Thermodinamic characterization of protein/ligand-interactions of herpes simplex virus type 1 thymidine kinase and the periplasmic domain of the histidine autokinase CitA

Author(s):
Perozzo, Remo

Publication Date:
1999

Permanent Link:
https://doi.org/10.3929/ethz-a-003886133

Rights / License:
In Copyright - Non-Commercial Use Permitted
Thermodynamic Characterization of Protein/Ligand-Interactions of Herpes Simplex Virus Type 1 Thymidine Kinase and the Periplasmic Domain of the Histidine Autokinase CitA

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

presented by

Remo Perozzo

Pharmacist (Eidg. dipl. Apotheker)
ETH Zurich

born November 19th, 1968
citizen of Zurich, Switzerland

accepted on the recommendation of

Prof. Dr. G. Folkers, examiner
Dr. L. Scapozza, co-examiner
Dr. I. Jelesarov, co-examiner

1999
# Table of Contents

**Table of Contents**

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table of Contents</td>
<td>1</td>
</tr>
<tr>
<td>Abbreviations</td>
<td>IV</td>
</tr>
<tr>
<td>Summary</td>
<td>VII</td>
</tr>
<tr>
<td>Zusammenfassung</td>
<td>XI</td>
</tr>
</tbody>
</table>

**PART 1**

**Isothermal Titration Calorimetry**

1.1. Introduction to Calorimetry | 3  
1.2. Historical Background | 3  
1.3. Calorimetric Principles and Properties | 7  
1.4. Isothermal Titration Calorimetry | 9  
1.5. Experimental Design | 12  
1.6. Data Analysis | 17  
1.7. Strategies for Measuring Tight Binding Affinity | 22  
1.8. Informational Content of ITC Data | 26  
1.9. Prediction of Binding Energetics | 31  
1.10. Thermodynamics and Rational Drug Design | 39  
1.11. References | 41  

**Herpes Virus Type 1 Thymidine Kinase**

2.1. Herpes Viruses | 51  
2.2. Thymidine Kinases | 52  
2.3. Structure and Function of HSV1 Thymidine Kinase | 55  
2.4. Medicinal Aspects of HSV1 TK | 62  
2.6. References | 64
# Table of Contents

**Periplasmic Histidine Autokinase CitA**

3.1. Citrate Metabolism 73  
3.2. Regulation of Citrate Fermentation 75  
3.3. References 77  

**Aims and Scope of the Presented Work**  
4.1. Objectives 79  
4.2. References 81  

## PART 2 83

**Method Development and Improvement**  
1.1. Introduction 85  
1.2. Expression and Purification 86  
1.3. Concentration Determination Procedures 91  
1.4. Kinetic and Activity Measurements 91  
1.5. Stability Measurements 94  
1.6. High Performance Liquid Chromatography 96  
1.7. References 99  

## PART 3 101

**Compulsory Order of Substrate Binding to HSV1 Thymidine Kinase: A Calorimetric Study**  
1.1. Summary 105  
1.2. Introduction 106  
1.3. Results 109  
1.4. Discussion 119  
1.5. Conclusion 123  
1.6. Material and Methods 124  
1.7. References 126
# Substrate Diversity of Herpes Simplex Virus Thymidine Kinase - Impact of the Kinematics of the Enzyme

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1. Summary</td>
<td>135</td>
</tr>
<tr>
<td>2.2. Introduction</td>
<td>135</td>
</tr>
<tr>
<td>2.3. Experimental Procedures</td>
<td>138</td>
</tr>
<tr>
<td>2.4. Results</td>
<td>143</td>
</tr>
<tr>
<td>2.5. Discussion</td>
<td>148</td>
</tr>
<tr>
<td>2.6. References</td>
<td>153</td>
</tr>
</tbody>
</table>

# The Periplasmic Domain of the Histidine Autokinase CitA Functions as a Highly Specific Citrate Receptor

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1. Summary</td>
<td>160</td>
</tr>
<tr>
<td>3.2. Introduction</td>
<td>160</td>
</tr>
<tr>
<td>3.3. Results</td>
<td>163</td>
</tr>
<tr>
<td>3.4. Discussion</td>
<td>172</td>
</tr>
<tr>
<td>3.5. Experimental Procedures</td>
<td>180</td>
</tr>
<tr>
<td>3.6. References</td>
<td>186</td>
</tr>
</tbody>
</table>

**PART 4**

<table>
<thead>
<tr>
<th>Appendix Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1. List of Publications</td>
<td>193</td>
</tr>
<tr>
<td>1.2. Posters</td>
<td>193</td>
</tr>
<tr>
<td>1.3. Oral Presentations</td>
<td>194</td>
</tr>
<tr>
<td>1.4. Curriculum Vitae</td>
<td>195</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACV</td>
<td>acyclovir, 9-(2-hydroxyethoxymethyl)guanine</td>
</tr>
<tr>
<td>ADK</td>
<td>adenylate kinase</td>
</tr>
<tr>
<td>ADP</td>
<td>adenosine 5'-diphosphate</td>
</tr>
<tr>
<td>ADPCP</td>
<td>β,γ-methyleneadenosine 5'-triphosphate</td>
</tr>
<tr>
<td>AMP</td>
<td>adenosine 5'-monophosphate</td>
</tr>
<tr>
<td>AMPCTF₆P</td>
<td>β,γ-difluoromethyleneadenosine 5'-triphosphate</td>
</tr>
<tr>
<td>Ap₅T</td>
<td>$P^T$-(adenosine-5')-$P^5$-(thymidine-5')-pentaphosphate</td>
</tr>
<tr>
<td>ATP</td>
<td>adenosine 5'-triphosphate</td>
</tr>
<tr>
<td>AZT</td>
<td>3'-azido-3'-deoxythymidine</td>
</tr>
<tr>
<td>CD</td>
<td>circular dichroism</td>
</tr>
<tr>
<td>ΔCp</td>
<td>heat capacity change</td>
</tr>
<tr>
<td>ΔG</td>
<td>free energy change</td>
</tr>
<tr>
<td>ΔH</td>
<td>enthalpy change</td>
</tr>
<tr>
<td>ΔS</td>
<td>entropy change</td>
</tr>
<tr>
<td>dC</td>
<td>2'-deoxycytidine</td>
</tr>
<tr>
<td>dCK</td>
<td>deoxycytidine kinase</td>
</tr>
<tr>
<td>dT</td>
<td>thymidine</td>
</tr>
<tr>
<td>dTMP</td>
<td>thymidine 5'-monophosphate, thymidylate</td>
</tr>
<tr>
<td>dTDP</td>
<td>thymidine 5'-diphosphate</td>
</tr>
<tr>
<td>dTTP</td>
<td>thymidine 5'-triphosphate</td>
</tr>
<tr>
<td>DEAE</td>
<td>diethylaminoethyl</td>
</tr>
<tr>
<td>DFT</td>
<td>density functional theory</td>
</tr>
<tr>
<td>DTT</td>
<td>D/L-dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediamine tetraacetic acid</td>
</tr>
<tr>
<td>E. coli</td>
<td>Escherichia coli</td>
</tr>
<tr>
<td>GSH</td>
<td>glutathione</td>
</tr>
<tr>
<td>GST</td>
<td>glutathione S-transferase</td>
</tr>
<tr>
<td>GST-TK</td>
<td>fusion protein of glutathione S-transferase and HSV1 TK</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance liquid chromatography</td>
</tr>
<tr>
<td>HSV1</td>
<td>herpes simplex virus type 1</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>HSV1 TK</td>
<td>herpes simplex virus type 1 thymidine kinase</td>
</tr>
<tr>
<td>hTK1</td>
<td>human cytosolic thymidine kinase</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-β-D-thiogalactopyranoside</td>
</tr>
<tr>
<td>ITC</td>
<td>isothermal titration calorimetry</td>
</tr>
<tr>
<td>Km</td>
<td>Michaelis-Menten constant</td>
</tr>
<tr>
<td>LDH</td>
<td>lactate dehydrogenase</td>
</tr>
<tr>
<td>NADH</td>
<td>nicotinamide-adenine dinucleotide, reduced form</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PK</td>
<td>pyruvate kinase</td>
</tr>
<tr>
<td>PEP</td>
<td>phosphoenol pyruvate</td>
</tr>
<tr>
<td>PMSF</td>
<td>phenylmethyisulfonyl fluoride</td>
</tr>
<tr>
<td>SDS</td>
<td>sodium dodecyl sulfate</td>
</tr>
<tr>
<td>TmpK</td>
<td>thymidylate kinase</td>
</tr>
<tr>
<td>TK</td>
<td>thymidine kinase</td>
</tr>
<tr>
<td>TK'</td>
<td>TK negative</td>
</tr>
<tr>
<td>Tris</td>
<td>tris(hydroxymethyl)aminomethane</td>
</tr>
<tr>
<td>$V_{\text{max}}$</td>
<td>rate of enzyme catalyzed reaction at infinite concentration of substrate</td>
</tr>
</tbody>
</table>

The one-letter code is used for amino acids.
Summary

Understanding of molecular recognition processes as in substrate and inhibitor binding, receptor and ligand binding or protein-protein interactions requires a complete characterization of the binding energetics and correlation of thermodynamic data with interacting structures involved. A quantitative description of the forces that govern molecular associations include determination of changes of all thermodynamic parameters, including free energy of binding (ΔG), enthalpy (ΔH) and entropy (ΔS) of binding and the heat capacity change (ΔCp). Such a close insight into the binding process is of significant and practical interest, since it provides the fundamental know-how for development of structure-based molecular design strategies. Isothermal titration calorimetry (ITC) has emerged as the premier tool for characterizing interactions in terms of thermodynamic parameters.

The main aspect of the present work is the investigation and elucidation of binding characteristics from a thermodynamic point of view, exemplified for two different systems. The first system discusses the ligand-enzyme interactions of herpes simplex virus type 1 thymidine kinase (HSV1 TK) with respect to all elementary steps forming the catalytically competent state. As for ligand-receptor interactions, the second focus of this work is the investigation of the sensory properties of the periplasmic domain (CitAP) of the histidine autokinase CitA.

Thymidine kinase is a key enzyme in the pyrimidine salvage pathway catalyzing the γ-phosphate transfer from ATP to thymidine (dT) in presence of magnesium cations to form thymidine monophosphate (dTMP). HSV1 TK has become an important target in medicinal chemistry because of its links with therapy of viral infections and gene therapy of cancer and AIDS. Thus, the accurate knowledge of the binding properties of HSV1 TK is a prerequisite for rational drug design of new and effective compounds for antiviral and antitumor therapy. The wild-type and mutant enzymes of HSV1 TK were expressed as thrombin cleavable glutathione S-transferase fusion proteins using E. coli. BL21 as the host. Activity studies and kinetic measurements showed enhanced stability of the fusion protein compared to cleaved HSV1 TK. Thus, a single-step purification protocol (affinity chromatography) has been established to yield high amounts of active and ligand free fusion protein. Conditions
for ITC measurements were subsequently optimized and verified in order to avoid possible systematic errors.

ITC has been used to investigate the binding energetics of the natural substrate dT and natural cofactor ATP to HSV1 TK, either in isolation or in presence of the other ligand, in the temperature range of 10-25°C. The results of the thermodynamic measurements show that dT is bound in the micromolar range ($K_D$ 5.3 µM; pH 7.5). Under the same conditions no binding of ATP is detectable. The measurement of the binary complexes showed a dramatic increase of affinity of both dT and ATP (46 nM and 0.26 µM, respectively). Associations are driven by favorable enthalpy changes while the entropic contributions are large and negative over the whole temperature range. The temperature dependent binding enthalpy revealed heat capacity changes in the range of 140 to 520 cal/K mol$^{-1}$.

Experimentally obtained thermodynamic quantities were linked by structure-based thermodynamic analysis to the known structure of the ternary HSV1 TK:dT:ATP complex. The pronounced changes in heat capacity as well as the observed unfavorable entropy changes indicate a sequential binding pathway, which is possibly realized by structural rearrangements of the enzyme coupled to binding of substrate and cofactor. In analogy to other nucleotide kinases, we propose a large movement of the dT binding domain and a smaller but still substantial movement of the LID domain, with the protein going from an empty (open) and less ordered conformation to a closed and compact one. Calorimetric studies of HSV1 TK harboring mutations in the residue triad H58/M128/Y172 demonstrate the strength of the calorimetric approach. Inactive mutants (M128F, M128F/Y172F) that were not distinguishable from each other by kinetic measurements, have shown to lose phosphorylation activity due to extensive alterations in binding mechanisms. Affinity is reduced by more than two orders of magnitude with much less favorable enthalpic contributions. In contrast, entropic contributions become less unfavorable, with the double mutant M128F/Y172F exhibiting even positive entropy of binding. The triple mutant H58L/M128F/Y172F regains phosphorylation activity which is also mirrored in the thermodynamic parameters revealing similar values as for the wild type enzyme. This further corroborates the importance of the interplay between binding and structural rearrangement of substrate and cofactor binding of HSV1 TK.

CitA is 547 amino acids in length and is composed of an aminoterminal periplasmic domain (131 amino acids), cytoplasmic domain, and a carboxyterminal cytoplasmic
kinase domain harboring the autophosphorylatable histidine residue. It is the sensor kinase of the two-component regulatory system CitAB of *K. pneumoniae*, which is responsible for the induction of responsible enzymes involved in citrate fermentation under anaerobic conditions. The activation of the two-component regulatory system occurs in presence of citrate.

For this work, the receptor was provided in form of the recombinantly overproduced periplasmic domain with a C-terminally attached histidine tag (CitAP<sub>His</sub>). Similarly to HSV1 TK, optimal ITC conditions have been established for this system to elucidates the binding characteristics of CitAP<sub>His</sub>.

Purified CitAP<sub>His</sub> binds citrate with high affinity ($K_D$ 5 $\mu$M, pH 7) in a 1:1 stoichiometry. In contrast, neither isocitrate nor tricarballylate exhibit demonstrable binding. Association is driven by a favorable enthalpy change while the entropic contribution is large and negative. The pH-dependence of the binding reaction indicated that the dianionic form H-citrate$^{2-}$ is the recognized species. In the presence of Mg$^{2+}$ ions the dissociation constant significantly increased, suggesting that the Mg-citrate complex is not bound by CitAP<sub>His</sub>. The result of this work reveals that the periplasmic domain of CitA is as a highly specific citrate receptor.

This work confirms that ITC is valuable method for characterizing recognition processes of small ligand and biological macromolecules. The informational content of thermodynamic data is large, and it is anticipated that it will play an important role in the elucidation of binding mechanisms and, through the link to structural data, also in rational drug design.
Zusammenfassung


Thymidinkinase ist das Schlüsselzernzym zur metabolischen Wiederverwertung von Pyrimidinen, wobei durch die Übertragung des endständigen Phosphatrestes von Adenosintriphosphat (ATP) auf Thymidin (dT) in Gegenwart von Magnesiumkationen Thymidinmonophosphat entsteht. Zwischen HSV1 TK und der Behandlung von viralen Erkrankungen sowie der Gentherapie von Krebs und AIDS wurde ein Zusammenhang gefunden, weshalb das Enzym zu einem interessanten Zielobjekt für die medizinische Forschung geworden ist. Unter diesen Umständen ist es von grossem Interesse, präzise und detaillierte Angaben über die Bindungseigenschaften von HSV1 TK zu erhalten, um mittels rationalem Wirkstoffdesign neue und effiziente Wirkstoffe für die antivirale und antineoplastische Therapie zu entwickeln. Dazu
Zusammenfassung

wurden Wildtyp und Mutanten der HSV1 TK in Form des Glutathion S-Transferase-Fusionsproteins exprimiert und gereinigt. Da Aktivitätsstudien und kinetische Messungen eine bessere Stabilität für das Fusionsprotein zeigten, wurde eine einstufige Reinigungsprozedur, basierend auf Affinitätschromatographie, zur Reinigung und Isolierung von aktivem und ligandfreiem Fusionsprotein etabliert. Kalorimetrische Messbedingungen wurden fortlaufend optimiert und mittels verschiedener Methoden auf systematische Fehler untersucht.


CitA ist ein 547 Aminosäuren langes Protein, bestehend aus einer aminoterminalen periplasmatischen Domäne (131 Aminosäuren), einer cytoplasmatischen Domäne, und einer carboxyterminalen Kinase-Domäne, die das autophosphorylierbare Histidin trägt. CitA ist die Sensorkinase des Zweikomponenten-Regulationssystems CitAB von K. pneumoniae, das für die Induktion der Enzyme des Citrat-Stoffwechsels unter anaeroben Bedingungen verantwortlich ist, wobei die Aktivierung in Gegenwart von Citrat erfolgt.

Für die vorliegende Arbeit wurde der Rezeptor in Form der rekombinant exprimierten und mit einem Polyhistidinrest modifizierten periplasmatischen Domäne (CitAP_His) eingesetzt. Analog zu HSV1 TK wurde ein experimentelles System etabliert, um die Bindungseigenschaften mittels ITC untersuchen zu können. Gereinigtes CitAP_His bindet Citrat mit hoher Affinität (KD 5 μM; pH 7) und mit einer 1:1 Stöchiometrie. Unter gleichen Bedingungen binden Citratanaloge wie Isocitrat und Tricarballylat nicht. Die Interaktion zeichnet sich durch günstige enthalpische und unvorteilhafte entropische Beiträge aus. In Gegenwart von Magnesiumkationen ist die Bindung signifikant reduziert. Die Affinität ist pH-abhängig, und es konnte gezeigt werden, dass die dianionische Form H-Citrat^{2-} erkannt wird und der Mg:Citrat-Komplex nicht an CitAP_His bindet. Die periplasmatische Domäne von CitA kann somit als ein hochspezifischer Citrat-Rezeptor bezeichnet werden.

