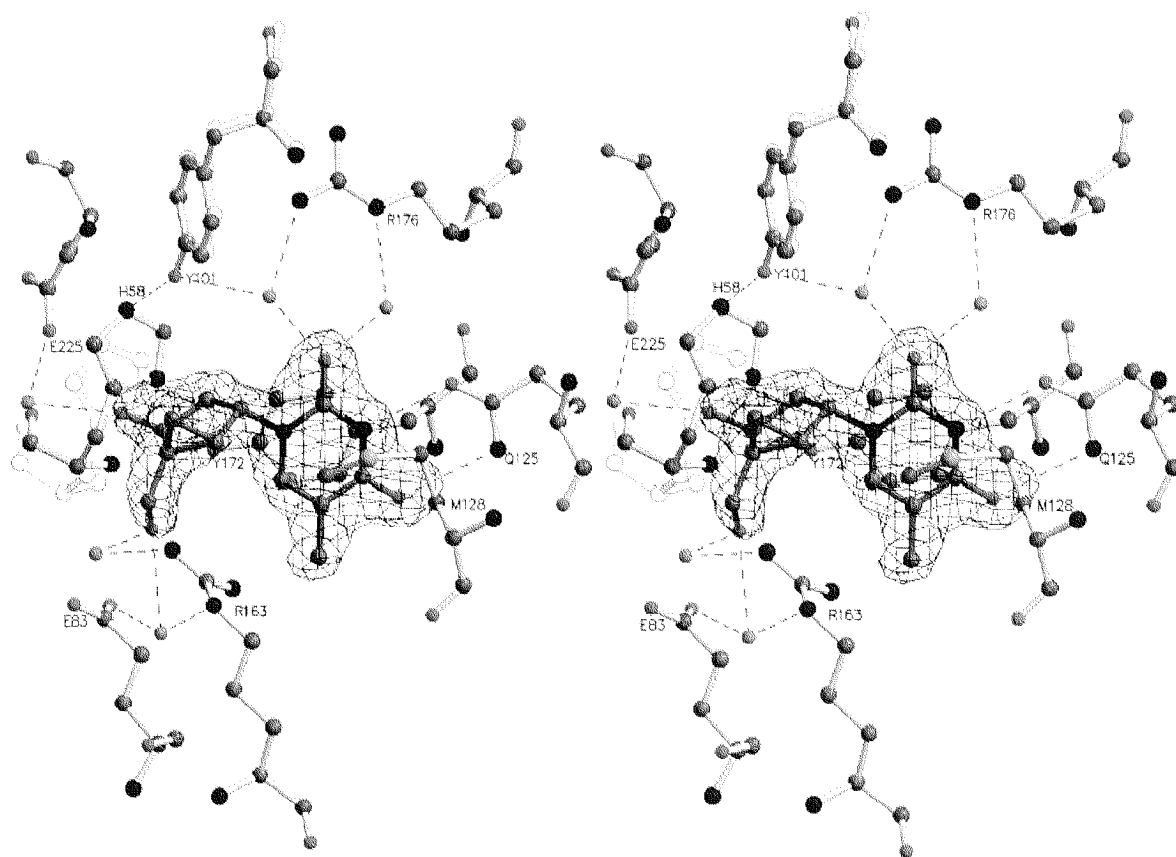
\*The  $K_i$ -data for the human tk could only be fitted by Lineweaver-Burk and indicated a very poor binding of (N)-MCT to the enzyme. The specificity is shown by the ratio  $K_i$  of (N)-MCT and  $K_M$  of dT.

### High resolution structure of $\text{TK}_{\text{HSV1}}:(\text{N})-\text{MCT}$

The complex of (N)-MCT with  $\text{TK}_{\text{HSV1}}$  crystallized in space group  $C222_1$  with two subunits A and B of  $M_r$  41 kDa per asymmetric unit. The structure has been refined to 1.7 Å with data collected at cryogenic temperatures. The R-value for this structure is 20.7 % ( $R_{\text{free}} = 25.4$ ). Only the Arg163 of both subunits is in a disallowed region of the Ramachandran diagram<sup>53</sup>. This is known from other analyses and may be caused by a participation of this residue in catalysis<sup>16-18,54,55</sup>. No higher resolved structure of HSV1 TK has been reported so far. The quality of the data was crucial for establishing the substrate binding geometry. Some protein parts could not be detected in the electron density map due to high mobility. These are in subunit A residues 1-45, 72-75, 149-152 and 266-278, in subunit B 1-45, 148-150, 221-224 and 268-273. A molecule of (N)-MCT per subunit could clearly be depicted.

The thymine ring of (N)-MCT is stacked between Tyr172 and Met128 (Figure 2) like in the wild-type structure and further fixed by a hydrogen bond network<sup>16</sup>. The strong hydrogen-bonds of N3 and O4 $\alpha$  with the carboxamide group of Gln125 are comparable in terms of distance and geometry with the complex  $\text{TK}:\text{dT}$ <sup>16,17</sup>.

are also present. The hydroxy group of Tyr172 forms a weak hydrogen bond (3.19Å) to Arg163, that forms a water-mediated hydrogen bond to the sulfate bound in the P-loop.



**Fig 2. Stereo view of the superposition of both the wildtype structure in complex with (N)-MCT and the mutant Y101F.** Only the mutation dependent shift of 1.Å of residue His58 away from the mutated residue 101 is shown. All other residues of the mutant which take the same conformation as in the wild type complex were omitted for clarity reason. The conformation of the compound is well defined by the electron density contoured at  $1.3\sigma$ . The hydrogen bonding pattern for the thymine moiety and for the 5' OH group is the same as observed for the natural substrate thymidine.

The well-defined electron density reveals clearly the ring pucker conformation. The ring pucker of the sugar-mimicking moiety of (N)-MCT is rather small (Note: there are two molecules of (N)-MCT found in the crystal structure: in subunits A and B. Molecule A has a ring pucker ( $v_{\max}$ ) of  $8.9^\circ$ , whereas molecule B is slightly more puckered with a  $v_{\max}$  of  $13.4^\circ$ . The values of  $P$  are, respectively,  $21.9^\circ$  and  $351.9^\circ$ , clearly in the northern hemisphere of the pseudorotational cycle<sup>56</sup>. Therefore, the cyclopentane ring of the bicyclo[3.1.0]hexane pseudosugar ring almost looks like a



and  $351.9^\circ$ , clearly in the northern hemisphere of the pseudorotational cycle<sup>56</sup>. Therefore, the cyclopentane ring of the bicyclo[3.1.0]hexane pseudosugar ring almost looks like a “planar pseudoboat”. The 2'-OH group of the carba-sugar moiety, the analogue of the 3'-OH of the deoxyribose of dT, is positioned at a distance of 3.7 Å to Glu225-OE1 and 4.1 Å to Tyr101-OH (4.0 Å in subunit B) indicating that, unlike dT, no direct hydrogen bonds are present for this moiety of the molecule. The direct hydrogen-bond between E225 and the 3' OH of dT reported in the structure of HSV1 TK in complex with dT is partially compensated by a water mediated H-bond between 2' OH group and E225 that has been depicted in both the viral TK structure in complex with (N)-MCT (Figure 2). These structural observations are in agreement with the kinetic data and explain the difference in affinity of ca. 60-fold between dT and (N)-MCT (Table 2).

The 5'-OH group of (N)-MCT is fixed by a direct hydrogen bond to Glu83 (3.11 Å in subunit A; 3.55 Å in subunit B) and two water mediated hydrogen bonds with Arg163 and the same Glu83 (Fig. 2). The water mediated hydrogen bonds are part of a large hydrogen bonding network involving also Asp162 and Arg222.

### **Mutant Y101F**

The complex of (N)-MCT with the active site mutant Y101F was refined to 2.4 Å resolution yielding a R-factor of 22.2% (Table 1). The ( $F_o - F_c$ ) density map confirmed the successful mutation Y101F. Neither the polypeptide chain nor (N)-MCT show significant changes in position or mobility.

The superposition of both complexes of (N)-MCT with the wild-type TK and the mutant TK shows 1 Å shift of His58 away from residue 101, causing a loss of the two hydrogen bonds involving the OH group of Tyr101 namely with His58 and the water molecule ligated with the O2 $\alpha$  of (N)-MCT (see Figure. 2). The reported structure of the TK<sub>HSV1</sub>-Y101F:(N)-MCT complex is in agreement with the kinetic data (Table 2). In the structure of TK<sub>HSV1</sub>:dT complex the 3'-OH group of the sugar moiety forms a hydrogen bond to Y101, so that its mutation to Phe weakens the binding of dT. The complex with ((N)-MCT is not affected by this mutation as this substrate does not form any hydrogen bonds to the hydroxyl group of Y101.