Diese Arbeit zeigt deutlich, wie wertvoll die ITC für die Charakterisierung von Wechselwirkungsprozessen von kleinen Ligandmolekülen mit biologischen Makromolekülen ist. Thermodynamische Daten liefern viele Informationen, und es darf angenommen werden, dass die ITC in Zukunft einen hohen Stellenwert einnehmen wird, wenn es um die Aufklärung von Bindungsmechanismen und um rationales Wirkstoffdesign geht.
PART 1

Introductions
Isothermal Titration Calorimetry

1.1. Introduction to Calorimetry

A fundamental principle of all biological processes is molecular organization and recognition. Biological macromolecules are able to interact with various small and large molecules, with a high degree of specificity and with high affinity, fascinating chemists and biologists from the very beginning of modern biochemistry. A prerequisite for a deeper understanding of the molecular basis of protein-ligand interactions is a thorough characterization and quantification of the energetics governing complex formation. Calorimetry is the only technique enabling us to study directly the basic physical forces between and within a macromolecule in sufficient detail by measuring heat quantities or heat effects.

1.2. Historical Background

The background of the development of calorimetry and thermodynamics has been the subject of a variety of historical studies, and this chapter is meant to be a short summary of the most interesting aspects thereof (Cobb & Goldwhite, 1995; Daumas, 1950; Guerlac, 1976; Hemminger & Höhne, 1979; Hudson, 1992; Lodwig & Smeaton, 1974; Partington, 1989; Pledge, 1939; Wintermeyer, 1974). Calorimetry is a very old science. In principle, the historical development of calorimetry and thermodynamics began with the description and definition of temperature and heat. The first known documents from the early 17th century witness for very crude attempts to describe temperature, most of them derived by perception: "heat of a breeding hen, heat of boiling water, heat of glowing charcoal". These estimations were too rough, and therefore it was necessary to develop objective
standards. The invention of the first thermometer had its origin in the same time period. The concept of expansion of gases and liquids due to heat was already known from antiquity and was used by Galileo and Drebbel. They independently used a bulb with an open-ended stem inverted over water to observe the expansion of air. The results of these so-called thermoscops were still very inaccurate, confused with the effects of barometric pressure and lacked of scaling. The next crucial step to make satisfactory thermometers was the use of (pure) liquids instead of air (water, ethanol, mercury) in a closed compartment. Until the end of the 17th century a reliable temperature scale was established by Fahrenheit and Celsius. It was around this time when the nature of heat and its quantitative aspects became of interest.

People had speculated on the nature of heat since ancient times. It was a widespread belief that heat was a substance, some held the view that it was composed of atoms. During the 18th century the foundations of calorimetry were laid by Joseph Black. He preferred an alternative explanation of heat being a fluid that can be absorbed or squeezed out of bodies and can flow from one place to another. He recognized that heat applied to melting ice did not change the temperature of the mixture but was consumed for the solid-liquid phase transition, for the first time clearly discriminating between the "strength" and "amount" of heat. Black introduced the concept of latent heat and showed that quantities of heat could be estimated from the amount of melted ice. This view brought him to first calorimetric experiments with a simple phase-transition calorimeter. A warm probe was placed in the cavity of an block of ice, covered with a plate of ice and brought to thermal equilibrium. Furthermore, he adopted the idea of mixing water of different temperatures (mixing calorimeter) from Brooke Tylor (1723) to determine a series of latent heats of different substances.

At the same time, A.L. Lavoisier and P.S. Laplace became interested in the theory of heat. They considered the widespread mixing calorimeters as unsuitable because of several disadvantages: the need of delicate corrections for the heat capacity of vessel and thermometer, heat loss by cooling, chemically reacting substances, inmiscible liquids. Moreover, this method did not allow the measurement of the heat produced during combustion and other chemical reactions, and during
respiration, topics in which they were mostly interested. They developed the first convenient phase transition calorimeter that led to reproducing results. It was a simple but ingenious ice calorimeter, a device for measuring heat release due to respiration and combustion (Fig.1). The instrument consisted of a chamber surrounded by an ice-packed jacket, and the whole device was further insulated with another ice-packed jacket to improve accuracy. The amount of water collected from the melted ice of the inner jacket was used as a measure of the heat evolved in the chamber. The handling was difficult and experiments could only be performed on days when the outside temperature was a few degrees above freezing. With this device, Lavoisier and Laplace determined the specific heat of various substances and found fairly good results compared to modern standards. The most famous experiments were conducted about 1780, when Lavoisier and Laplace measured the heat generated by a guinea pig and determined the amount of carbon dioxide in its exhaled air during the experiment. They compared it to the heat release and to carbon dioxide formation when burning charcoal. The results were accurate enough to conclude that respiration was a form of combustion.
Despite this interesting experimental work, the resulting interpretations of the nature of heat remained unclear. Lavoisier still treated heat as a weightless substance and called it *caloric*, matter of fire. Laplace favored a mechanical explanation of heat as motion of particles of matter, a view that emerged toward the end of the 18th century out of experimental evidence provided by Count Rumford, formerly known as Benjamin Thompson. Rumford noticed that a large quantity of heat was generated when a cannon was being bored and that there was no limit of amount that could be produced by simple drilling. He concluded that heat was motion and not matter (or caloric), otherwise it had to stop when the cannon was run out of caloric. But there was a big controversy about this theory, and it was not until the middle of the 18th century when the caloric theory was finally overthrown. The kinetic gas theory was established and the concept of energy arose.

With the Industrial Revolution beginning in the 19th century, the nature of matter became of more than academic interest. With the realization that heat from combustion could produce work, the science of thermodynamics was born. It is concerned with the rules governing the interconversion of energy and is able to predict the feasibility of chemical processes.

Calorimetric measuring techniques remained more or less the same during this time period, although there were some modifications and improvements. Until the last few decades, calorimetric techniques have started to become of interest to biochemists and biologists outside a few specialized laboratories. Since practically every process, be it physical, chemical or biological, is accompanied by heat changes, it is obvious that calorimetry could serve as powerful analytical tool for a variety of applications, particular in biological sciences. With concurrent advances in molecular biology, expression and purification techniques, that made available significant amounts of homogeneous protein, there was an increasing need for more and reliable thermodynamic data. This gave the inputs to develop new and very sensitive calorimeters, requiring only small sample quantities and being able to detect accurately very small heat quantities.

Since the middle of the 20th century several calorimetric principles of different practical design have emerged. But it is only since the last few years, with the development and improvement of sufficiently sensitive, stable, user friendly and affordable commercial calorimeters, that allowed calorimetry to become an almost
routine analytical procedure in biochemical and biophysical research. Since modern instruments are very sensitive, detecting heat changes in the range of microcalories, requiring only 10 to 100 nmol of sample in a volume of 0.2 to 1.4 ml, they are usually denominated as microcalorimeters.

1.3. Calorimetric Principles and Properties

Unfortunately, there is a rather confusing collection of names describing the principles of calorimetry and calorimeters. In general, it is useful to separate three important areas: the measuring principle, the operating mode, and the type of construction (Hemminger & Höhne, 1979; Oscarson & Izatt, 1992; Wadsö, 1975; Wadsö, 1994).

1.3.1. Measurement Principles

Calorimeters are instruments used to quantify heat effects or heat effects. Several measurement principles have found use. In principle, solution calorimeters form two main groups: adiabatic calorimeters and heat conduction calorimeters.

With an ideal adiabatic calorimeter there is no heat exchange between the calorimeter and the surroundings, and the heat quantity $Q$ evolved during the experiment is directly proportional to the observed temperature change $\Delta T$, and to the heat capacity $c$ of the reaction vessel and its contents:

$$ Q = c \cdot \Delta T $$

Thus, in an experiment the heat quantity is determined by measuring the temperature change.

In an ideal heat conduction calorimeter the heat evolved is quantitatively transferred from the reaction vessel to the heat sink, a body surrounding the calorimeter which is usually made of metal. With this type of calorimeter, some property proportional to the heat flow between vessel and heat sink, is measured. Normally the heat flow is recorded by placing a thermopile wall between the vessel and the surrounding sink. The temperature difference over the thermopile gives rise to a potential or voltage signal $S$ which is proportional to the heat flow. The time integral for the heat flow,
times a calibration constant $\varepsilon$, is proportional to the heat quantity released in the experiment:

$$Q = \varepsilon \cdot \int Sdt$$  \hspace{1cm} (2)

The heat quantity is thus proportional to the area under the signal time curve.

1.3.2. Operating Mode

The most common type of calorimeter in use is the isoperibol calorimeter, also called "constant temperature environment" calorimeter. The vessel is separated by thermal insulation from the surrounding thermostated bath which forms the isothermal jacket. The insulation is usually filled with air or vacuum. Exothermic or endothermic processes will result in a temperature change that is recorded by a thermometer. In practice, there will always be a small heat loss from the vessel into the surrounding. Therefore, this calorimeters are not truly adiabatic, but quasi-adiabatic. The heat exchange cannot be neglected and must be corrected for. Isoperibol calorimeters are very simple and for fast processes also very precise instruments. They are used as reaction or solution calorimeters and as combustion calorimeters, but have not found widespread use in biochemical or biological studies.

In adiabatic shield calorimeters, the reaction vessel is additionally enclosed by a thin-walled metal envelope, the adiabatic shield, which is placed in the vacuum or air space between the reaction vessel and the thermostated bath. The temperature difference between the shield and the vessel is kept at zero during the experiment by automatically applying a suitable heat effect on the shield.

1.3.3. Construction Design

Calorimeters can be in a single or a twin arrangement. Although the single arrangement is simpler, the twin calorimeter has advantages which makes it very attractive for microcalorimetry. One of the calorimetric vessels, the reaction cell, contains the system of interest, whereas the other vessel, the reference cell, contains water or buffer. With such an arrangement the recorded signal is a differential signal,
of which the effects of thermal disturbances from the surroundings are expected to cancel out.

1.4. Isothermal Titration Calorimetry

1.4.1. Introduction

Calorimetric techniques have had a great impact to the current understanding of the mechanisms ruling association and stability of interacting systems at the molecular level. The principle calorimetric techniques applied to investigate biological macromolecules are differential scanning calorimetry (DSC) and isothermal titration calorimetry (ITC). DSC measures the enthalpy and heat capacity of thermal denaturation, and researches have learned about stability of biological macromolecule (proteins and nucleic acids) and of macromolecular assemblies (Freire, 1995; Privalov, 1989; Privalov & Potekhin, 1986; Sturtevant, 1987).

In contrast, ITC measures the heat evolved during molecular association. The direct thermodynamic observable is the heat associated with a binding event, i.e. a ligand is titrated into a solution containing the macromolecule of interest and the heat evolved or absorbed is detected. It allows the simultaneously determination of the equilibrium binding constant (Kₐ) and thus the standard Gibbs free energy change (ΔG), the enthalpy change (ΔH), the entropy change (ΔS), as well as the stoichiometry of the association event. Moreover, experiments performed at different temperatures yield the heat capacity change (Δ Cp) of the binding reaction (Chen & Wadso, 1982; Freire et al., 1990; Wiseman et al., 1989). As almost any interacting system is characterized by changes in enthalpy, there is a vast range of potential ITC applications.

1.4.2. Titration Microcalorimeters

Titration calorimeters are characterized by the addition (injection) of one binding partner into a solution of the other binding partner which is placed in the sample cell of the calorimeter. It is worth emphasizing that calorimetric binding experiments are very challenging since non-covalent binding heats are intrinsically small, typically in the range of 5 kcal/mol to 10 kcal/mol, and must be liberated stepwise during the binding experiment. Furthermore, ligand addition has additional heat effects arising
from dilution and mixing, for which corrections must be made, and which are frequently comparable to the binding heat of interest.

Considering the case of a typical reaction of interest which exhibits the heat effect of -5 kcal/mol in a 1 ml to 2 ml solution containing $10^{-7}$ moles of protein (a few milligrams), the experiment would liberate about 0.5 mcal of heat after complete saturation of all binding sites. Assuming that this heat is released upon 10 injections, the mean individual contribution would be 50 µcal. Accurate detection, with an accuracy of 10% or better, of such small quantities would require instrumental sensitivity and noise levels down to 5 µcal and less. This corresponds to temperature changes in the sample solution of just a few millionths of a degree and is comparable to the inevitable heat effects of dilution, mixing and stirring.

1.4.3. The OMEGA Titration Calorimeter

The use of the calorimeter in the titration mode allows fast and accurate determination of the thermodynamics of ligand binding to macromolecules and related processes. Recently, several microcalorimeters designed for direct measurement of the energetics of biological processes have been reported (Chen & Wadso, 1982; Freire et al., 1990; McKinnon et al., 1984). With the introduction of the commercially available Omega titration calorimeter from MicroCal Inc. (Northhampton, Massachusetts, USA) in 1989 (Wiseman et al., 1989), titration calorimetry has had a broad impact throughout biotechnology, which is reflected by the large body of publications during the last decade.

This type of calorimeter consists of two cells, the sample cell and the reference cell, that are insulated by an adiabatic shield (Fig.2). The cells are coupled by a thermoelectric device which measures the temperature difference $\Delta T_1$ between the cells. A second temperature difference $\Delta T_2$ is monitored between the two cells and the adiabatic shield. Both cells and the jacket are supplied with heaters that are connected to a feedback circuit controlling $\Delta T_1$ and $\Delta T_2$ to be zero. During an experiment, a small constant power (< 1 mW) is dissipated in the heater of the reference cell. This activates the feedback circuit to drive $\Delta T_1$ back to zero, thus slowly increasing the temperature during a measurement (typically less than 0.1°C/h). The resting power applied to the sample cell is the baseline signal.
Exothermic reactions as a result of an addition of ligand will decrease the necessary feedback power to maintain $\Delta T_1=0$, and endothermic reactions lead to an increase in feedback power. The enthalpy of reactions are obtained by integration of the deflections from the resting baseline.

Both cells are accessible by long narrow access tubes through which samples are introduced or removed using long-needled syringes. Typically, the reference cell is filled with water and the sample cell with the system of interest. The ligand is applied by injection syringes with long needles having a stirring paddle attached to the extreme end. The syringe is continuously rotated during an experiment, leading to complete mixing in the cell within a few seconds after an injection. The mechanical heat of stirring is constant and becomes part of the resting baseline.

With optimal performance (short equilibration time) a complete binding isotherm may be determined within 30 minutes, although in practice it takes usually 80 to 100
minutes to obtain reliable data. The absolute detection limit of this calorimeter, expressed as the minimum detectable heat quantity, is reported to be 0.3 μcal (Wiseman et al., 1989).

1.5. Experimental Design

1.5.1. General Experimental Setup

The setup of an ITC experiment is largely dependent on the thermodynamic characteristics of the system of interest, i.e. the expected binding affinity and the heat effect of the interaction.

The appropriate concentration range for the macromolecule placed in the cell depends on the binding constant of reaction. The shape of the binding curve is dependent on the product of the binding constant $K_B$ (in M$^{-1}$) and the concentration of macromolecule $[M_T]$ (in M) being titrated (Wiseman et al., 1989):

$$C = K_B [M_T]$$

The sensitivity of the shape of the binding isotherm to the dimensionless parameter $C$ is crucial for determination of the binding constant. At high values for the so-called $C$-value, the shape of the curve approaches a step function and becomes increasingly insensitive to changes in $K_B$ (Fig.3). Therefore, very strong binding ($10^7$ to $10^8$ M$^{-1}$) requires low macromolecule concentrations to measure at lower $C$-values. With decreasing concentrations of the reactants, the signal arising from the interactions will also become smaller, leading to the detection limit of ITC. This gives the technical limit of the highest affinity constant determinable.

At low $C$-values, the binding curve becomes a horizontal trace that again yields very little information about $K_B$. Consequently, it is necessary to use high macromolecule concentrations to obtain informative binding isotherms. In principle, there is no lower limit to the $C$-value that allows accurate determination of the binding constant, but in practice there are problems with solubility, stability and often with availability of the macromolecule. From Fig.3, it is evident that even at low ligand concentration, only a minor fraction of ligand is bound to the macromolecule, making it difficult to detect
Fig. 3: Simulated calorimetric binding curves illustrating the dependence of the shape of the curve on the product of the association constant $K_0$ and the total macromolecule concentration ($C = K_0[M_T]$). The curves are simulated for several $C$-values (as indicated in the plot) according to equation (16) and (18) for $\Delta H = -10$ kcal/mol. For high $C$-values the binding isotherms approach a step function, becoming increasingly insensitive to changes in $K_0$. At low $C$-values, the binding curve becomes a horizontal trace that yields very little information about $K_0$, making it necessary to use high macromolecule concentrations to obtain suitable binding isotherms.

sufficient heat and to determine $\Delta H$ accurately. This often sets the limit of the lowest affinity constant measurable in the range of $10^4$ M$^{-1}$.

The correct choice of reactant concentrations depends not only on the magnitude of $K_0$, but also on the objective of the experiment. Considering the limiting sensitivity of 0.5 $\mu$cal, each injection should produce an average heat change of 5 to 10 $\mu$cal in the 1.4 ml cell. For a series of 10 injections, each of 10 $\mu$l, a total $Q$ of 100 $\mu$cal in the sample volume are required to define a total binding curve:

$$Q = \Delta H [M_T] V_0$$

(4)

where $\Delta H$ is the enthalpy of binding and $V_0$ is the reaction volume of the sample cell. Solving equation (4) for $[M_T]$ gives a minimum concentration of ca. 7 $\mu$M for a protein with a $\Delta H$ of -10 kcal/mol (or +10 kcal/mol), needed to generate a complete binding isotherm to yield $n$, $K$ and $\Delta H$. According to equation (3), a $K_0$ of $10^6$ M$^{-1}$ would result
in a $C$-value of 7 which is reasonable. Nevertheless, in practice it is often more convenient to use higher concentrations so that the heat signals will be larger and the $C$-values will be in the ideal range of 10 to 100.