## DISCUSSION

The sugar ring of nucleosides and nucleotides equilibrates in solution between two extreme forms, a 2'-exo/3'-endo conformation and a 2'-endo/3'-exo conformation<sup>56</sup>. When a nucleoside or nucleotide binds to its target enzyme, only one of these conformations is expected to be present at the active site. These conformations can be mimicked in solution by introduction of a fixed sugar ring pucker. A variety of nucleobases with a bicyclo[3.1.0]hexane carbocyclic analog has been synthesized in the past<sup>23,26-28</sup> and their antiviral activity has been tested against several types of DNA viruses, such as HSV1 and HSV2, human CMV and Epstein-Barr virus. One of these compounds, namely (N)-MCT, showed an excellent and reproducible antiviral activity against HSV1 and HSV2 measured by plaque reduction assays using aciclovir as a positive control. The potency of this compound surpassed that of ACV<sup>23</sup>.

Since antiviral activity involves multi-enzymatic processes, the pathway of action has to be outlined. Cellular or viral kinases play a decisive role in these processes, as nucleosides constitute prodrugs that must be activated to the requisite 5'-phosphorylated analogs needed for antiviral activity. For this purpose we studied the interaction of the potent carbocyclic nucleoside analog, (N)-MCT, with HSV1 TK, the prodrug activator in classical antiviral therapy.

With HPLC analysis we could qualitatively demonstrate the phosphorylation of the compound by HSV1 TK showing that the antiviral activity is most likely linked to the thymidine salvage pathway in analogy to ACV. Knowing the crystal structure of the enzyme we constructed an active site mutant of the enzyme that would show the role of the fixed sugar ring pucker during the enzyme reaction and would have increased substrate specificity towards (N)-MCT and concomitant decrease of thymidine utilization. The mutant Y101F was indeed able to phosphorylate the compound to the same extent as the wild-type enzyme.

Kinetic studies with (N)-MCT were performed with wild-type HSV1 TK, mutant Y101F and human cytosolic TK. On the one hand, the binding behavior of the three enzymes towards dT has been characterized by the Michaelis constant  $K_M$ . On the other hand, the affinity of (N)-MCT has been measured ( $K_I$ ) resulting

from inhibition study because the radioactively labeled (N)-MCT was not available. In the case of HSV1 TK, it has been previously demonstrated that  $K_M$  of dT correspond to the dissociation constant of dT, and thus,  $K_M$  equals  $K_i$ <sup>52</sup>. Furthermore, for ACV, a representative of prodrugs, it has been shown that the  $K_M$  value of 0.2 mM<sup>52</sup> corresponds with the  $K_i$  values ranging from 100 to 200  $\mu$ M reported in literature<sup>57-59</sup>. Thereof,  $K_i$  and  $K_M$  can both be used as binding constant. This represents a peculiarity of HSV 1 TK by which the rate constant of disintegration of the intermediate complex, Enzyme-Substrate (ES), is negligible as compared to the corresponding dissociation rate constant<sup>52</sup>. The difference in  $K_i$  of (N)-MCT between virus and human enzymes indicates the selectivity of the compound for the viral protein and explains the low toxicity measured by in vitro cell systems<sup>23</sup>.

The binding mode of the (N)-MCT to wild-type HSV1 TK is comparable to that of dT and the difference in affinity between dT and (N)-MCT is clearly related to the loss of direct hydrogen bonds between the substrate and Tyr101 as well as Glu225. The affinity of the mutant is only 4.5 times lower for (N)-MCT as compared to the wild-type, whereas the affinity for the natural substrate, dT, decreases more than 20 fold. It should be stressed that dT still binds stronger to both wild-type and mutant although the selectivity of the mutant is now shifted towards the prodrug (N)-MCT.