Above treatment yields $\Delta H$ as a fitting parameter (see 1.6.). A better practice is to measure $\Delta H$ at concentrations when the binding partners are fully associated and the saturation is still low: full association at partial saturation (Bains & Freire, 1991). At these conditions ($C$-value $> 100$), the amount of heat released or absorbed is directly determined by the amount of ligand injected. This allows the determination of $\Delta H$ without deconvolution of the total binding isotherm with a single injection:

$$Q = \Delta H [L_T] V_{\text{inj}}$$  \hspace{1cm} (5)

where $[L_T]$ is the concentration of the ligand solution in the syringe, and $V_{\text{inj}}$ is the injection volume.

To obtain high quality data, an appropriate protocol has to be established by optimizing ligand and protein concentrations and the injection volume. Typically, the ligand concentration is much higher since several equivalents must be added in the sample cell. The titration experiment should be planned to approach or reach complete saturation of the binding sites at the end of the experiment. To generate a sufficient number of data points, which will improve data analysis, the ligand has to be added in small aliquots. However, the heat signal should not become too small to maintain high precision of each data point. If the interaction heat is small, it will be necessary to choose larger injection volumes. These prerequisites define the titration protocol, and it is up to the experimentator to find the ideal compromise. As a rule of thumb, 25 injections, each of 5 $\mu$l, of a ligand solution with a concentration 25 times higher than the protein solution will result in adequate binding isotherm. If the ligand is poorly soluble, it is possible to place it in the sample cell and to inject the macromolecule. As long as the binding stoichiometry is 1:1, either interacting molecule can act as the titrant without adjusting the binding model. For more complicated cases where this assumption does not hold, the model must be modified accordingly (Bhatnagar & Gordon, 1995).

The time between successive injections is another important parameter. If association is rapid, the instrument baseline will be equilibrated in a short time,
depending on the response time of the calorimeter. Under such conditions 3 to 4 min are sufficient to reach baseline again after injection. In contrast, heat signals of slow processes required much more time to reach thermal equilibrium. Several other problems of experimental design should be mentioned. It is crucial that solutions of ligand and macromolecule are pure and exactly match with respect to pH, buffer capacity, and salt concentration. This means that macromolecule and ligand are preferably dissolved in the same buffer. To achieve this goal, it is good practice to dialyze the protein prior to the experiment and dissolve the ligand in the dialysis buffer. This procedure will prevent spurious heat effects resulting from mixing of different buffers. Both interacting components, often purified form biological source, must be free of contaminating enzymatic activity which could affect the association event under investigation. Furthermore, the formation of air bubbles has to be avoided. Thus it is very important to thoroughly degas all solutions prior the experiment. Any air in the syringe can cause variation in the injected volume or lead to additional heat signals, and bubbles in the sample cell interfere with the thermal contact of solution and cell wall. Finally, in most experiments the heat effect of the first injection of a series of injections is obviously to small. This results from diffusion while equilibrating the system. Even if care is taken to avoid this leakage, the problem persists, therefore it is common proactive to make a small first injection of 1 µl and then to remove the first data point before data analysis.

1.5.2. Control Experiments

ITC measures heat released or absorbed during binding reactions. Since heat is an ubiquitous phenomenon associated with any molecular interaction process, this methodology is being widely adopted. However, calorimetric techniques detects the total heat effect in the calorimetric cell upon addition of ligand. Thus, the experimental heat effect contains contributions arising from non-specific effects, such as dilution of ligand in the buffer, dilution of the protein sample, temperature differences between the cell and the syringe, and mixing of buffers of slightly different composition. These contributions can be determined by performing a control experiment. To do this, the buffer is filled in the sample cell and titrated according to the same protocol with the ligand solution. This experiment will provide the heat of dilution of ligand and macromolecule, which should be subtracted.
In most cases these heat contributions are found to be small and frequently negligible. If the titration experiment is designed to ensure complete saturation of the enzyme before the final injection, and above blanks show that the heat of dilution of the ligand is concentration-independent, the non-specific heat effects can be estimated very well by averaging the small heats at the end of the titration.

1.5.3. Evaluation of Protonation Effects

Whenever binding is coupled to changes in protonation state of the system, the measured heat signal will contain the heat effect due to ionization of buffer. If binding changes the protonation state of free or bound ligand as well as of free or complexed macromolecule, proton transfer with the buffered medium occurs. As a consequence, the heat of protonation/deprotonation will contribute to the overall heat of binding and \( \Delta H_{\text{obs}} \) will depend on the ionization enthalpy of the buffer (\( \Delta H_{\text{ion}} \)). Repeating the calorimetric experiment at the same pH in buffers of different \( \Delta H_{\text{ion}} \) allows to determine the number of protons \( n_{H^+} \) that are released (\( n_{H^+} > 0 \)) or taken up (\( n_{H^+} < 0 \)) by the buffer, and thus to calculate the intrinsic binding enthalpy, \( \Delta H_{\text{bind}} \), corrected for protonation heats:

\[
\Delta H_{\text{obs}} = \Delta H_{\text{bind}} + n_{H^+} \Delta H_{\text{ion}}
\]

In practice, it is recommended to perform ITC experiments in a series of buffers of different ionization heats under otherwise the same conditions. Values of \( \Delta H_{\text{ion}} \) have been described (Christensen et al., 1976; Fukada & Takahashi, 1998) or can be determined by ITC (Jelesarov & Bosshard, 1994). If the same \( \Delta H_{\text{obs}} \) is observed, there is no protonation event coupled to binding. Deviations of \( \Delta H_{\text{obs}} \) with different buffers point to a protonation event, and the intrinsic enthalpy of binding (\( \Delta H_{\text{bind}} \)) can be obtained from the intercept (\( \Delta H_{\text{ion}} = 0 \)) of the regression line described by equation (6) (Cooper & Johnson, 1994b).

This calorimetric buffer effect is by no means a nuisance. On the contrary, it is one of the most powerful means to investigate binding mechanisms, and this property can be exploited to increase the signal strength of an otherwise undetectable binding
event by simply changing the buffer system to a different pH and to buffers with high ionization enthalpies (Jelesarov & Bosshard, 1994).

1.6. Data Analysis

Ligand binding may be studied by a variety of experimental techniques (Connors, 1987), but titration microcalorimetry is a particularly useful method since it offers not only the possibility of determining the association constant $K_a$, but also the enthalpy ($\Delta H$), entropy ($\Delta S$) and stoichiometry ($n$) in a single experiment.

The signal monitored by ITC is the differential power applied to the sample cell. The total heat released or absorbed upon an injection of ligand into the cell corresponds to the area under the signal vs. time curve (Fig.5, panel A). Integration of the raw data yields the heat per injection which is plotted against the total ligand

![Diagram](image)

Fig.5: Calorimetric data for the exothermic binding of cytidine 2'-monophosphate (2'CMP) to ribonuclease A (RNase A) at pH 5.5 (0.2 M K-acetate, 0.2 M KCl) and 28° C a) Raw data obtained for 25 automatic injections of 5 μl. Concentrations of RNase A and 2'CMP are 0.145 mM and 3.72 mM respectively. The area of each peak represents the total heat evolved upon addition of a single aliquot of 2'CMP. B) Titration plot derived from the integrated and for heats of dilution corrected heats of binding. The solid line represents the best least-squares fit to the data assuming a single site binding model.
concentration to generate the binding isotherm (Fig. 5, panel B). The instrumental baseline and other unspecific heat effects must be carefully subtracted from the raw data. In general, the solution in the sample cell is diluted between 10% and 20%, whereas the ligand solution, typically much more concentrated, undergoes a dilution of up to several hundred fold when injected into the cell.

Experimental data can be presented as a sigmoid plot (differential mode) or as a hyperbolic saturation curve (integral mode). The differential mode treats each injection as an independent point and is plotted as heat evolved per injection versus total ligand concentration or the ratio of the total ligand concentration to the concentration of macromolecule (Fig. 5). In the integral mode, the total cumulative heat is plotted against the total ligand concentration. Fitting a binding model to the calorimetric data plotted in either mode yields equivalent results. Generally, random errors tend to cancel out in the integral mode, whereas systematic errors tend to be amplified. Comparative statistical analysis of both modes can give information about the accumulation of systematic errors (Bundle & Sigurskjold, 1994; Sigurskjold et al., 1991).

1.6.1. Ligand Binding in Titration Calorimetry

The reversible association between a macromolecule M and a ligand L,

\[ M + L \leftrightarrow ML \]  

is characterized by its binding constant \( K_b \):

\[ K_b = \frac{[ML]}{[M][L]} \]  

There are many techniques available to measure \( K_b \). Equilibrium dialysis, radio-ligand binding assays or ultracentrifugation directly yield values for \([M]\), \([L]\) or \([ML]\) to calculate \( K_b \). Spectroscopic methods are more indirect by detecting an observable \( p \) the change of which is proportional to the degree of saturation (Bundle & Sigurskjold, 1994; Cooper & Johnson, 1994a; Fisher & Singh, 1995; Freire et al., 1990).
ITC is the most direct method to measure the heat change on complex formation at a constant temperature. The observable response of an ITC experiment is the heat change associated with each addition of ligand. For each injection, the heat released or absorbed is directly proportional to the total amount of formed complex. This can be expressed by

\[ q = V_0 \Delta H \Delta [ML] \]  

(9)

where \( q \) is the heat associated with the change in complex concentration, \( \Delta [ML] \), \( \Delta H \) is the molar enthalpy of binding, and \( V_0 \) is the reaction volume of the sample cell.

In a calorimetric experiment, each addition of ligand gives rise to a heat change depending on the reaction volume, concentrations, molar enthalpy, binding constant, heat of dilution, stoichiometry and the amount of previously added ligand. As the concentration of unoccupied binding sites begins to decrease, the heat changes decrease correspondingly as ligand is added. The total cumulative heat after the \( i \)th addition, \( Q \), will be

\[ Q = V_0 \Delta H \sum \Delta [ML] = V_0 \Delta H [ML]_i \]

(10)

where \( [ML]_i \) is the total concentration of complex after the \( i \)th injection.

Evaluation of microcalorimetric data requires the consideration of the observable response in terms of total ligand added or the total ligand concentration. Therefore, the binding equations must be expressed as a function of total ligand and macromolecule concentration:

\[ [M_r] = [ML] + [M] \]

(11)

\[ [L_r] = [ML] + [L] \]

(12)

where \([M_r]\) and \([L_r]\) are total macromolecule and ligand concentrations, respectively, and \([M]\) and \([L]\) are free concentrations of macromolecule and ligand, respectively. \([ML]\) is the concentration of the formed complex.
1.6.2. Single Set of Independent Sites Model

In the simplest case of ligand binding, each macromolecule consists of only one type of binding sites with a finite number of identical noninteracting binding sites, all of which exhibiting the same intrinsic affinity for the ligand. For such a system, the binding constant $K_b$ is given by

$$K_b = \frac{\Theta}{(1 - \Theta)[L]}$$

where $\Theta$ is the fractional saturation and $[L]$ is the concentration of free ligand. It is related to the total ligand $[L_T]$ and macromolecule concentration $[M_T]$, by mass conservation:

$$[L] = [L_T] - n\Theta[M_T]$$

Combining equations (11) and (12) gives the quadratic equation

$$\Theta^2 - \Theta \left(1 + \frac{1}{nK_b[M_T]} + \frac{[L_T]}{n[M_T]}\right) + \frac{[L_T]}{n[M_T]} = 0$$

whose only meaningful root is

$$\Theta = \frac{1}{2} \left(1 + \frac{1}{nK_b[M_T]} + \frac{[L_T]}{n[M_T]} - \sqrt{\left(1 + \frac{1}{nK_b[M_T]} + \frac{[L_T]}{n[M_T]}\right)^2 - 4\frac{[L_T]}{n[M_T]}}\right)$$

The integral heat of reaction $Q$ after the $i^{th}$ injection is given by

$$Q = n[M_T]V_0 \Delta H \Theta_i$$

where $V_0$ is the cell volume and $\Delta H$ is the molar heat of ligand binding. The differential heat of the $i^{th}$ injection is
A nonlinear fit based on equation (17) to the hyperbolic saturation curve in the integral mode (Q vs. [L_i]) yields the parameters $K_B$, $\Delta H$ and $n$ from a single experiment. Based on equation (18), the titration data can be fit to the sigmoid saturation curve in the differential heat mode ($q_i$ vs. $[L_i]$, or vs. $[L_i]/[M_i]$). The same parameters are obtained.

### 1.6.3. Other Models

Similar relationships as described above exist for other models, i.e. a model for multiple sets of independent binding sites, single set of interacting sites (cooperative sites), multiple sets of interacting binding sites. By use of statistical thermodynamic treatment it is possible to deconvolute a binding isotherm of such complex systems (Connors, 1987; Di Cera, 1995; Wymann & Gill, 1990). Instructive examples from literature demonstrate the strength of this approach (Bruzzone & Connelly, 1997; Eisenstein et al., 1994; Ferrari & Lohman, 1994; Gopal et al., 1997; Hyre & Spicer, 1995). However, the success strongly depends on the quality and reliability of the experimental data.

### 1.6.4. Basic Thermodynamic Relationships

The binding enthalpy of protein-ligand interactions can be determined accurately by means of ITC. The association constant $K_B$ is related to the Gibbs free energy $\Delta G$ by the well known relation

$$\Delta G = -RT \ln K_B$$

(19)

where $R$ is the universal gas constant equal to 1.987 calK$^{-1}$mol$^{-1}$ and $T$ is the temperature in degrees Kelvin. $\Delta G$ is again composed of an enthalpy term ($\Delta H$) and an entropy term ($\Delta S$), related by another fundamental equation:

$$\Delta G = \Delta H - T\Delta S$$

(20)
The Gibbs free energy is temperature dependent and is described by

$$\Delta G(T) = \Delta H(T_0) + \int_{T_0}^{T_n} \Delta C_p dT - T \Delta S(T_0) - \int_{T_0}^{T_n} \Delta C_p d\ln T$$  \hspace{1cm} (21)$$

where $\Delta C_p$ is the heat capacity change and $T_0$ is an appropriate reference temperature. With $\Delta C_p$ being independent of temperature in the range of interest, equation (21) simplifies to

$$\Delta G(T) = \Delta H(T_0) - T \Delta S(T_0) + \Delta C_p \left[ T - T_0 - T \ln \frac{T}{T_0} \right]$$  \hspace{1cm} (22)$$

Equation (22) shows that enthalpy and entropy changes are dependent on temperature through the heat capacity change $\Delta C_p$:

$$\Delta H(T) = \Delta H(T_0) + \Delta C_p(T - T_0)$$  \hspace{1cm} (23)$$

$$\Delta S(T) = \Delta S(T_0) + \Delta C_p \ln \frac{T}{T_0}$$  \hspace{1cm} (24)$$

In a thermodynamic analysis the goal is to determine $\Delta G$, $\Delta H$, $\Delta S$ and their temperature dependence by $\Delta C_p$, since these four parameters provide a full description of the energetics governing molecular interactions.

1.7. Strategies for Measuring Tight Binding Affinity

High affinity binding constants for protein-ligand interactions are inherently difficult to measure. With increasing affinity, it becomes necessary to work at low concentration of macromolecule, leading to difficulties in detecting the signal specific to the analytical method. In the case of ITC, the largest binding constant ($K_b$) that can be reliably measured, approaches $10^9$ M$^{-1}$ for a typical macromolecule-ligand interaction (Doyle et al., 1995; Freire et al., 1990; Wiseman et al., 1989).

The thermodynamic approach offers a powerful advantage for measuring tight binding affinities and thus Gibbs free energy changes ($\Delta G$). Free energy changes are state functions, i.e. their values are defined by the initial and final thermodynamic
states, regardless of the pathway connecting the two states. This being the case, it is possible to determine the binding constant for a protein-ligand interaction under different conditions that allow to measure the affinity. The result can be corrected to other conditions of more physiological relevance, if it is known by what parameters the binding free energy is linked with the conditions being varied.

In principal, any change of physical or chemical conditions that influence ligand association to a macromolecule is useful, but in most cases linkage of binding with pH and temperature are exploited.

1.7.1. Linked Protonation Effects in Ligand Binding

Often molecular interactions are very tight (> 10^9 M^-1), and are not accurately measurable even with the most sensitive calorimeters available. Generally, molecular interactions occur to some degree in dependence of pH, reflecting the linkage between the association of a ligand and the binding of protons (proton linkage). The molecular basis of the linkage is the result of alterations of pKa values of ionizable amino acid groups concomitant with binding (Doyle et al., 1995).

If ligand binding is coupled with uptake of a single proton, the observed ligand binding constant $K_{\text{obs}}$ is given as

$$K_{\text{obs}} = \frac{K_{\text{int}} 10^{-pH}}{1 + K_{\text{pH}} 10^{-pH}}$$

where $K_{\text{int}}$ is the intrinsic binding constant, $K_{\text{pH}}$ and $K_{\text{pC}}$ are the proton binding constants for the complex and free form of the protein (equal to $10^{pKa}$ of the ionizing group) (Baker & Murphy, 1996; Bradshaw & Waksman, 1998). According to equation (25) proton linkage can be viewed as change in proton affinity, thus protons will either be released or absorbed due to ligand binding.