Therefore, because of the observed shift in specificity Y101F HSV1 TK represents a rationally designed tailor-made enzyme for (N)-MCT and constitutes a new lead for all enzyme-prodrug gene therapy approaches for which the bystander effect, that is reduced by pyrimidine-based prodrugs<sup>61</sup>, is not a essential for increasing the therapeutic efficacy because of the very high efficiency of the transduction of the target cells with HSV1 TK. An example of such an approach is the TK-controlled graft-versus-host disease in allogeneic bone marrow transplantation.

The crystal structures clearly demonstrate that no major movements or reorientation of the substrate occurs and confirmed the kinetic measurements, showing that the conformationally restricted sugar ring pucker stabilizes (N)-MCT

in a substrate-like orientation without using additional hydrogen bonds. The loss of two hydrogen bonds compared to the natural substrate thymidine is obviously compensated by the fixed sugar moiety, which brings the 5'-OH group into the phosphorylation position and leads to a gain in entropy. Although it is known from *ab initio* calculations and from the crystal structure of the compound that the ring is almost a perfect 2-exo envelope ( $P = 343.3$  with a max puckering amplitude of 30 degrees)<sup>60</sup> (deposited Cambridge Database), the high resolution data at 1.7 Å allowed us to define the ring pucker as being more flat than expected. With this template, direct H-bond interactions between the 2'-OH of the pseudosugar moiety are absent explaining the lower affinity of wildtype HSV1 TK towards (N)-MCT and the low susceptibility of (N)-MCT towards changes at position 101 that strongly influence binding of the natural substrate, thymidine.

## CONCLUSIONS

The biochemical characterization and the structures of (N)-MCT in complex with HSV1 TK and mutant Y101F reveal the role of a fixed form of sugar ring pucker during catalysis in the TK active site. They provide a lead for the design of resistance-repellent drugs for antiviral therapy.

The result of this study can also be used for improving the efficacy of enzyme-prodrug mediated gene therapy by constructing new mutants that tend to lose entropy and can be stabilized by such a rigid pseudosugar.

## ACKNOWLEDGEMENT

We thank the team of the EMBL-outstation Hamburg for help with synchrotron data collection.

## REFERENCES

1. Roizman, B. The function of herpes simplex virus genes: a primer for genetic engineering of novel vectors. *Proc Natl Acad Sci U S A* **93**, 11307-12 (1996).
2. Roizman, B. & Sears, A. *Herpes simplex viruses and their replication*, 2231-2295 (Lippincott-Raven Publishers, Philadelphia, 1996).
3. Umene, K. & Sakaoka, H. Evolution of herpes simplex virus type 1 under herpesviral evolutionary processes. *Arch Virol* **144**, 637-56 (1999).
4. Okazaki, R. & Kornberg, A. Deoxythymidine Kinase of Escherichia coli. *Journal of Biological Chemistry* **239**, 275-284 (1964).
5. Keller, P.M. *et al.* Enzymatic phosphorylation of acyclic nucleoside analogs and correlations with antiherpetic activities. *Biochem Pharmacol* **30**, 3071-7 (1981).
6. Fyfe, J.A., Keller, P.M., Furman, P.A., Miller, R.L. & Elion, G.B. Thymidine kinase from herpes simplex virus phosphorylates the new antiviral compound, 9-(2-hydroxyethoxymethyl)guanine. *J Biol Chem* **253**, 8721-7 (1978).
7. Elion, G.B. *et al.* Selectivity of action of an antiherpetic agent, 9-(2-hydroxyethoxymethyl) guanine. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5716-20 (1977).
8. Cheng, Y.C., Grill, S.P., Dutschman, G.E., Nakayama, K. & Bastow, K.F. Metabolism of 9-(1,3-dihydroxy-2-propoxymethyl)guanine, a new anti-herpes virus compound, in herpes simplex virus-infected cells. *J Biol Chem* **258**, 12460-4 (1983).
9. McGuirt, P.V. & Furman, P.A. Acyclovir inhibition of viral DNA chain elongation in herpes simplex virus-infected cells. *Am J Med* **73**, 67-71 (1982).
10. Coen, D.M. The implications of resistance to antiviral agents for herpesvirus drug targets and drug therapy. *Antiviral Research* **15**, 287-300 (1991).
11. Wutzler, P. Antiviral therapy of herpes simplex and varicella-zoster virus infections. *Intervirology* **40**, 343-56 (1997).