If proton transfer occurs during binding, the $\Delta H_{\text{obs}}$ is determined by the ionization enthalpy of the buffer and the enthalpy of binding corrected for buffer effects (see 1.5.3., equation (6)), and both the number of protons ($n_{H^+}$) and the intrinsic binding enthalpy ($\Delta H_{\text{bind}}$) will vary as a function of pH. Thus $n_{H^+}$ is given by

$$n_{H^+} = f^- - f^+$$
where \( f^c \) and \( f^f \) are the fractional saturation of protons at a given pH of the bound and free protein. In case of a single protonation event, \( f^c \) and \( f^f \) can be expressed as

\[
f^c = \frac{K_p^c 10^{-pH}}{1 + K_p^c 10^{-pH}} \quad \text{and} \quad f^f = \frac{K_p^f 10^{-pH}}{1 + K_p^f 10^{-pH}}
\]  

(27) and (28)

The change in the number of protons bound by the protein upon binding of the ligand is the difference between (27) and (28):

\[
n_{H^+} = f^c - f^f = \frac{K_p^c 10^{-pH}}{1 + K_p^c 10^{-pH}} - \frac{K_p^f 10^{-pH}}{1 + K_p^f 10^{-pH}}
\]  

(29)

Equation (29) clearly shows that at a minimum of two pH values the pKa in the liganded and free protein can be determined if \( n_{H^+} \) at the corresponding pH are known, even when the ligand affinity is too tight to be measured (Baker & Murphy, 1996).

In practice, \( n_{H^+} \) is determined in a series of buffers of different ionization enthalpies as a function of pH (see 1.5.3.). Equation (27) is used to determine the pKa of the protein in the free and complexed state. With these values \( K_B \) at the tight binding conditions can be calculated by equation (25) (Baker & Murphy, 1996; Bradshaw & Waksman, 1998).

The power of the calorimetric approach in evaluating proton linkage lies in the fact that \( \Delta H \) can be determined with high precision under conditions where \( K_B \) is not measurable, and thus the contributions of the linkage.

### 1.7.2. Thermodynamic Linkage to Temperature

The fundamental equations \( \Delta G = \Delta H - T \Delta S \) and \( \Delta G = -RT \ln K \) show that the equilibrium constant for a process is related to standard entropy and enthalpy changes, and to the absolute temperature. The temperature dependence of the changes in free energy (\( \Delta G \)) of the Gibbs-Helmholtz equation for a thermodynamic system is described as
where \( T \) is the absolute temperature and \( \Delta H \) is the reaction enthalpy. Substitution of equation (30) with \( \Delta G = -RT\ln K \) yields the familiar van't Hoff equation:

\[
\frac{\delta \ln K}{\delta (1/T)} = -\frac{\Delta H}{R}
\]  

(31)

where \( K \) is the binding constant. The temperature dependence of \( K \) is commonly analyzed by means of the van't Hoff plot, whose basis is equation (31). By measuring \( K \) over a temperature range and plotting \( \ln K \) versus \( 1/T \), a van't Hoff \( \Delta H (\Delta H_{vH}) \) is calculated from the slope of the plot, according to equation (31).

When \( \Delta H \) is known, then integration of equation (31) gives the temperature dependence of the equilibrium constant:

\[
K(T) = K(T_0)\exp\left[-\frac{\Delta H}{R}\left(\frac{1}{T} - \frac{1}{T_0}\right)\right]
\]  

(32)

Equation explicitly assumes that \( \Delta H \) is constant over the temperature range \( T \) to \( T_0 \). It has been shown for many biological protein-ligand interactions that this assumption is not valid, with \( \Delta H \) often being temperature dependent.

The temperature dependence of the binding enthalpy, \( \Delta Cp \), is described by equation (23), and combination with equation (32) leads to the extended form of the van't Hoff equation that accounts for the temperature dependence of \( \Delta H \):

\[
K(T) = K(T_0)\exp\left[-\frac{\Delta H(T_0)}{R}\left(\frac{1}{T} - \frac{1}{T_0}\right) + \frac{\Delta Cp}{R}\left(\ln \frac{T}{T_0} + \frac{T_0}{T} - 1\right)\right]
\]  

(33)

where \( K(T) \) is the binding constant to be calculated at temperature \( T \), \( K(T_0) \) is the binding constant experimentally determined at temperature \( T_0 \), and \( \Delta H(T_0) \) is the experimental enthalpy change at temperature \( T_0 \).
In cases where affinity is too tight to measure directly by ITC under ambient conditions, the binding constant becomes accessible by changing to a temperature where binding is lowered. After determining $\Delta H$ and the temperature dependence thereof, the binding constant can easily be calculated for the temperature of interest (Doyle & Hensley, 1998).

1.7.3. Displacement Experiments

As an alternative approach for measuring tight binding constants, a displacement experiment can be carried out. The protein of interest is presaturated with a weaker binding ligand whose binding parameter can be determined directly, and this ligand is displaced by injecting a stronger binding one. The first ligand will inhibit the second one and thus reduce the apparent binding constant (Hu & Eftink, 1994; Khalifah et al., 1993; Sigurskjold et al., 1994).

A first experiment yields the thermodynamic parameters for the first ligand ($\Delta H_1$, $K_1$), the second titration gives apparent values for the second ligand ($\Delta H_{\text{obs}}$, $K_{\text{obs}}$). These are related to the values of the first ligand, which yields binding constant and enthalpy by following relations

$$\Delta H_2 = \Delta H_{\text{obs}} + \Delta H_1$$

$$K_2 = K_{\text{obs}} K_1$$

If binding of both ligands is exothermic, $\Delta H$ of the tight binder will be reduced by the endothermic contribution of the dissociation of the first ligand. Therefore, it is necessary that the interaction enthalpies for both ligands differ significantly.

1.8. Informational Content of ITC Data

The gain of knowledge through thermodynamic data of binding reactions is large and can have a dramatic impact on characterizing the molecular mechanism of binding. The thermodynamic profile of an interaction process reflects various types of forces that drive binding, including enthalpic contributions of bond formation, entropic effects such as restrictions of degrees of freedom, the release and uptake of water and ion
molecules, the burial of water-accessible surface area and changes in vibrational content.

As outlined before, calorimetry measures a global property of a system, and thus reflects the sum of all concomitant phenomena which must be carefully analyzed and quantified in order to yield parameters conforming to the binding event properly. Moreover, the quality of the deconvoluted parameters depends on the appropriate model used. Nevertheless, once reliable data are available, the informational content of thermodynamic data can have an intriguing impact on the characterization of molecular mechanisms of binding.

1.8.1. Binding Free Energy

The Gibbs free energy of binding is the most important thermodynamic description of binding, since it determines the stability of any given biological complex, and it has been (and still is) a useful analytical tool for the phenomenological characterization of structure-function relationships.

The typical analysis of calorimetric data involves fitting an appropriate model to the data, i.e. simple single-site or two-site binding model, which yields the binding constant ($K_b$). But often, the system under investigation exhibits a more complex behavior and more sophisticated models must be applied (multiple interacting-site models). The macromolecule may undergo ligand-induced changes, be they conformational adaptations in the binding site or self-association of the receptor, which will contribute to the total free energy of binding. These effects should be evaluated independently with companion methods, i.e. analytical ultracentrifugation, analytical gel filtration chromatography and spectroscopic methods.

In several recent publications it has been proposed to dissect binding free energy into several contributing terms. The total binding free energy contains a contribution typically associated with the formation of secondary and tertiary structure (van der Waals interactions, hydrogen bonding, hydration, conformational entropy), electrostatic and ionization effects, contributions due to conformational transitions, loss of translational and rotational degrees of freedom, and others that must be accounted for on an individual basis (Aqvist et al., 1994; Cummings et al., 1995; Horton & Lewis, 1992; Krystek et al., 1993; Searle et al., 1992; Wallqvist et al., 1995; Williams et al., 1991; Williams et al., 1993). For example, an observed $\Delta G$ can be the
same for an interaction with positive $\Delta S$ and $\Delta H$ (binding dominated by hydrophobic effect) and an interaction with negative $\Delta S$ and $\Delta H$ (when specific interactions dominate). Moreover, interacting systems tend to compensate enthalpic and entropic contributions to $\Delta G$, making binding free energy relative insensitive to changes in the molecular details of the interactions process (Eftink et al., 1983; Lumry & Rajender, 1970; Williams et al., 1993). Thus, consideration of $\Delta H$ and $\Delta S$ are crucial for a detailed understanding of the free energy of binding.

1.8.2. Binding Enthalpy

The observed heat effect of a binding reaction is a global property of the whole system under investigation, reflecting the total heat change in the calorimetric cell upon addition of ligand. On the one hand, the measured heat contains contribution from unspecific heat effects, on the other hand, there might be protonation effects coupled to binding. Therefore it is of outmost importance to determine possible contributions to the intrinsic binding enthalpy and to correct for them. But even corrected heat changes are itself composed of different contributions, which is the reason why the enthalpy is an apparent or observed ($\Delta H_{\text{obs}}$) quantity (Cooper & Johnson, 1994a).

At first appearance, the physical meaning of $\Delta H$ seems to be simple: it represents the changes in non-covalent bond energy occurring during the interaction. This interpretation is too simple to describe observed $\Delta H$. The measured enthalpy must be the result of the formation and breaking of many individual bonds, since it is barely conceivable to form bonds without breaking any, especially in aqueous medium. The enthalpy change of binding reflects the loss of protein-solvent hydrogen bonds and van der Waals interactions, formation of protein-ligand bonds, salt bridges and van der Waals contacts, and solvent reorganization near protein surfaces. These individual components may produce either favorable or unfavorable contributions, and the resultant is likely to be smaller than the specific interactions (Fisher & Singh, 1995).

From calorimetric studies carried out in water and D$_2$O it was concluded that a large part of the observed enthalpy change is due to bulk hydration effect (Chervenak & Toone, 1994; Connelly et al., 1993). Often, water molecules are placed in the complex interface, improving the complementarity of the complex surfaces and
extending H-bond networks. This can make enthalpies more favorable, but is often counterbalanced by an entropic penalty (Bhat et al., 1994; Holdgate et al., 1997; Ladbury, 1996). The role of interfacial water was directly shown by lowering water activity by means of glycerol or other osmolytes. Complexes with a low degree of surface complementarity and no change in hydration are tolerant to osmotic pressure (Goldbaum et al., 1996; Jelesarov & Bosshard, 1994; Kornblatt et al., 1993; Robinson & Sligar, 1993; Xavier et al., 1997).

Besides the unspecific hydration effects, all direct non-covalent bonds at the binding interface contribute to $\Delta H$, actually reflecting the binding enthalpy in a strict sense. The dissection of each non-covalent interaction is very difficult since the net heat effect of a particular bond is the balance between the reaction enthalpy of the ligand to the macromolecule and to the solvent. Moreover, structural alterations at the binding site due to the binding event may contribute to the binding enthalpy. Several mutational approaches have been applied to investigate the energetics of individual bonds: alanine scanning mutagenesis (Pearce et al., 1996), removal of particular H-bonds at the active site (Connelly et al., 1994), construction of double mutant cycles (Frisch et al., 1997). All these approaches suffer from the problem whether there is a direct relation between the change in $\Delta H$ and the removal of a specific contact in the active site. On a theoretical basis it has been argued that decomposition of $\Delta H$ is not possible (Mark & van Gunsteren, 1994), but other favor a dissection into specific contributions (Boresch & Karplus, 1995; Brady & Sharp, 1995).

1.8.3. Binding Entropy

The entropy of binding is directly calculated from $\Delta G$ and $\Delta H$ according to equation (20). In general, it represents all other positive and negative driving forces that contribute to the free energy.

Recently, it has been proposed that the total entropy change associated with binding can be expressed as the sum of several contributing effects. The main contribution upon complex formation is caused by hydration effects. Since the entropy of hydration of polar and apolar groups is large, the burial of water accessible surface area on binding results in solvent release which contributes often large and positive to the total entropy of interaction. Another important, though unfavorable contribution reflects the reduction of rotational degrees of freedom around torsion angles of
protein and ligand side-chains. An additional entropy term accounts for the reduction of the number of particles in solution and their degrees of freedom (Murphy et al., 1995; Murphy et al., 1993; Murphy et al., 1994). A negative entropy change can entail different contributions, and it does not necessarily indicate increased or unchanged hydration interfaces, but a positive entropy change is a strong indication that water molecules have been released from the complex surface (Jelesarov & Bosshard, 1999).

1.8.4. Binding Stoichiometry

The determination of interaction stoichiometries (n) is of central importance for the characterization of binding mechanisms of biological macromolecules. ITC has emerged as an important tool, since it enables high-precision analysis with high reproducibility because of the computer-controlled injection of definite volumes. Assuming that the concentrations of both interacting species are known, the binding stoichiometry can be determined from the molar ratio of the interacting species at the equivalence point. During fitting procedures the parameter n can either be fixed as equal to the number binding sites per macromolecule, or it can be treated as an additional floating parameter that is determined by iterative fitting. There are a number of possible sources which lead to deviation from the expected values of n: high experimental uncertainty of the data set, error in concentration of either the ligand or the macromolecule, unspecific binding, degradation of ligand, low protein quality (unfolded, missfolded).

If these systematic errors can be ruled out and the fitted values for n still deviate significantly from expected values known from additional independent information, they should be reexamined. It has been proposed to fit the data to the hyperbolic equation and then to reanalyze it by means of a double reciprocal plot (Fisher & Singh, 1995). However, although very intriguing, it is recommended to corroborate stoichiometric data obtained by microcalorimetry with additional independent information.

Once the model is verified, ITC can be used as an excellent quality control tool for the analysis of the fractional binding activity of different lots of protein (for example antibodies), for stability testing (freeze-thaw) and more.
1.8.5. Heat Capacity Changes

If ΔH is determined at a range of temperatures (modern ITC instruments allow measurements between 5°C and 70°C), the change in the constant pressure heat capacity, ΔC_p, for an interaction is given by the slope of the linear regression analysis of ΔH_{obs} plotted versus temperature. Often ΔC_p does not depend on temperature within the small physiological temperature range, although several publications reported weak or strong dependencies (Bruzzese & Connelly, 1997; Ferrari & Lohman, 1994).

For binding reaction, ΔC_p is almost always negative when the complexed state of the macromolecule is taken as the reference state. Its origin lies in the fact that there is strong correlation between ΔC_p and the surface area buried on forming a complex (Gomez & Freire, 1995; Gomez et al., 1995; Murphy et al., 1992; Murphy & Freire, 1992; Spolar et al., 1992). It has been shown that the removal of protein surface area from contact with solvent results in a large negative ΔC_p. The basis of this observation is the different behavior of solvent on the surface of a macromolecule and to that in the bulk, particularly with respect to water molecules interacting with hydrophobic surfaces. This means that for any process, in which water is released from the surface, ΔC_p will be substantial and would be proportional to the amount of surface involved.

Through this correlation of ΔC_p and burial of surface area, the heat capacity provides a link between thermodynamic data and structural information of macromolecules. Therefore in the last years, there has been considerable progress in the parameterization of all thermodynamic parameters and the predictions (Baker & Murphy, 1998; Luque & Freire, 1998).

1.9. Prediction of Binding Energetics

A long standing goal of biophysical chemistry is the prediction of binding energetics from the 3D-structure of protein-ligand complexes, and it is a key element in the field of structure based drug design. The rapidly increasing availability of high resolution protein structures from X-ray crystallography and nuclear magnetic resonance (NMR) opened the field to combine structural information with thermodynamic data of the binding process.
In the past few years, considerable progress has been made in characterizing molecular aspects of protein-ligand interactions by means of thermodynamic data. It has been shown that the major polar and apolar contributions to the enthalpy, entropy and heat capacity changes for protein folding and unfolding can be described in terms of changes in solvent-accessible polar and apolar surface area (Makhatadze & Privalov, 1993; Murphy et al., 1992; Murphy & Freire, 1992; Privalov & Gill, 1988; Privalov & Makhatadze, 1993; Spolar et al., 1989; Spolar et al., 1992; Xie & Freire, 1994).

The empirical parameterization based on calculation of changes in solvent accessible surface areas was first applied to the prediction of protein folding energetics (Makhatadze & Privalov, 1993; Murphy et al., 1992; Murphy & Freire, 1992; Privalov & Makhatadze, 1993; Spolar et al., 1992; Xie & Freire, 1994). Since the atomic interactions involved in associations reactions are similar (Janin & Chothia, 1990), it was suggested to apply this approach to protein-protein interactions, peptide binding to proteins, and to small ligand binding. The parameterization has now reached the state in which accurate prediction of protein folding energetics and binding energetics is possible (Baker & Murphy, 1997; Burrows et al., 1994; Gomez & Freire, 1995; Murphy et al., 1993).

1.9.1. Solvent Accessible Surface Areas

According to Lee and Richards (Lee & Richards, 1971), the solvent-accessible surface area (ASA) is defined as the surface traced out by the center of a solvent probe (frequently taken as a sphere with a radius of 1.4 Å) as it moves over the surface of the protein.