12. Vere Hodge, R.A., Sutton, D., Boyd, M.R., Harnden, M.R. & Jarvest, R.L. Selection of an oral prodrug (BRL 42810; famciclovir) for the antiherpesvirus agent BRL 39123 [9-(4-hydroxy-3-hydroxymethylbut-1-yl)guanine; penciclovir]. *Antimicrobial Agents & Chemotherapy* **33**, 1765-73 (1989).
13. Ida, M. *et al.* Emergence of resistance to acyclovir and penciclovir in varicella-zoster virus and genetic analysis of acyclovir-resistant variants. *Antiviral Res* **40**, 155-66 (1999).
14. Grant, D.M., Mauskopf, J.A., Bell, L. & Austin, R. Comparison of valaciclovir and acyclovir for the treatment of herpes zoster in immunocompetent patients over 50 years of age: a cost-consequence model. *Pharmacotherapy* **17**, 333-41 (1997).
15. Fife, K.H., Barbarash, R.A., Rudolph, T., Degregorio, B. & Roth, R. Valaciclovir versus acyclovir in the treatment of first-episode genital herpes infection. Results of an international, multicenter, double-blind, randomized clinical trial. The Valaciclovir International Herpes Simplex Virus Study Group. *Sex Transm Dis* **24**, 481-6 (1997).
16. Wild, K., Bohner, T., Folkers, G. & Schulz, G.E. The structures of thymidine kinase from herpes simplex virus type 1 in complex with substrates and a substrate analogue. *Protein Sci* **6**, 2097-106 (1997).
17. Champness, J.N. *et al.* Exploring the active site of herpes simplex virus type-1 thymidine kinase by X-ray crystallography of complexes with aciclovir and other ligands. *Proteins* **32**, 350-61 (1998).
18. Bennett, M.S. *et al.* Structure to 1.9 Å resolution of a complex with herpes simplex virus type-1 thymidine kinase of a novel, non-substrate inhibitor: X-ray crystallographic comparison with binding of aciclovir. *FEBS Lett* **443**, 121-5 (1999).
19. Holy, A., Dvorakovy, H., Declercq, E.D.A. & Balzarini, J.M.R. Antiretroviral enantiomeric nucleotide Analogs. (, 1994).
20. Holy, A. & Masojdikova, M. Synthesis of enantiomeric N-(2-phosphonomethoxypropyl)derivatives of purine and pyrimidine bases. *Collection of Czechoslovak Chemical Communications* **60**, 1196 (1995).

21. De Clercq, E. *et al.* A novel selective broad-spectrum anti-DNA virus agent. *Nature* **323**, 464-7 (1986).
22. Ono, N. *et al.* Mode of action of (1'S,2'R)-9-[[1',2'-bis(hydroxymethyl)cycloprop-1'-yl]methyl]guanine (A-5021) against herpes simplex virus type 1 and type 2 and varicella-zoster virus. *Antimicrob Agents Chemother* **42**, 2095-102 (1998).
23. Marquez, V.E. *et al.* Nucleosides with a twist. Can fixed forms of sugar ring pucker influence biological activity in nucleosides and oligonucleotides? *J Med Chem* **39**, 3739-47 (1996).
24. Scapozza, L. & Folkers, G. Molecular mechanisms ruling drug resistance of Herpes Simplex Virus Type-1 linked to thymidine kinase mutations. *International Antiviral News* **6**, 210-213 (1998).
25. Sekiyama, T. *et al.* Synthesis and antiviral activity of novel acyclic nucleosides: discovery of a cyclopropyl nucleoside with potent inhibitory activity against herpesviruses. *J Med Chem* **41**, 1284-98 (1998).
26. Ezzitouni, A., Barchi, J.J.J. & Marquez, V.E. Simple approach to 1,1'a-Methano carbocyclic thymidine. *J. Chem. Soc., Chem. Commun.* , 1345-1346 (1995).
27. Siddiqui, M.A., Ford, H.J., George, C. & Marquez, V.E. Synthesis, conformational analysis, and biological activity of a rigid carbocyclic analogue of 2'-Deoxyaristeromycin built on abicyclo[3.1.0]hexane template. *Nucleosides Nucleotides* **15**, 235-250 (1996).
28. Rodriguez, J.B., Marquez, V.E., Nicklaus, M.C., Mitsuya, H. & Barchi, J.J., Jr. Conformationally locked nucleoside analogues. Synthesis of dideoxycarbocyclic nucleoside analogues structurally related to neplanocin C. *J Med Chem* **37**, 3389-99 (1994).
29. Bonini, C. *et al.* HSV-TK gene transfer into donor lymphocytes for control of allogeneic graft-versus-leukemia [see comments]. *Science* **276**, 1719-24 (1997).
30. Culver, K.W. *et al.* In vivo gene transfer with retroviral vector-producer cells for treatment of experimental brain tumors [see comments]. *Science* **256**, 1550-2 (1992).