There are a number of implementations described and used to determine ASA (Hubbard & Thornton, 1996; Lee & Richards, 1971), and it is very important to recognize that each implementation yields slightly different results. The original description of the predicting parameters are based on the Lee and Richards algorithm as implemented in the program ACCESS, using a probe radius of 1.4 Å and a slice width of 0.25 Å, but there exist reparameterized values for other implementations. When performing calculations, it must be assured that the appropriate parameters are used, because they are dependent on the algorithm used.
1.9.2. Calculation of Thermodynamic Parameters

In general, changes in solvent-accessible surface area (ΔASA) are determined as the difference of ASA of the final state and of the initial state. For a molecular interaction process, this is the difference between the ASA of the complex and the sum of the ASA of the macromolecule and the ligand, resulting in negative values of ΔASA. ΔASA is further subdivided into nonpolar and polar contributions by simply defining which atoms take part in the surface. Oxygen, nitrogen and sulfur are treated as polar and all carbon atoms as apolar. If structured water molecules take part in the binding interface, they must be accounted for, since their presence will contribute to the amount and type of surface area buried (Baker & Murphy, 1998).

In the ideal case, structural information will be available for the complex and for both interacting species, i.e. the free macromolecule and the ligand. This will account for any structurally defined conformational differences that occur on binding. If there are no conformational changes linked with the association (rigid body binding), it is good practice to extract the free state of the protein by removing the coordinates of the ligand from the complex. However, the assumption of rigid body binding must be verified by other independent methods because any structural rearrangement will contribute to the energetics. For the case where binding is linked to order/disorder of particular regions which are not defined in the coordinate file of the macromolecule, it may be necessary to add a model of the region. This is especially found in protein-

![Fig. 6: Schematic representation of solvent-accessible surface area (ASA). (A) solvent probe is rolled over the structure of a complex. The ASA is defined as the surface traced out by the center of the probe (indicated as solid line). (B) The ASA of the two molecules forming the complex. When the complex is formed, the surface of the shaded regions is buried. The change in solvent-accessible surface area (ΔASA) upon binding is the surface area of the complex minus the sum of the surface areas of both molecules in B (from (Baker & Murphy, 1998)).](image-url)
peptide interactions, for which it is necessary to define a solution structure of the free peptide (Gomez & Freire, 1995; Murphy et al., 1993).

1.9.3. Calculation of $\Delta C_p$

The largest contribution to the $\Delta C_p$ for a binding process arises from dehydration of protein and ligand surface with negative contributions due to burial of apolar surfaces ($\Delta A SA_{np}$) and positive contributions due to burial of polar surfaces ($\Delta A SA_{p}$). From studies of the dissolution of solid model compounds, the following relationship has been proposed (Murphy et al., 1992; Murphy & Freire, 1992):

$$\Delta C_p = \Delta C_{np} \Delta A S A_{np} + \Delta C_p \Delta A S A_{p}$$  \hspace{1cm} (36)

The parameters $\Delta C_{np}$ and $\Delta C_p$ that are suitable for calculations based on ASAs determined by the program NACCESS is given in Table 1. Equation (36) is used to predict the $\Delta C_p$ for ligand binding.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>ACCESS (Presnell)$^6$</th>
<th>ACCESS (Richards)$^7$</th>
<th>NACCESS (Hubbard)$^8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Delta C_{np}$</td>
<td>0.45</td>
<td>0.36</td>
<td>0.43</td>
</tr>
<tr>
<td>$\Delta C_p$</td>
<td>-0.26</td>
<td>-0.25</td>
<td>-0.26</td>
</tr>
<tr>
<td>$\Delta H_{np}$</td>
<td>-8.43</td>
<td>-3.63</td>
<td>-7.26</td>
</tr>
<tr>
<td>$\Delta H_p$</td>
<td>31.29</td>
<td>23.75</td>
<td>29.14</td>
</tr>
</tbody>
</table>

$^6$Values from (Baker & Murphy, 1998); $^7$(Presnell, ); $^8$from (Lee & Richards, 1971); $^9$(Hubbard & Thornton, 1996).

1.9.4. Calculation of $\Delta H$

The change in enthalpy of binding reflects the loss of protein-solvent hydrogen bonds, van der Waals interactions, and formation of protein-ligand bonds, salt bridges and van der Waals contacts, and solvent reorganization near protein surfaces. $\Delta H$ is calculated with reference to the temperature at which the apolar contribution is assumed to be zero ($T_{H^*}$). The value of $T_{H^*}$ is the enthalpy convergence temperature, which is obtained from protein unfolding studies and is
determined to be 100°C (373 K) (Murphy & Gill, 1991). Thus the enthalpy change is calculated from

\[ \Delta H = \Delta H^* + \Delta C_p(T - T_{h^*}) \]  

(37)

where \( \Delta H^* \) is the polar contribution to \( \Delta H \) at \( T_{h^*} \). The value for \( \Delta H^* \) is directly proportional to the burial of polar surface area (Murphy et al., 1992; Murphy & Freire, 1992) and is described by

\[ \Delta H^* = 35(\pm 6)\Delta A_{SA_p} \]  

(38)

where \( \Delta A_{SA_p} \) is the change in polar accessible surface area and is negative for a binding reaction.

Above dissection contains a linear extrapolation of protein unfolding enthalpies as a function of polar and apolar surface area (Murphy et al., 1992; Murphy & Freire, 1992; Xie & Freire, 1994). A regression analysis at the medium unfolding temperature of proteins (60°C) minimizes the extrapolation error and yields the elementary contributions per Å² of apolar (\( \Delta h_{ap} \)) and polar (\( \Delta h_{pol} \)) surface to the enthalpy function at the reference temperature of 60°C (\( \Delta H_{\text{bind}}(60°C) \)):

\[ \Delta H_{\text{bind}}(60°C) = \Delta h_{ap} \Delta A_{SA} + \Delta h_{pol} \Delta A_{SA_p} \]  

(39)

Values for \( \Delta h_{ap} \) and \( \Delta h_{pol} \) are given in Table 1. At any other temperature \( T \), \( \Delta H(T) \) is given by the standard equation

\[ \Delta H_{\text{bind}}(T) = \Delta H_{\text{bind}}(60°C) + \Delta C_p(T - 333.15) \]  

(40)
1.9.5. Calculation of $\Delta S$

The total entropic contributions associated with a binding reaction can be expressed as the sum of three terms (Murphy et al., 1995):

$$\Delta S_{\text{tot}} = \Delta S_{\text{solv}} + \Delta S_{\text{conf}} + \Delta S_{\text{r/t}}$$  \hspace{1cm} (41)

where $\Delta S_{\text{solv}}$ describes the change in entropy resulting from solvent release upon binding, $\Delta S_{\text{conf}}$ is a configurational term reflecting the reduction of rotational degrees of freedom around torsion angles of protein and ligand. $\Delta S_{\text{r/t}}$ entails the loss of translational and rotational degrees of freedom when a complex is formed from two molecules free in solution.

The most important contribution to the entropy change arises from the solvation term ($\Delta S_{\text{solv}}$), primarily due to burial of apolar surface area and is approximated for any temperature ($T$) by following equation:

$$\Delta S_{\text{solv}} = \Delta C_p \ln \left( \frac{T}{T^*_S} \right)$$  \hspace{1cm} (42)

where $T^*_S$ is the temperature at which there is no solvent contribution to the hydrophobic entropy change and is equal to 112°C (385 K) (Baldwin, 1986; Murphy et al., 1990; Murphy et al., 1994).

The translational/rotational entropy term ($\Delta S_{\text{r/t}}$) accounts for the reduction of the number of particles in solution and their degrees of freedom. Very different estimates of the magnitude of $\Delta S_{\text{r/t}}$ have been discussed (Finkelstein & Janin, 1989; Murphy et al., 1993). Empirical and theoretical considerations have suggested that this term contributes -4 to -10 cal K$^{-1}$ mol$^{-1}$ to the overall entropy of a bimolecular binding event (Amzel, 1997; Murphy et al., 1994; Tamura & Privalov, 1997). This value is numerically close to the cratic entropy of -8 cal K$^{-1}$ mol$^{-1}$ (Kauzmann, 1959) even if there is no sound physical reason to equate translational/rotational entropy to the statistical part of the mixing entropy (Gilson et al., 1997; Holtzer, 1995).

Finally, the configurational entropy $\Delta S_{\text{conf}}$ reflects contributions from changes in side-chain conformational entropy as well as all other structural rearrangement of protein.
and ligand induced by complex formation. Since $\Delta S_{\text{tot}}$ is experimentally accessible, and $\Delta S_{\text{solv}}$ and $\Delta S_{\text{nt}}$ can be estimated, we have

$$
\Delta S_{\text{conf}} = \Delta S_{\text{solv}} - \Delta S_{\text{nt}} - \Delta S_{\text{c/s}}
$$

(43)

The contribution to $\Delta S$ from side chains involved in binding can be estimated by considering an average contribution of -4.3 calK$^{-1}$mol$^{-1}$ per residue to $\Delta S_{\text{conf}}$ (Baldwin, 1986; Murphy & Freire, 1995; Murphy et al., 1993). On the assumption that only minor configurational entropic contributions of ligand occur, a rough estimate of the number of amino acids ($X_{\text{res}}$) participating in the interaction is available by

$$
X_{\text{res}} = \frac{\Delta S_{\text{conf}}}{-4.3 \text{calK}^{-1}\text{mol}^{-1}}
$$

(44)

If the amino acid side chains directly participating in the interaction process are known from theoretical or experimental studies, it is possible to calculate each contribution to $\Delta S_{\text{conf}}$ as the sum over the amino acids involved in binding (Creamer & Rose, 1992; Creamer & Rose, 1994; D’Aquino et al., 1996; Lee et al., 1994). A binding process involves two contributions to $\Delta S_{\text{conf}}$: restrictions around side-chain torsion angles and immobilization of the peptide backbone. If binding is not involved in order/disorder transitions, only the side-chain component applies. The side-chain contribution is then calculated by assuming that the entropy is zero when fully buried and scales linearly with ASA, resulting in a maximum value when fully exposed. By applying this model, a term $\Delta S_{\text{bur}}$ must be introduced to account for the buried-to-exposed entropy gain which differs for each side chain. The contribution is then calculated as

$$
\Delta S_{\text{conf}} = \sum_i \frac{\Delta ASA_{\text{sc},i}}{\Delta ASA_{\text{ Ala-X-Ala}}^i} \Delta S_{\text{bur-ex},i}
$$

(45)

where $\Delta ASA_{\text{sc},i}$ is the change in ASA of side chain $i$ on binding, and $\Delta ASA_{\text{ Ala-X-Ala}}^i$ is the ASA of the side chain in an extended Ala-X-Ala tripeptide (Baker & Murphy, 1998). The summation is carried out over all side chains participating in the interface.
Values for $\Delta \text{ASA}_{sc,i}$ are determined from structural data, estimates for $\Delta \text{ASA}_{XAX,i}$ and $\Delta S_{bu>ex,i}$ (Lee et al., 1994) are available, and they have been adapted for different implementations (Baker & Murphy, 1998). These data are presented in Table 2.

For interaction processes involving transitions, two additional terms associated with the backbone entropy ($\Delta S_{bb}$) and with the entropy of going from the exposed/folded to the exposed/unfolded state ($\Delta S_{ex>u}$) must be included (Baker & Murphy, 1998). Estimates of these contributions are also available (D'Aquino et al., 1996), and are summarized in Table 2.

### Table 2: Total side-chain ASA ($\Delta \text{ASA}_{sc}$) of Ala-X-Ala tripeptides used in parameterization for three ASA implementations (units in Å$^2$), and side-chain and backbone conformational entropy values ($\Delta S_{bu>ex}$, $\Delta S_{bb}$, $\Delta S_{ex>u}$; units in cal/K·mol$^{-1}$)

<table>
<thead>
<tr>
<th></th>
<th>ACCESS ACCESS ACCESS</th>
<th>NACCESS ACCESS ACCESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Presnell) (Richards) (Hubbard)</td>
<td></td>
</tr>
<tr>
<td>Ala</td>
<td>52.1</td>
<td>60.4</td>
</tr>
<tr>
<td>Arg</td>
<td>187.9</td>
<td>210.2</td>
</tr>
<tr>
<td>Asn</td>
<td>113.8</td>
<td>123.0</td>
</tr>
<tr>
<td>Asp</td>
<td>102.4</td>
<td>106.5</td>
</tr>
<tr>
<td>Cys$^a$</td>
<td>70.8</td>
<td>69.1</td>
</tr>
<tr>
<td>Cys$'$</td>
<td>91.9</td>
<td>94.5</td>
</tr>
<tr>
<td>Gln</td>
<td>128.7</td>
<td>142.8</td>
</tr>
<tr>
<td>Glu</td>
<td>117.5</td>
<td>135.8</td>
</tr>
<tr>
<td>Gly</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>His</td>
<td>144.3</td>
<td>152.2</td>
</tr>
<tr>
<td>Ile</td>
<td>123.8</td>
<td>138.5</td>
</tr>
<tr>
<td>Leu</td>
<td>134.5</td>
<td>150.6</td>
</tr>
<tr>
<td>Lys</td>
<td>156.9</td>
<td>177.7</td>
</tr>
<tr>
<td>Met</td>
<td>158.5</td>
<td>160.8</td>
</tr>
<tr>
<td>Phe</td>
<td>176.7</td>
<td>179.4</td>
</tr>
<tr>
<td>Pro</td>
<td>90.7</td>
<td>104.8</td>
</tr>
<tr>
<td>Ser</td>
<td>68.6</td>
<td>76.6</td>
</tr>
<tr>
<td>Thr</td>
<td>105.3</td>
<td>112.1</td>
</tr>
<tr>
<td>Trp</td>
<td>222.7</td>
<td>218.7</td>
</tr>
<tr>
<td>Tyr</td>
<td>188.5</td>
<td>196.4</td>
</tr>
<tr>
<td>Val</td>
<td>105.6</td>
<td>118.0</td>
</tr>
</tbody>
</table>

$^a$Values from (Baker & Murphy, 1998); $^b$(Presnell); $^c$from (Lee & Richards, 1971); $^d$(Hubbard & Thornton, 1996); $^e$Values for disulfide-bonded cystine; $^f$Values for free cysteine.
For non-peptide ligands a special empirical parameterization has been proposed to account for changes in conformational degrees of freedom between free and complexed forms of ligand (Bardi et al., 1997). As a first approximation, it is assumed that the conformational entropy will be proportional to the number of rotatable bonds. Since effects from excluded volume increase with the number of atoms for a given number of rotatable bonds, the conformational entropy change for a non-peptide ligand ($\Delta S_{\text{conf, np}}$) is considered to be a linear function of the number of rotatable bonds ($N_{rb}$) and the total number of atoms ($N_{at}$):

$$\Delta S_{\text{conf, np}} = k_1 N_{rb} + k_2 N_{at} \quad (46)$$

To date, the applicability of equation (44) has not been widely used, and it is only tested for the analysis of HIV-1 protease inhibitors (Luque & Freire, 1998). However, for these interactions $k_1$ was found to be -1.76 calK$^{-1}$mol$^{-1}$, whereas $k_2$ equals 0.414 calK$^{-1}$mol$^{-1}$.

### 1.9.6. Linkage Effects

Above calculations apply for systems that do not involve linked equilibria. If binding of a second ligand is coupled to binding of a first one, the contributions of the second equilibrium must be considered. Protonation linkage is a common phenomenon in interaction processes. Its contributions to the binding energetics can be determined experimentally, using a global analysis of experimental data as a function of pH, temperature and buffer ionization enthalpy (Baker & Murphy, 1996; Baker & Murphy, 1997).

### 1.10. Thermodynamics and Rational Drug Design

The rapidly increasing availability of high-resolution protein structures has opened the possibility to use structural information in the design of new drugs. The central problem of structure-based design studies is the understanding of the features dictating the energetics of the interaction of a ligand with a macromolecule, i.e. the accurate prediction of the Gibbs free energy that determines the binding affinity.
1.10.1. The Thermodynamic Approach

The prediction of binding energetics is greatly complicated by the effects of enthalpy-entropy compensation (Dunitz, 1995; Gilli et al., 1994) which means that an increase in $\Delta H$ does not contribute to the binding affinity, as the improvement is only achieved by a compensation cost in the $T\Delta S$ term. The sheer number of entropic and enthalpic effects contributing to the observed $\Delta G$ makes it difficult to rationalize and predict binding affinities. Additional energetic effects will arise from any differences in ligand conformations in the free state in solution and bound to the macromolecule. Moreover, water molecules play an important role in adapting the binding pocket to different ligands.

The strength of the thermodynamic approach is that it has become possible to dissect the contributing forces ($\Delta H$, $\Delta S$, $\Delta C_p$) which make up the free binding energy of the interaction by direct calorimetric measurements, yielding effective energetic contributions, including all interactions that are found in the system (protein-, ligand, and solvent interactions).

1.10.2. Current Status of the Thermodynamic Approach

Until recently, thermodynamic data have not played an important role in molecular design, since no theoretical framework that relates structural and thermodynamic data has been established. The situation has changed with the realization that changes in solvent-accessible surface area ($\Delta$ASA) are related to the thermodynamic parameters. The semi-empirically derived set of structural energetic parameters (summarized in Table 3), based on the parameterization of $\Delta G$, $\Delta H$, $\Delta S$, and $\Delta C_p$ in terms of $\Delta$ASA, is a promising thermodynamic approach to drug design.

It is possible now to predict the energetics of folding of a globular protein (Makhatadze & Privalov, 1993; Murphy et al., 1992; Murphy & Freire, 1992; Privalov & Makhatadze, 1993; Spolar et al., 1992; Xie & Freire, 1994) with an accuracy within 9% to 12% (Baker & Murphy, 1998). The parameterization has reached the state in which accurate predictions of binding energetics is possible as well, as shown for peptide-protein and protein-protein association (Baker & Murphy, 1997; Burrows et al., 1994; Gomez & Freire, 1995; Murphy et al., 1993), and also for nonpeptide ligand-protein interactions (Bardi et al., 1997). However, the thermodynamic approach is still in its infancy, as there are no comprehensive reports documenting
the design of a high-affinity drug using a "thermodynamic-directed rational drug-design" approach.

1.10.3. Outlook

The structural thermodynamic approach provides a set of promising new tools for the development of rational strategies for drug design. However, current parameterization is still based on protein unfolding data, and there is need for more thermodynamic and structural characterization of protein-ligand interactions, including not only $\Delta H$, $\Delta S$, and $\Delta C_p$, but also thorough investigation of phenomena linked to binding, i.e. protonation, ion binding, conformational changes. As the research in this field makes progress, the structural thermodynamic database is expanded and the set of energetic parameters can be refined to the point that will allow accurate prediction of effects of small molecular changes. Thus, this will allow a much easier assessment of better binding drugs.

1.11. References


PART 1: Isothermal Titration Calorimetry


Presnell, S. R. ACCESS 3.1. edit. ZymoGenetics, Seattle, WA.


Sigurskjold, B. W., Berland, C. R. & Svensson, B. (1994). Thermodynamics of inhibitor binding to the catalytic site of glucoamylase from Aspergillus niger
PART 1: Isothermal Titration Calorimetry


Herpes Virus Type 1 Thymidine Kinase

2.1. Herpes Viruses

Herpes viruses belong to the family Herpesviridae which is divided in three subfamilies. The alphaherpesviridae include herpes simplex virus type 1 and 2 (HSV1 and HSV2), varicella zoster virus, bovine herpes virus, marmoset herpes virus, equine herpes virus, pseudorabies virus and feline herpes virus. The betaherpesviridae entails the cytomegalovirus, and the gammaherpesviridae includes Eppstein-Barr virus, herpes virus saimiri and Marek's disease herpes virus. Herpes viruses are coated, species-specific DNA viruses, carrying highly evolved and extremely specialized genes exclusively coding for virus specific proteins. Infections with herpes viruses can cause severe and recurrent diseases in animals and humans, which requires effective antiviral therapy.

2.1.1. Herpes Simplex Virus Infections

Herpes viruses are endemic in most animals, and it has been estimated that 60 to 95% of the world population is infected by members of the human herpes family, especially by HSV1 (WHO, 1985). Typically, the first HSV1 infection occurs in childhood by way of epithelial surfaces in the oral cavity, manifesting as gingivostomatitis. The virus enters the nerve endings innervating these tissues and enters the peripheral nervous system by retrograde transport to neurons in the trigeminal ganglion, where the viral latency is established (Roizman & Batterson, 1985; Spruance, 1995). In contrast to other viruses, HSV1 does not integrate its DNA into the host DNA, but stays dormant in the nerve cell until it is reactivated. On reactivation it causes a variety of diseases like the common fever blisters, but also genital skin lesions, blindness and encephalitis.
2.1.2. HSV1 Latency

Herpes viruses encode a variety of proteins involved in the replication of viral DNA, of which the viral DNA polymerase is essential for virus replication (Challberg & Kelly, 1989). In contrast, the viral thymidine kinases (TK) have been shown to be dispensable for virus replication in cell culture. Early reports describe a crucial role of TK in reactivation of the virus from latency. Mutations such as deletions leading to complete inactive TK (TK⁻ mutations) severely impaired acute ganglionic replication and reactivation in mice (Coen et al., 1989a; Coen et al., 1989b; Efstathiou et al., 1989; Jacobson et al., 1993), and TK inhibitors are able to suppress the reactivation in mice (Gebhardt et al., 1996; Jacobson et al., 1995; Leib et al., 1990). Treatment with acyclovir, a HSV TK specific nucleoside analog, significantly reduces the rate of recurrences of genital and labial herpes (de Ruiter & Thin, 1994).

More recently, a publication reported mutations resulting in truly TK⁻ virus strains that were still able to reactivate from latency (Horsburgh et al., 1998). It seems that HSV1 consists of a genetic background that permits reactivation from latency independent of the thymidine kinase pathway, yet the underlying mechanism remains to be discovered.

However, cells of the neuronal system contain little or no cellular TK, making the formation of dTMP the rate limiting step for virus reactivation. This was shown by constructing HSV either encoding for their own TK, for thymidylate kinase (TmpK), for deoxycytidine kinase (dCK) and for cellular TK. Cellular TK was found to be sufficient for reactivation, TmpK led to partially competent viruses, and dCK exhibited a TK⁻ phenotype which failed to reactivate from latency (Chen et al., 1998).

Taken together, TK appears still to be essential for reactivation from latency, but the precise mechanism remains unclear.

2.2. Thymidine Kinases

Thymidine kinases (EC 2.7.1.21) are highly conserved enzymes that occur in nearly all organisms. They catalyze the phosphorylation of thymidine (dT) to thymidine monophosphate (dTMP) in presence of magnesium cations by transferring the γ-phosphate group of adenosine triphosphate (ATP) to the 5'-OH group of dT. The systematically correct name for TK is ATP-thymidine-5'-phospho-transferase. The product is subsequently rephosphorylated by cellular thymidylate kinase and
nucleoside diphosphate kinases to finally build the triphosphate (dTTP) used for DNA synthesis.

The cell does not depend on TK activity to form dTMP, since it is provided by the de novo synthesis pathway through methylation of deoxyuridine monophosphate (dUMP) in presence of methylene tetrahydrofolate by thymidylate synthetase (Fig.1). TK is therefore designated as a salvage pathway enzyme, recycling endogenous dT released from metabolic decomposition or resorbed from nutrition.

2.2.1. Cytosolic Thymidine Kinase

This type of enzyme is mainly found in fetal tissues, especially in the liver. Therefore, it is also termed TK1 or fetal TK. Prior to birth the concentration declines sharply. In adults, high concentrations are only found in dividing cells (Kit, 1976), in the serum of patients suffering certain neoplastic diseases, and concomitant with a number of viral infections.

TK activity levels are tightly regulated by both the growth state and the cell cycle position of the cell, with dividing cells revealing high activities and resting cells having low or undetectable activities (Machovich & Greengard, 1972). In mitogenically stimulated cells, TK activity remains low in the G1 phase, increases dramatically as the cell enters the S phase, and remains elevated during S and G2 phase (Carozza & Conrad, 1994; Kauffman & Kelly, 1991; Sherley & Kelly, 1988).

2.2.2. Mitochondrial Thymidine Kinase

This type of TK, also named TK2, is localized in the mitochondrial matrix and exhibits low but constant levels of TK activity during the entire cell cycle. It is transcribed in the cytosol and subsequently translocated, most probably guided by a signaling peptide that is cleaved off after having reached the mitochondrial matrix. TK2 is not involved in cell growth, but is used for thymidine nucleotide synthesis of mitochondrial DNA, since dTMP is negatively charged and exhibits extremely poor penetration of the mitochondrial membrane.

TK2 is able to phosphorylate dC, dU, AZT and fluoroarabinofuranosyl-iodouracil (FAIU), and it is therefore thought to be responsible for mitochondrial toxicity (Johansson & Karlsson, 1997; Lewis & Dalakas, 1995).
Fig. 1: Biosynthesis pathways of thymidine triphosphate

- "salvage pathway"
  - thymidine (dT)
  - ATP to ADP via thymidine kinase

- "de novo synthesis"
  - 2'-deoxyuridine-5'-monophosphate (dUMP)
  - thymidylate synthetase
  - 5,10-methylene-tetrahydrofolate reductase
  - 7,8-dihydrofolate
  - thymidine-5'-monophosphate (dTMP)
  - ATP to ADP via thymidylate kinase
  - dTDP
  - ATP to ADP via nucleoside diphosphate kinase
  - dTTP
  - incorporated in DNA by DNA polymerase
2.3. Structure and Function of HSV1 Thymidine kinase

2.3.1. Crystal Structure of HSV1 TK

HSV1 TK is a homodimer with 376 residues per subunit and is an $\alpha/\beta$ protein consisting of a central five stranded parallel $\beta$-sheet flanked on either side by $\alpha$-helices (Fig. 2). The crystal structure (Wild et al., 1995; Wild et al., 1997) comprises residues 34 to 376, whereas residues 34-45, 150-152 and 265-279 are missing because of insufficient electron density most probably due to segmental mobility. To date, there are structures known in complex with the natural substrates, substrate analogs and cofactor (Bennett et al., 1999; Brown et al., 1995; Champness et al., 1998; Wild et al., 1995; Wild et al., 1997).

HSV1 TK shows the classical mononucleotide (NMP) binding fold (Schulz, 1992) with the Walker A-motif (Walker et al., 1982) forming the characteristic phosphate-binding loop (Saraste et al., 1990). The precise ordering of the central five $\beta$-sheets classifies HSV1 TK as belonging to the same structural family as the NMP-kinases for which structures are known for adenylate kinase (ADK) (Dreusicke et al., 1988; Schlauderer & Schulz, 1996), guanylate kinase (Stehle & Schulz, 1990), uridylate kinase (Muller-Dieckmann & Schulz, 1994) and bacteriophage T4 deoxynucleotide kinase (Vonrhein et al., 1995). For homology analysis of NMP-kinases the central five $\beta$-sheets has been dubbed CORE domain whereas the remaining domains are the so-called LID domain and NMP-bind domain (Vonrhein et al., 1995). Superposition of the backbones of HSV1 TK and ADK reveals substantial similarity with respect to the CORE domain. Major differences occur in the topology of the nucleoside/NMP-bind domain, which is formed by extensive insertions, and in the size of the LID domain. The LID domain of HSV1 TK consists of only eight residues reminiscent of small variants of NMP-kinases. Taken together, the largest differences are due to the terminal residues (Wild et al., 1997). Interestingly, many of these differences appear to be connected to the dimer interface that is similar to the dimer interface of the dimeric thymidylate kinase (Lavie et al., 1997), but differs significantly of the dimeric bacteriophage T4 deoxynucleotide kinase (Teplyakov et al., 1996).
2.3.2. Quaternary Structure

In contrast to TK1, which appears to be active as a tetramer (Munch-Petersen et al., 1993), HSV1 TK is found to be dimeric in solution (Fetzer et al., 1994) and in the crystal state (Wild et al., 1995; Wild et al., 1997). The solvent accessible surface area buried in the HSV1 TK dimer is 1800 Å² per subunit, i.e. 14% of the total subunit surface (Wild et al., 1997), and is, with 65% hydrophobic residues, exceptionally non-polar. The large interface is in agreement with values from other dimers of comparable molecular weight, but differs grossly from the small interface of 900 Å² of bacteriophage T4 deoxynucleotide kinase.

The interface is almost planar with only one substantial protrusion formed by W310 anchoring in a hollow of the other subunit. The dimer interface of HSV1 TK consists of mainly hydrophobic helix/helix interactions with α4 and α6 forming the center and is completed by helix α2, parts of loop α2-α3, loop α12-α13 and the C-terminal end of helix α15. Most of these elements are part of the NMP bind domain, but there are also contributions from the CORE domain and from the additional mobile segment of 72 residues to the interface.
2.3.3. Thymidine Binding Site

The substrate binding site is located in the so called NMP\textsubscript{bind} domain which accommodates either dT or dTMP (Fig.3), being responsible for the additional thymidylate kinase function of TK (Chen \textit{et al.}, 1979a). The carboxamide of Q125 is interacting via H-bonds with N3 and O4 of the thymine ring of the substrate, the O2 of dT is linked to R176 by two water-mediated H-bonds. The thymine ring is held between M128 and Y172, forming a sandwich-like complex. The ribose moiety is interacting with Y101 and E225.

The nucleoside/nucleotide binding site is deeply buried in the protein interior, in agreement with the low $K_M$ value of 0.2 \(\mu\text{M}\) for dT (Fetzer \textit{et al.}, 1993; Kussmann-Gerber \textit{et al.}, 1999; Michael \textit{et al.}, 1995). Although tightly bound, dT does not completely fill its binding pocket. It leaves a 35 \(\text{Å}^3\) void close to its C5 position.

![Diagram](image-url)

*Fig.3: Representation of the nucleoside/nucleotide binding site of HSV1 TK with bound dT (from PDB entry 2VTI). Amino acids directly involved in substrate binding are labeled and shown as capped sticks. Water molecules are presented as small spheres and hydrogen-bonds are displayed as dashed lines. Tubes represent the secondary structure of the enzyme. The LID domain as well as the glycine-rich loop (P-loop) are indicated. The figure has been prepared using SYBYL V.6.3. (Tripos Associates).*
2.3.4. ATP Binding Site

A crucial part of the ATP binding site is the glycine rich motif connecting \( \beta \)-strand \( \beta 1 \) with helix \( \alpha 1 \). It contains the sequence fingerprint \( ^{56}GXXGXGKT^{63} \) (P-loop (Saraste et al., 1990)), forming a giant anion hole that accommodates the \( \beta \)- and \( \gamma \)-phosphoryl group of ATP in a number of ATP-binding proteins (Schulz, 1992). The binding site is much more solvent exposed (Fig.4), and the adenosine moiety is only weakly bound as reflected in the high \( K_M \) value of 16 to 70 \( \mu \text{M} \) (Chen et al., 1979a; Kussmann-Gerber et al., 1999). The adenine moiety is bound by one H-bond through Q331 and flanked by R216. Besides the glycine rich loop, residues R212, R220, and R222 are involved in ATP binding. The bottom of the binding site includes the sequence motif FD\(_{162}\)RH, which is the so-called Mg\(_{2+}\)-NTP binding motif, originally identified by Walker (Walker et al., 1982). It is highly conserved among thymidine kinases (Remond et al., 1995), and D162 is assumed to take part in binding Mg\(_{2+}\) together with T63 and three water molecules (Kussmann-Gerber et al., 1998). As deduced from ADK, K62 is essential for phosphoryl transfer during catalysis (Muller & Schulz, 1992).

2.3.5. Substrate Specificity

In contrast to the highly specific TK1, HSV1 TK is a multifunctional enzyme and exhibits a broad substrate diversity. It shows additional deoxycytidine kinase and thymidylate kinase activities and converts a broad spectrum of pyrimidine as well as purine analogs (Fig.5). Moreover, HSV1 TK displays low stereoselective and stereochemical demands, since it accepts modified ribose moieties and phosphorylates even acyclic side chains as well as the L-dT instead of the natural occurring D-dT. The preferred phosphate donor is ATP, but HSV1 TK shows high affinities for cytidine triphosphate (CTP), uridine triphosphate (UTP) and guanosine triphosphate (GTP) as well.

The substrate binding pathway is not yet clear. It is generally accepted that NMP kinases follow a random bi bi mechanism (Roads & Lowenstein, 1968; Yan & Tsai, 1999). Kinetic analysis of HSV1 TK revealed different results. On the one hand a random bi-bi mechanism is proposed (Chen et al., 1979b), and on the other hand a preferred, sequential binding order has been described (Gerber, 1997; Kussmann-Gerber et al., 1999). From the structural point of view it would be assumed that dT
must bind first, since its binding site is localized deeply buried in the protein, with ATP placed in front like a plug. This view is further corroborated by mutagenesis studies and thermodynamic measurements (this work).

2.3.6. Structural Rearrangement due to Binding

Catalysis by kinases is known to be accompanied by large conformational changes that occur during substrate and cofactor binding. Since in most cases kinases transfer phosphoryl groups to hydroxyl groups, they need to protect their active sites from the surrounding water to avoid ATP hydrolysis (ATPase function). This is achieved by induced-fit movements that exclude strongly competing water efficiently from the reaction center of the kinases, as it is observed for several kinases. A detailed crystallographic analysis of conformational changes of ADK showed a rigid
Fig. 5: Formula of various HSV1 TK substrates and analogs thereof.
body rotation for the NMP-binding domain and a LID domain movement accompanied by large hinge bending rotation (Fig.6) (Muller et al., 1996). With respect to the NMP kinases, structures are known for complexes with NMP or with ATP, either separated or in form of a covalent connection of two nucleotides.

To date, there is no structure known with natural ATP alone, with the only exception of a mutated ADK in complex with an ATP derivative (ADPCF₂P), showing the conformational changes related to ATP binding (Schlauderer et al., 1996).

All structures of HSV1 TK solved to date are ternary complexes with natural and non-natural substrates, inhibitors and cofactors or sulfate ions (mimicking the β-phosphate of ATP), i.e. they correspond to the closed conformation of NMP kinases (Vonrhein et al., 1995), whereas structural data corresponding to the native apo form has not yet been reported. Therefore, the exact mechanism of substrate and cofactor binding of HSV1 TK with respect to possible structural rearrangement remains to be solved.
2.4. Medicinal Aspects of HSV1 TK

2.4.1. HSV1 TK as Drug Target

Chemotherapy of viral infections is difficult, since viruses possess no metabolism, and the use of antimetabolites is limited to nucleoside analogs. However, virus replication is based on host cell enzymes, thus infected and non-infected cells are equally damaged. A certain virus selectivity is only given by the DNA synthesis rate which is elevated in infected cells. Several DNA viruses encode for their own enzymes which are essential for replication, i.e. TK and DNA polymerase.

First treatment of herpesvirus infections started at the end of the 1970's with the introduction of acyclovir, and HSV1 TK was soon identified as the target enzyme for the activation of selective and effective therapeutic prodrugs (Fyfe et al., 1978). The most widely used are the purine analogs acyclovir (ACV) and gancyclovir (GCV). Both virustatics are activated by HSV1 TK to their monophosphates (Fyfe et al., 1978), further phosphorylated to their triphosphates by cellular enzymes (Miller & Miller, 1980) and incorporated into DNA, resulting in chain termination and formation of dead-end complexes with viral DNA polymerase (Reardon & Spector, 1989). The difference in substrate specificity between TK1 and HSV1 TK establishes the molecular basis for this selective antiviral therapy.