31. Moolten, F.L. Drug sensitivity ("suicide") genes for selective cancer chemotherapy. *Cancer Gene Ther* **1**, 279-87 (1994).
32. Panis, Y. & Houssin, D. [Gene therapy. A new prospect in the treatment of liver tumors (editorial)]. *Presse Med* **24**, 1681-3 (1995).
33. Nagy, H. *et al.* Are hepatomas a good target for suicide gene therapy? An experimental study in rats using retroviral-mediated transfer of thymidine kinase gene. *Surgery* **123**, 19-24 (1998).
34. Tiberghien, P. "Suicide" gene for the control of graft-versus-host disease. *Curr Opin Hematol* **5**, 478-82 (1998).
35. Sturtz, F.G. *et al.* Parameters influencing the efficiency of the thymidine kinase/ganciclovir strategy in human glioblastoma cell lines. *Stereotact Funct Neurosurg* **68**, 252-7 (1997).
36. Klatzmann, D. Gene therapy for metastatic malignant melanoma: evaluation of tolerance to intratumoral injection of cells producing recombinant retroviruses carrying the herpes simplex virus type 1 thymidine kinase gene, to be followed by ganciclovir administration. *Hum Gene Ther* **7**, 255-67 (1996).
37. Black, M.E., Newcomb, T.G., Wilson, H.M. & Loeb, L.A. Creation of drug-specific herpes simplex virus type 1 thymidine kinase mutants for gene therapy. *Proc Natl Acad Sci U S A* **93**, 3525-9 (1996).
38. Encell, L.P., Landis, D.M. & Loeb, L.A. Improving enzymes for cancer gene therapy. *Nat Biotechnol* **17**, 143-7 (1999).
39. Kokoris, M., Sabo, P., Adman, E. & Black, M. Enhancement of tumor ablation by a selected HSV-1 thymidine kinase mutant. *Gene Ther* **6**, 1415-1426 (1999).
40. Baretino, D., Feigenbutz, M., Valcarcel, R. & Stunnenberg, H.G. Improved method for PCR-mediated site-directed mutagenesis. *Nucleic Acids Research* **22**, 541-2 (1994).
41. Steinberg, R.A. & Gorman, K.B. A high-yield method for site-directed mutagenesis using polymerase chain reaction and three primers. *Analytical Biochemistry* **219**, 155-7 (1994).



42. Fetzer, J., Michael, M., Bohner, T., Hofbauer, R. & Folkers, G. A fast method for obtaining highly pure recombinant herpes simplex virus type 1 thymidine kinase. *Protein Expr Purif* **5**, 432-41 (1994).
43. Silva, N.L., Haworth, R.S., Singh, D. & Fliegel, L. The carboxyl-terminal region of the Na<sup>+</sup>/H<sup>+</sup> exchanger interacts with mammalian heat shock protein. *Biochemistry* **34**, 10412-20 (1995).
44. Pilger, B. *et al.* Substrate Diversity of Herpes Simplex Virus Thymidine Kinase - Impact of the kinematics of the enzyme. *J Biol Chem* **274**, 31967-31973 (1999).
45. Gerber, S. & Folkers, G. A new method for quantitative determination of tritium-labeled nucleoside kinase products adsorbed on DEAE-cellulose. *Biochem Biophys Res Commun* **225**, 263-7 (1996).
46. Furlong, N.B. A rapid assay for nucleotide kinases using C14- or H3-labeled nucleotides. *Analyt. Biochem.* **5**, 515-522 (1963).
47. Garman, E.F. & Mitchell, E.P. Glycerol concentrations required for cryoprotection of 50 typical protein crystallization solutions. *J Appl Cryst* **29**, 584-587 (1996).
48. CCP4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760-763 (1994).
49. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard. Improved methods for binding protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47**, 110-9 (1991).
50. Kraulis, P.J. MOLSCRIPT - a program to produce both detailed and schematic plots of protein structures. *Journal of Applied Crystallography* **24**, 946-950 (1991).
51. Merritt, E.A. & Bacon, D.J. Raster3D: Photorealistic Molecular Graphics. *Methods in Enzymology* **277**, 505-524. (1997).
52. Kussmann-Gerber, S. *et al.* Interaction of the recombinant herpes simplex virus type 1 thymidine kinase with thymidine and aciclovir: a kinetic study. *Nucleosides Nucleotides* **18**, 311-30 (1999).

53. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK - a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291 (1993).
54. Brown, D.G. *et al.* Crystal structures of the thymidine kinase from herpes simplex virus type-1 in complex with deoxythymidine and ganciclovir. *Nat Struct Biol* **2**, 876-81 (1995).
55. Wild, K., Bohner, T., Aubry, A., Folkers, G. & Schulz, G.E. The three-dimensional structure of thymidine kinase from herpes simplex virus type 1. *FEBS Lett* **368**, 289-92 (1995).
56. Saenger, W. *Principles of Nucleic Acid Structure*, (Springer-Verlag, New York, 1984).
57. Fyfe, J.A., McKee, S.A. & Keller, P.M. Altered thymidine-thymidylate kinases from strains of herpes simplex virus with modified drug sensitivities to acyclovir and (E)-5-(2-bromovinyl)-2'-deoxyuridine. *Mol Pharmacol* **24**, 316-23 (1983).
58. Larder, B.A., Derse, D., Cheng, Y.C. & Darby, G. Properties of purified enzymes induced by pathogenic drug-resistant mutants of herpes simplex virus. Evidence for virus variants expressing normal DNA polymerase and altered thymidine kinase. *J Biol Chem* **258**, 2027-33 (1983).
59. Larder, B.A., Cheng, Y.C. & Darby, G. Characterization of abnormal thymidine kinases induced by drug-resistant strains of herpes simplex virus type 1. *J Gen Virol* **64 Pt 3**, 523-32 (1983).
60. Marquez, V.E. *et al.* Conformational analysis of nucleosides constructed on a bicyclo[3.1.0]hexane template. Structure-antiviral activity analysis for the northern and southern hemispheres of the pseudorotational cycle. *Nucleosides Nucleotides* **16**, 1431-1434 (1997).
61. Degrève, B., De Clercq, E. & Balzarini, J. Bystander effect of purine nucleoside analogues in HSV-1tk suicide gene therapy is superior to that of pyrimidine nucleoside analogues. *Gene Therapy* **6**, 162-170 (1999).

## PUBLICATIONS AND TALKS

### PUBLICATIONS

Folkers, G., Prota, A., Merz, A., Modelling of guanine-derivative-protein interaction complexes as a basis of drug design, *Farmaco* 1995 Jun 50:6 449-54

Prota, A., Vogt, J., Perozzo, R., Pilger, B., Wurth, C., Marquez, V.E., Russ, P., Schulz, G.E., Folkers, G. and Scapozza, L., Kinetics and high resolution structure of HSV1 TK and the engineered Y101F mutant in complex with an antiviral drug with conformationally restricted sugar ring pucker, *submitted to Biochemistry*, 1999

Vogt, J., Perozzo, R., Pautsch, A., Prota, A., Schelling, P., Pilger, B., Folkers, G., Scapozza, L. and Schulz, G., The rational of substrate binding and diversity of *Herpes simplex* type 1 TK studied by X-ray crystallography, *submitted to PROTEINS, Structure, Function & Genetics*, 1999

### POSTERS

Perozzo, R., Prota, A., Pautsch, A., Schulz, G.E., Folkers, G. and Scapozza, L. (1998) " *Crystallographic Study of the Native Full Length HSV1 Thymidine Kinase*", GPEN 98, Department of Pharmacy, Swiss Federal Institute of Technology (ETH), Zürich, Switzerland