Although TK is not generally required for virus replication, it is required for full pathogenicity in animal models and for reactivation from latency (Coen et al., 1989b). Consequently, a number of groups have synthesized compounds that specifically inhibit HSV1 TK (Gebhardt et al., 1996; Hildebrand et al., 1990; Martin et al., 1989; Martin et al., 1998; Martin et al., 1983), without being phosphorylated, and which are able to diminish reactivation of viruses. Since non-neuronal tissues exhibit cellular TK activity, specific viral TK inhibitors will have little or no effect on the host cell (Gebhardt et al., 1996; Leib et al., 1990). This class of antivirals acts towards prophylaxis of infection and reactivation of latency, but will be ineffective once the virus is reactivated from latency because non-neuronal tissues will provide dTMP for viral DNA synthesis.
2.4.2. HSV1 TK and Gene Therapy

With the development of biotechnological methods for gene manipulation, it became feasible to insert foreign genes into viral or bacterial gene vectors, and also transfer of new genes into a human cell. The general concept of suicide gene therapy is based on introducing a new metabolic property into a target cell in order to make it susceptible to therapeutic drugs, that will be activated only in these cells. HSV1 TK features some unique properties which are exploited for several different gene therapeutical approaches.

For the treatment of neoplastic diseases, the so-called virus-directed enzyme/prodrug therapy has been developed (Huber et al., 1991). This approach transduces the tk gene into a population of cells, conferring a lethal sensitivity to the a drug. The major problem of this approach is the specific targeting of malignant cells, selective gene delivery, specific gene expression, specific gene activity and, if possible, specific drug action. Although recent progress, these problems remain a formidable challenge (Dachs et al., 1997).

To date, a number of cancer types are under evaluation for this kind of treatment, i.e. cell lung cancer (Kumagai et al., 1996), ovarian cancer (Tong et al., 1999; Tong et al., 1997; Tong et al., 1998), uterine adenomcarcinoma (Kunishige et al., 1999), gastrointestinal tumors (Yang et al., 1998), and colon cancer (Wildner et al., 1999). Further applications might be treatment of retinoblastoma (Hurwitz et al., 1999), osteosarcoma (Charissoux et al., 1999), or pituitary tumors (Lee et al., 1999). Even in the field of AIDS therapy, the gene therapy approach has been adopted (Caruso & Klatzmann, 1994; Christians et al., 1999; Guettari et al., 1997).

2.4.3. Phosphorylation of New DNA Building Blocks

The exceptional substrate diversity of HSV1 TK can be exploited for the phosphorylation of non standard nucleosides. Since forming the monophosphate of such compounds is the rate limiting step, TK can be used as a device to for the first phosphorylation step. Subsequent phosphorylation to the triphosphate would lead to new DNA building blocks, thus extending the genetic alphabet. For such purpose, HSV1 TK may be engineered to allow phosphorylation, if wild type enzyme is not active (Pilger, 1999).
2.5. References


PART 1: Herpes Virus Type 1 Thymidine Kinase


3.1. Citrate Metabolism

Citrate is an ubiquitous natural compound that occurs in all living cells. It is therefore not surprising that many bacteria are able to utilize citrate as carbon and energy source. Under aerobic conditions citrate is usually metabolized via the tricarboxylic acid cycle (TCA cycle). Citrate fermentation under oxic conditions requires only an additional citrate carrier for citrate uptake and, if citrate is the only substrate, a set of enzymes that allow formation of acetyl-CoA and phosphoenol-pyruvate from TCA cycle intermediates.

Anaerobic conditions requires a different pathway for citrate dissimilation, since the TCA cycle is usually not operative due to suppressed synthesis of the key enzyme 2-oxoglutarate dehydrogenase (Park et al., 1997). A number of different pathways for citrate fermentation have been discovered so date (Antranikian & Giffhorn, 1987), of which the only enzyme in common is citrate lyase. This enzyme catalyzes the cleavage of citrate into acetate and oxalacetate. Consequently, its presence is indicative for anaerobic citrate fermentation capabilities.

3.1.2. Citrate Fermentation in Klebsiella Pneumoniae

Klebsiella pneumoniae belongs to the family of Enterobacteriaceae and is able to utilize citrate as the sole carbon and energy source under anaerobic conditions. A prerequisite for citrate fermentation is its uptake by carriers, of which at least three are known (Schwarz & Oesterhelt, 1985). Two have been analyzed genetically and biochemically.

During aerobic growth, citrate uptake by the CitH carrier protein is sodium independent and is driven mainly by a transmembrane chemical gradient of H⁺ and
by a transmembrane electrical gradient. The divalent anion H-citrate\textsuperscript{2-} is assumed to be the transported species, and a symport with three protons has been proposed (Van der Rest \textit{et al.}, 1991; Van der Rest \textit{et al.}, 1990). Under fermentative growth, citrate uptake has been shown to be dependent on a chemical gradient of Na\textsuperscript{+} and/or H\textsuperscript{+} across the cytoplasmic membrane. The process is mediated by the dimeric CitS carrier protein (Pos \textit{et al.}, 1994; Pos & Dimroth, 1996; Van der Rest \textit{et al.}, 1992).

The first step of citrate degradation of citrate is cleavage by citrate lyase to form acetate and oxalacetate (Fig.1). Latter is decarboxylated to pyruvate by the membrane bound oxalacetate decarboxylase Na\textsuperscript{+} pump (Bott, 1997; Dimroth, 1988), which generates an transmembrane electrochemical Na\textsuperscript{+} gradient. Pyruvate is further degraded by pyruvate formate lyase to yield acetyl-CoA, which is converted to acetyl-phosphate by phosphotransacetylase. Finally, acetate kinase is used to form ATP from ADP and acetyl-phosphate. The pathway allows the synthesis of 1 mol of ATP by substrate level phosphorylation in the acetate kinase reaction, and 0.3 mol ATP per mol of citrate by converting the electrochemical Na\textsuperscript{+} gradient established by oxalacetate decarboxylase (Bott, 1997). These are the only reactions that provide ATP during fermentative growth on citrate under anaerobic conditions.
3.2. Regulation of Citrate Fermentation

The environmental conditions of microorganisms are often subject to rapid changes, and microorganisms must be able to detect and adapt to these changes. Bacteria monitor their surroundings either directly by membrane-bound sensors for indirectly by intracellular receptors, detecting fluctuations in the cytoplasmic metabolite concentrations that result from extracellular changes. Once a stimulus is detected, a signaling cascade is initiated and leads to altered gene expression, enzymatic activities, chemotactic behavior, and more. A number of transducing systems have been found, i.e. the methyl-accepting chemotaxis proteins, the phosphoenolpyruvate:carbohydrate phosphotransferase systems, and the two-component regulatory systems (Goudreau & Stock, 1998; Hazelbauer et al., 1993; Postma et al., 1993).

3.2.1. Two-Component Regulatory Systems

Two component systems are formed by ubiquitous proteins found in eubacteria, archaea and several eukaryotic organisms (Posas et al., 1998; Swanson et al., 1994) and are involved in the regulation of citrate fermentation. A prototypical two-component regulatory system consists of a sensor kinase and a response regulator, representing the signal-communicating modules. Either of these two components contains at least two functional domains and form the backbone of the signal transduction pathway. Information transduction occurs by direct protein/protein interaction and by protein phosphorylation (Appleby et al., 1996; Goudreau & Stock, 1998).

Extracellular changes are detected by periplasmic domains of the sensor kinases. In presence of the signal, the homodimeric sensor kinase autophosphorylates a conserved histidine residue of the transmitter domain in trans, using ATP as phosphoryl donor. The kinase function is included in the transmitter domain and is generally 240 amino acids in length (Stock et al., 1995). From the phosphorylated transmitter domain the signal is passed onto a conserved aspartate residue of the response regulator by means of a reversible Mg$^{2+}$ dependent phosphotransfer reaction. The receiver domain is about 120 amino acids in length and harbors the phosphotransfer function (Wanner, 1995). The phosphorylated receiver domain activates the second domain of the regulatory protein, leading to the response
according to the type of regulatory system, i.e. binding to a target DNA sequence or methylesterase activity. Inactivation is achieved either by the response regulator through its autophosphatase activity, or by the cognate sensor kinase, or by an alternative protein with autophosphatase activity.

3.2.2. Regulation of Citrate Fermentation of *Klebsiella Pneumoniae*

In an early study it was noticed that the responsible enzymes in citrate fermentation are induced under anaerobic conditions in presence of citrate, and that they are subject to catabolite repression (Dagley & Dawes, 1953). Recent investigations of the genes required for expression of the Na⁺ dependent CitS citrate carrier led to the identification of a two-component regulatory system essential for anaerobic growth of *Klebsiella pneumoniae* (Bott et al., 1995). The system consists of the sensor kinase CitA and the response regulator CitB.

The sensor kinase CitA is 547 amino acids in length. Analysis of the primary structure indicated that the protein is membrane bound and composed of an aminoterminal periplasmic domain enclosed by two transmembrane helices, a central cytoplasmic domain, and carboxyterminal cytoplasmic kinase domain harboring the autophosphorylatable histidine residue. The periplasmic domain has been defined as a highly specific citrate receptor (this work).

The response regulator is composed of 234 amino acids and consists of an aminoterminal receiver domain containing the phosphorylatable aspartate residue, and a carboxyterminal output domain with a helix-turn-helix motif typical for DNA binding proteins. Recently, it has been shown that the carboxyterminal domain binds to DNA of the control region of the citS operon of *Klebsiella pneumoniae* (Meyer et al., 1997). In general, citrate fermentation is achieved in presence of citrate and Na⁺ which triggers the phosphorylation cascade catalyzed by the CitA/CitB two-component system. Phosphorylized CitB activates transcription of the citS promoters, leading to elevated levels of the citrate carrier protein CitS. Adequate levels of the housekeeping enzymes are controlled by mRNA processing and partial transcription termination (Bott et al., 1995).
3.3. References


Aims and Scope of the Presented Work

4.1. Objectives

Specific binding is a fundamental process on which virtually all biological systems depend critically, be it macromolecule-macromolecule or macromolecule-small molecule interactions. In order to understand and manipulate these processes, it is very important to obtain a close and detailed description of the forces that drive complex formation. This includes the determination of changes of all thermodynamic parameters, including free energy of binding ($\Delta G$), enthalpy ($\Delta H$) and entropy ($\Delta S$) of binding and the heat capacity change ($\Delta C_p$). The main aspect of the presented work is the investigation and elucidation of binding characteristics from a thermodynamic point of view with respect to the ligand-enzyme interactions of herpes simplex virus type 1 thymidine kinase (HSV1 TK) and of ligand-receptor interactions of the periplasmic domain (CitAP) of the histidine autokinase CitA.

HSV1 TK has become an important target in medicinal chemistry because of its links with therapy of viral infections, gene therapy of cancer (Culver et al., 1994; Kunishige et al., 1999; Tong et al., 1998) and AIDS (Caruso & Bank, 1997; Smith et al., 1996), vascular proliferative disorders (Ohno et al., 1994; Xu et al., 1998). In addition, it is being used as control system in allogeneic bone marrow transplantation (Bonini et al., 1997) and AIDS vaccine (Chakrabarti et al., 1996), for non-invasive diagnosis of viral encephalitis (Iwashina et al., 1988) and as expression reporter gene (Gambhir et al., 1999). For above applications, a variety of compounds is required that will comfort to the individual needs of the system different systems. The fact that these compounds are not readily synthesized clearly shows the lack of knowledge concerning ligand interactions of HSV1 TK. A deeper understanding of these
interactions from a thermodynamic point of view will contribute to the development for the design of new and effective compounds.

To achieve this, we applied isothermal titration calorimetry (ITC) which has emerged as the premier tool for characterizing interactions in terms of thermodynamic parameters. Therefore, several specific objectives are formulated.

As a first objective, the expression, isolation and purification had to be optimized. The major goal was the establishment of a strategy that led to pure and active fusion protein, free of substrate and cofactor. Recently, in our laboratory the expression and purification of recombinant HSV1 TK has been established and optimized for obtaining highly pure protein suitable for kinetic measurements and crystallization (Bohner, 1996; Fetzer et al., 1994). However, it has been shown that the kinase is very unstable, and it was necessary to develop modified and refined protocols with respect to stability, quality and high yield of the macromolecule.

The second objective was devoted to the development of proper experimental conditions suitable for microcalorimetric experiments. Conditions were subsequently optimized and verified in order to avoid possible systematic errors (pH-dependent activity, kinetic behavior, CD spectroscopy, HPLC).

The thermodynamic characterization and exploration of ligand binding to HSV1 TK denoted the third objective. ITC has been used to investigate the binding energetics of all elementary steps forming the catalytically competent state, in the temperature range of 10-25°C. Experimentally obtained thermodynamic quantities should then be linked with structure-based thermodynamic analyses based on the known structure of the ternary TK:dT:ATP complex of HSV1 TK.

Objective four was concerned with the elucidation of binding energetics of several mutants of the residue triad H58/M128/Y172 that were inactive and not distinguishable from each other by kinetic measurements. The background of this objective is whether loss in phosphorylation activity is due to non-binding or altered binding behavior of the mutants.

CitA is the sensor kinase of the two-component regulatory system CitAB of K. pneumoniae, which is responsible for the induction of responsible enzymes involved in citrate fermentation under anaerobic conditions. The activation of the two-component regulatory system occurs in presence of citrate, and it was proposed that CitA might function as a citrate receptor (Bott et al., 1995). The main focus of this
PART 1: Aims and Scope of the Presented Work

System was the investigation of the sensory properties of the periplasmic domain of the histidine autokinase CitA.

Therefore, the fifth objective was the establishment of proper experimental conditions to perform ITC. For this work, the receptor was available in form of the recombinantly overproduced periplasmic domain with a C-terminally attached histidine tag (CitAP\textsubscript{His}).

The seventh objective was devoted to the thermodynamic characterization of citrate binding to CitAP\textsubscript{His} dependent on pH. Specificity was explored by means of citrate derivatives isocitrate and tricarballylate.

The thermodynamic approach turned out to be very efficient and useful, comprising a lot of information about recognition processes between small ligands and biological macromolecules.

4.2. References


PART 2

Experimentals
Method Development and Improvement

1.1. Introduction

When performing microcalorimetric experiments, one should keep in mind that usually relatively high protein concentrations will be required, particularly in case of low affinity interactions. Thus, stable and reliable purification protocols must be established to provide the macromolecule in high yields.

In theory, calorimetric experiments can tolerate any amount of impurities or contaminating material provided they do not interfere with the reactions of interest. On the one hand, this can be particularly useful with turbid samples which are difficult to measure by spectroscopic methods. On the other hand, at the high protein concentrations required, even trace contamination with other enzymes (proteases, nucleases, phosphatases, and so on) may catalyze unwanted side reactions with ligands or macromolecules, leading to additional heat effects not due to specific binding. Therefore, it is strongly recommended to use adequately purified protein.

The outcome of calorimetric experiments is strongly dependent on the knowledge of the exact concentrations of ligand and macromolecule used. As \( \Delta H \) is directly dependent on the amount of ligand applied, the error in determination of its concentration will directly influence the reliability of this parameter. Moreover, if stoichiometric data are to be extracted, not only the protein concentration must be exactly known, but also the amount of active protein in the sample. Therefore, specific protocols for concentration determination must be established.
1.2. Expression and Purification

Recently, in our laboratory the expression and purification of recombinant HSV1 TK has been established and optimized for obtaining highly pure protein suitable for kinetic measurements and crystallization (Bohner, 1996; Fetzer et al., 1994). However, it has been shown that the kinase is very unstable, and it was necessary to develop altered and refined protocols with respect to stability, quality and high yield of the macromolecule.

1.2.1. Expression

Recombinant protein expression was achieved using the isopropyl β-D-thiogalactopyranoside (IPTG) inducible plasmid vector pGEX2T-TK (Pharmacia) as a thrombin cleavable glutathione S-transferase fusion protein as described earlier (Fetzer & Folkers, 1992; Fetzer et al., 1994). The TK deficient E.coli strain KY895 was used as host, which was changed to the expression strain BL21 later. After being transformed with the plasmid, bacteria were grown overnight at 37°C in Luria-Bertani (LB) medium containing 100 μg/ml ampicillin. The cultures were diluted 1:10 in fresh medium and grown for 3 hours at 25°C. Gene expression was induced by addition of IPTG to a final concentration of 100 μM. After 36 hours for KY895 or 20 hours (over night) for BL21, the cells were harvested by centrifugation at 4°C and frozen at -70°C.

1.2.2. Isolation of HSV1 TK

The protein was isolated mainly according to a previously reported procedure (Fetzer et al., 1994), with minor changes. Briefly summarized, the frozen pellet was thawed and suspended in buffer (50 mM Tris/HCl pH 7.5, 5 mM EDTA, 10% glycerol, 1% Triton X-100) before adding 1 mM PMSF, 10 mM DTT, 10 mM MgCl₂, 1 mM MnCl₂, and lysed on ice in presence of 150 μg/ml lysozyme and 2000 units DNaseI for 15 min. Afterwards, the mixture was further treated by continuous sonication at 4°C (Bohner, 1996). The lysate was substituted with 10 mM EDTA to inactivate DNaseI, clarified by centrifugation at 12,000xg for 20 min, filtered (0.45 μM) and subjected to purification or frozen at -70°C, where it remained stable for months.
1.2.2. Purification of HSV1 TK Fusion Protein

HSV1 TK was expressed as glutathione S-transferase fusion protein (GST-TK) and purified by glutathione affinity chromatography (Hengen, 1996; LaVallie & McCoy, 1995; Smith & Johnson, 1988). Formerly established protocols (Fetzer et al., 1994) were modified since they were not suitable for ITC experiments. For stability reasons (see 4.5.) it was decided to carry out all measurements with the GST-fusion protein of HSV1 TK since the kinetic properties are identical to thymidine kinase (Fetzer et al., 1994).