A.E. Prota, R. Crameri, T. Bohner, L. Scapozza, M. Breitenbach, M. Susani, K. Blaser and G. Folkers, (10/2-5/97) " *The role of Enolase of Alternaria Alternata in Allergies*", 4<sup>th</sup> Meeting on Peptide and Protein Drugs/ Jahreskongress der SGPhW and DPhG. Zurich, Switzerland

A.E. Prota, A. Pautsch, B.D. Pilger, P. Schelling, G.E. Schulz, G. Folkers and L. Scapozza (7/17-23/99), " *From Hydrogen Bonds To Gene Therapy*", 24<sup>th</sup> International Herpesvirus Workshop, MIT, Boston, USA

L. Scapozza, B. D. Pilger, F. Alber, R. Perozzo, A. Prota, J. Vogt , C. Wurth, G.E. Schulz and G. Folkers, "*Extended substrate Acceptance of HSV1-TK as new chance for gene and antiviral therapy*", 24<sup>th</sup> International Herpesvirus Workshop, MIT, Boston, USA

### **TALKS**

"On the way to the three dimensional structure of the human cytosolic thymidine kinase" (4/1/98) Doktorandentag Campus University Irchel, Zurich, Switzerland

**CURRICULUM VITAE**

- 1969                    born on January 12<sup>th</sup> in Baden, Switzerland
- 4/1989                High school diploma (Matura Typus B)    Kantonsschule Baden,  
Switzerland
- 1989-1995            Study of pharmacy at the Swiss Federal Institute of Technology  
(ETH), Zürich, Switzerland
- 1991-1992            12 month compulsory practical training at the pharmacy  
ROTZENBÜHL APOTHEKE AG, Spreitenbach, Switzerland
- 12/1995              Graduation as federal qualified pharmacist (Eidg.  
Staatsexamen); Acquisition of exams in pharmaceutical  
chemistry, analytical chemistry, inorganic chemistry,  
pharmaceutical biology, biopharmacy, pharmaceutical  
technology and pharmacology
- since 3/1996        Ph.D. student at the Swiss Federal Institute of Technology (ETH)  
Zürich, Switzerland, Department of Pharmacy, section  
Pharmaceutical Chemistry, Prof. Dr. Gerd Folkers
- 4/1996-12/98        teaching assistant in analytical chemistry, Swiss Federal Institute  
of Technology, Zürich
- 12/1999              final examination to obtain the degree of Doctor of Natural  
Sciences, Swiss Federal Institute of Technology, Zürich

## ***Thank you all, grazie a tutti.....***

.....for the great support during all this time.

First of all I would like to thank my supervisor Professor Gerd Folkers for giving me the opportunity to join his lab. With enthusiasm he imparted his extensive way of looking at things and learned me always to look straight ahead and not to complain about defeats. Besides science, we had many opportunities to discuss about other important things in life.

For the financial support during my thesis I would like to thank Professor P.A. Schubiger.

I am grateful to Dr. Leonardo Scapozza and Dr. Oliver Zerbe for being my co-referees. Leonardo had always the right answer to solve a problem, learned me to keep cool in any situation and was a huge leadership during all these years.

Also I want to express my gratitude to Dr. Alfred Merz for introducing me in the TK group and in the secrets of computer technology, to Dr. Thomas Bohner and Isabelle Amrhein for revealing me the secrets of macromolecular crystallization.

Further I want to thank Remo for the interesting discussions and adventures in crystallizing TK, Bea for the great help in molecular biology, lab organization and support whenever English was a problem, Christine for the exciting conversations about kinetics, Pavel and Francesca for the assistance in computer graphics, Uli for his great helpfulness whenever I needed any help and the rest of the group for the whole support.

Furthermore, I want to thank the crystallographer team in Freiburg, namely Klemens Wild, Alex Pautsch and Joachim Vogt for the great collaboration in solving the structures and the innumerable discussions about crystallography.

Grazie a mamma per avermi dato la possibilità di studiare, Massimo e Luca per essere sempre stati pronti ad intervenire in momenti di bisogno ....quando la tecnica non funzionava a dovere.

Un grandissimo grazie a Carla, per essermi sempre stata accanto anche nei momenti più duri. Grazie mille per i panini e le bibite. Senza il tuo grande sostegno non ce l'avrei mai fatta!

Grazie infine a tutti coloro, che hanno contribuito al successo di questo lavoro.