The major goal was the establishment of a strategy that yields pure and active fusion protein, free of substrate and cofactor. Therefore, a single-step purification procedure has been developed. Generally, the crude extract containing the fusion protein was applied to the glutathione sepharose column. After thoroughly washing, the fusion protein was eluted by addition of 5 mM glutathione (GSH). Purification was monitored by SDS-PAGE using the PhastSystem from Pharmacia with precasted gels (PhastGel Gradient 10-15) and PhastGel SDS buffer strips. Expression in strain KY895 was found to be ineffective, always leading to low expression levels of the fusion protein, as seen in the crude extracts, and minor yields of fusion protein (MW of 67 kDa) contaminated with unknown impurities (46 kDa and 30 kDa, respectively), as seen after purification. Figure 1 (A and B) represents typical results of different batches.

![Fig. 1: SDS-PAGE analysis of the purification procedure. Vertically arranged numbers indicate molecular weights of the marker (lane M), horizontal number depict lanes; FP = fusion protein; GST = glutathione S-transferase (26 kDa); TK = HSV1 TK (42 kDa); F1 = fragment 1 (30 kDa); F2 = fragment 2 (46 kDa). Lane 1 shows crude extract, lane 2 flow-through from affinity purification. (A), (B) Standard purification protocol yielded low expression and truncated fusion protein (A3; B3-5).]
Judged on the molecular weight of the impurities and the overall low yield, it was concluded that the fragmentation of the fusion protein could be due to protease activity in the host or due to extract preparation. This would result in one fragment consisting of GST and a second fragment with HSV1 TK origin. The origin of the additional bands in the gels was further investigated by a thrombin-cleavage assay (Fig. 2A) and Western blot analysis (Fig. 2B).

The 30 kDa fragment (F1) was further analyzed for the presence of a thrombin cleavage site. Thrombin cleavage was achieved by adding a small amount of thrombin and CaCl₂ to the protein fraction containing F1 (Fig. 1A, F1), and was monitored for 5 hours by SDS-PAGE. From figure 2A it is clearly seen that there is a thrombin cleavage site in F1. Cleavage results first in a fragment of 26 kDa that became the main product after 5 hours. Most probably, F1 is the GST part of the fusion protein including the engineered thrombin cleavage site and parts of the N-terminal sequence of HSV1 TK. This region has been recognized to be exceptionally susceptible to protease cleavage just recently (Pilger, 1999).

These findings are further corroborated by Western blot analysis. They showed that fusion protein (FP) and fragments descending from the GST part (F1) and HSV1 TK
Fig. 3: SDS-PAGE analysis of the optimized purification procedures. Horizontally arranged numbers indicate molecular weights of the marker (lane M), vertical number depict lanes. Lane 1 shows crude extract after the isolation process, lane 2 flow-through after affinity purification. FP = fusion protein; F1 = fragment 1 (30 kDa); F2 = fragment 2 (46 kDa). (A) Results of typical expression using strain BL21. The fusion protein is highly expressed, forming a predominant band on the gel at a molecular weight of 66 kDa (lane 1-2). Elution yields the protein of ~70% purity (lane 3). (B) The quality of fusion protein achievable with the new procedure is shown in lane 3 and 4.

part (F2) of the fusion protein are present in the crude extract and the purified protein (Fig.2B). Obviously, proteolysis is not mainly due to the purification protocol, but occurs during expression in the host.

Expression strain KY895 was changed to BL21 which is deficient in several proteases. This let to high yields of fusion protein that was still contaminated by several unknown proteins (Fig.3A), amongst which the GST part and the HSV1 TK part of the fusion protein could be detected by Western blot analysis (Fig.2B). Subsequent optimization of the purification protocol with respect to protein contamination through unspecific column interactions, proteolytic cleavage and DnaK co-purification, finally let to about 80% pure fusion protein as detected by SDS-PAGE and quantified by gel densitometry (Fig.3B) The final purification procedure was following: the crude extract is applied twice to the glutathione sepharose column, washed with buffer A (50 mM Tris/HCl pH 7.5, 150mM NaCl, 4 mM EDTA, 10% glycerol, 1% Triton X-100), then washed with buffer A containing 10 mM MgATP, and finally, the protein was directly (on-column) exchanged into the experimental buffer by thoroughly rinsing the column with excess of buffer. The protein was eluted by addition of 5 mM GSH and was directly used for titration experiments. This protocol yielded about 10 to 20 mg protein per liter of culture.
1.2.3. Expression and Purification of CitAP₇His

For the calorimetric binding studies in respect of periplasmic domain of the histidine autokinase CitAP, no protocol had to be established since pure and stable protein was available in high amounts.

Briefly, strain BL21, harboring expression plasmid pET-CitAP, served as host for overproduction of CitAP₇His and were grown in LB medium containing 50 μg/ml kanamycin. The cultures were incubated at 37°C until the optical density at 600 nm (OD₅₀₀) reached a value between 0.6 and 0.8. Expression was induced by the addition 1 mM IPTG and the cultures were incubated for another three hours at 30°C. Subsequently, cells were harvested by centrifugation, washed once in buffer A (20 mM Tris/HCl pH 7.9, 500 mM NaCl) containing 5 mM imidazole, and stored at -20°C. For disruption, the cells were resuspended in the same buffer (4 ml/g wet weight) supplemented with 0.25 mg/ml DNasel and a protease inhibitor cocktail (Complete, EDTA-free) at the concentration recommended by the supplier (Boehringer Mannheim). After passing the cell suspension twice through a French pressure cell (SLM Aminco) at 108 MPa, intact cells and cell debris were removed by centrifugation (30 min at 27000xg, 4°C). The cell-free extract was subjected to ultracentrifugation (1 h at 150000xg, 4°C) to sediment the membranes. The supernatant was passed through a 0.2 μm filter and used for the isolation of the desired proteins.

CitAP₇His was purified by Ni²⁺ chelate affinity chromatography essentially as described previously for CitB₇His (Meyer et al., 1997). The soluble fraction obtained from 1-2 g cells (wet weight) was loaded onto a column with 2 ml (bed volume) His-bind resin (Novagen) pre-equilibrated with buffer A (20 mM Tris/HCl pH 7.9, 500 mM NaCl) containing 5 mM imidazole. Weakly bound proteins were removed by washing with five bed volumes of buffer A containing 30 mM imidazole. Elution was performed with 400 mM imidazole in buffer A. Subsequent buffer exchanges were performed by gel filtration with Sephadex G-25 (PD-10 columns, Pharmacia). The purification of all proteins was monitored by SDS-PAGE (Laemmli, 1970) and staining with Coomassie brilliant blue.

This protocol yielded about 10 to 20 mg protein per g cell weight of high purity (>98%).
1.3. Concentration Determination Procedures

1.3.1. Concentration of HSV1 TK
Since the protein was impure, concentration could not be accomplished by amino acid analysis (Gill & von Hippel, 1989), and it was not possible to establish a precise standard curve on the basis of the fusion protein. Under these circumstances, it is even more difficult to estimate the amount of active protein in the sample. Therefore, protein concentrations were estimated using a dye-binding assay (Bradford, 1976) with bovine serum albumin as standard. The content was corrected for impurities detected by SDS-PAGE and quantified by gel densitometry using a CAMAG Electrophoreses Scanner II with CATS software.

1.3.2. Concentration Determination of CitAPHis
The protein was essentially pure and concentrations could easily be calculated by amino acid analysis (Gill & von Hippel, 1989). CitAPHis concentrations were determined spectrophotometrically at 276 nm using $\varepsilon = 7,25 \text{ mM}^{-1}\text{cm}^{-1}$.

1.3.3. Concentration of dT and ATP
Determination of dT and ATP was straightforward and could easily be done by spectroscopic methods. For experiments where both dT and ATP were present in the final ligand solutions, concentrations were determined by an HPLC assay (see 4.6.)

1.3.4. Concentration of Citrate and Citrate Analogs
Citrate was directly used without further purification, and concentration were calculated by weight.

1.4. Kinetic and Activity Measurements
The purification protocol led to samples containing 5 mM glutathione, the influence of which on the dT/enzyme interaction was tested based on kinetic measurements. Furthermore, the experiments were planned to take place at the pH of maximum activity of HSV1 TK. To date, no data were available with respect to pH dependence
of the fusion protein. Consequently, these parameters were investigated using a radioactive enzyme assay previously established (Gerber & Folkers, 1996). This assay is based on the DEAE-cellulose method (Furlong, 1963) which represents a rapid assay technique for nucleoside kinases. The anion-exchange paper retains labeled $^3$H-dTMP but allows the unreacted $^3$H-dT to be washed away. The reaction products adsorbed on the paper are counted in a liquid scintillation counter after digestion of the DEAE paper with cellulase prior to scintillation counting. This procedure is especially useful if solution conditions like pH, salt, buffer substances have to be changed.

1.4.1. Effect of Glutathione

The effect of GSH on the kinetics of HSV1 TK at 37°C was evaluated by double reciprocal analysis of initial velocity data obtained with variable concentrations of dT from 20 μM to 800 μM. The reaction mixture contained 50 mM Tris/HCl pH 7.2, 0.2% BSA, 5 mM ATP, 5 mM, 5 mM MgCl$_2$ and 10 ng/ml purified cleaved HSV1 TK. For measurements in presence of 5 mM GSH a $K_m$ value of 0.18 μM has been determined, corresponding very well to the value of 0.14 μM without GSH (Fig.4). $K_m$ reported in literature is 0.2 μM (Chen et al., 1979; Michael et al., 1994). With respect to $V_{max}$, no significant difference was found for reactions in presence of 5 mM GSH ($V_{max} = 493$ pmol/mg⁻¹/min⁻¹) and without GSH ($V_{max} = 443$ pmol/mg⁻¹/min⁻¹).

1.4.2. pH-Dependent Activity

The effect of pH on thymidine kinase activity was studied by varying the pH of the reaction buffer with 100 mM Tris/HCl in the range of 6.5 to 9.5 (Fig.5). Reaction mixtures contained 0.175% BSA, 5 mM ATP, 5 mM MgCl$_2$ and 18 ng/ml purified fusion protein. Since the buffering capacity is low beyond pK ± 1 (pKa Tris/HCl = 8.0), pH of the actual solutions was controlled at 37°C.

The catalytic activity of HSV1 TK of fusion protein increases from pH 6.5 to 7.5 and reaches a maximum between pH 7 to 8. At higher pH values, the activity decreases sharply. Suitable conditions for calorimetry would be at pH of maximum activity. Thus, measurements were carried out at pH 7.5.
Fig. 4: Kinetics of substrate phosphorylation of recombinant HSV1 TK fusion protein without added GSH (filled squares) and in presence of 5 mM GSH (open circles). $v$ is given as $1/$pmol$\cdot$$\mu$g$^{-1}$$\cdot$min$^{-1}$. $dT$ is given in $\mu$M. The solid lines represent linear regression fits to the data, yielding $K_{m}$ for $dT$ of 0.14 $\mu$M and 0.18 $\mu$M without and in presence of GSH, respectively.

Fig. 5: Effect of pH on thymidine kinase activity of HSV1 TK fusion protein measured at 37°C. The fusion protein exhibits maximal activity in the pH range of 7.5 to 8.
1.5. Stability Measurements

1.5.1. Time Dependent Activity

For rapid and easy screenings a enzyme coupled UV-spectrophotometric assay has been established (Gerber, 1997) which monitors the time dependent change in $A_{340}$ due to oxidation of NADH and corresponds to the ADP formation during the phosphorylation reaction. The method has limitations since it is associated with an enzyme cascade that prevents modifications of the reaction conditions such as pH. Activity measurements were carried out at 37°C with 75 $\mu$l of a reaction mixture containing 50 mM Tris/HCl pH 7.2, 5 mM MgCl$_2$, 1 mM DTT, 0.3 mM NADH, 0.4 mM phosphoenol pyruvate, 0.45 U pyruvate kinase, 0.5 U lactate dehydrogenase, 1 mM dT, 160 $\mu$M ATP, and 1.7 $\mu$g fusion protein. No loss of activity could be detected for at least 20 hours. Afterwards, the catalytic activity decreased sharply to a residual value of 20% of the starting conditions.

Fig. 7: Decrease in activity of HSV1 TK fusion protein at 25°C as a function of time. Conditions were: 50 mM Tris/HCl pH 7.5, 4 mM EDTA, 5 mM GSH, 1 mM DTT. For at least 24 hours, no loss in activity could be detected, thus the enzyme remained stable under these conditions. Activity decreased sharply after 30 hours to a residual activity of 20% of the starting value.
1.5.2. Thermal Stability

Modern calorimeters allow precise measurements in the temperature range of about 5°C to 70°C which can be exploited to calculate ΔCp from the temperature dependence of ΔH or the thermodynamic linkage of KB with temperature. The fundamental assumption of the method is that the measured thermodynamic parameters are the resultant of the ligand binding equilibrium of the natively folded form of the biological macromolecules. However, this assumption should be tested whenever calorimetric experiments are conducted since proteins unfold at higher temperatures.

Therefore the thermal stability of HSV1 TK was studied by monitoring ellipticity at 217 nm by circular dichroism as a function of temperature. Figure 8 shows that global unfolding occurs with a Tm of 42°C with transition starting near 35°C. The maximum temperature useful for the present ITC binding study would be about 30°C.

Fig.8: Thermal unfolding of HSV1 TK monitored by circular dichroism. Measurements were done at a scan rate of 20°C/h. Conditions were 10 mM Tris/HCl pH 7.5, 1 mM EDTA and a protein concentration of 0.4 mg/ml. The signal is normalized to the monomer. Unfolding begins to occur near 35°C, giving the maximum useful temperature for titration experiments as 30°C (figure kindly provided by Ch. Wurth).
1.6. High Performance Liquid Chromatography

High performance liquid chromatography was applied for concentration determination of thymidine and ATP in the final ligand solutions, and to monitor potential phosphorylation products during calorimetric experiments. The HPLC system is based on ion-pair chromatography using a modified protocol of a previously published method (Masson et al., 1993). It is able to separate ADP, ATP, dT, dTMP, dTDP and dTTP in a single run (Fig.5).

Fig.8: HPLC system for monitoring ADP and nucleoside mono-, diphosphate and triphosphate formation during substrate catalysis.
Nucleotides were determined by reverse-phase ion-paired chromatography using a C18 column (LiChromspher 100 RP-18, 5 μm, 250x4 mm, Merck), with the mobile phase (0.2 M NaH₂PO₄, 25 mM tetrabutylammonium, 3% (v/v) methanol) delivered at a rate of 1.0 ml/min and detection at 254 nm.

1.6.1. Concentration Determination of Ligand Mixtures

Ligand concentrations were calculated by means of calibration curves from standard solutions showing linearity in the range of <0.04 to at least 2 mM. The detection limit for dT, dTMP and ATP has been determined to be < 20 nmol (Pilger et al., 1999).

1.6.2. Monitoring Phosphorylation Products

Under the experimental conditions of microcalorimetry, phosphorylation can occur due to traces of divalent cations in the solution. The amount of EDTA needed to complex metal cations in order to avoid any phosphorylation products during titration experiments was determined by means of the HPLC assay. HSV1 TK fusion protein (64 μg/ml) was incubated with 42 μM dT and ATP and various concentrations of EDTA (1-7.5 mM). The protein was added last to the reaction mix. The effect of EDTA on phosphorylation was monitored immediately after addition of the protein (0 min) and after 80 min, the average duration of a titration experiment (Fig. 9). In addition, a control containing no protein was prepared and examined by the same procedure.

As depicted in Figure 9, without any EDTA phosphorylation takes place immediately after addition of HSV1 TK to the reaction mix. After 80 min almost all ATP (retention time Rₜ=10.2 min) is transformed to ADP (Rₜ=4.8 min), whereas most of dT (Rₜ=6.3 min) has reacted to dTMP (Rₜ=8.1 min). Since catalytic activity depends on Mg²⁺ that has never been added to the solutions, the observed phosphorylation is due to traces of divalent cations in the solutions used. Subsequent examination of the effect of increasing EDTA concentrations showed that 3 to 4 mM EDTA efficiently prevented phosphorylation activity during the titration of dT in presence of ATP. All calorimetric experiments were carried out at 4 mM EDTA concentrations. The control run exhibited no catalytically independent ATP hydrolysis.
Fig. 9: Examination of various EDTA concentrations on phosphorylation activity of HSV1 TK fusion protein due to traces of divalent cations. The retention times ($R_t$) are: ADP, 4.8; dT, 6.3; dTMP, 8.1; ATP, 10.2; GSH, 4.4. In conditions containing no EDTA, phosphorylation takes place immediately after adding protein to the mix. After 80 min most of ATP and dT is consumed to build dTMP and ADP. Subsequent examination of various EDTA concentrations shows that in the range of 3 to 4 mM EDTA efficiently prevents catalytic activity. Higher concentrations do not improve the result. The control experiments in which no HSV1 TK was present, show that dT and ATP are stable over the time of experiment.
1.7. References


PART 3

Results
Seite Leer / Blank leaf
Compulsory Order of Substrate Binding to HSV1 Thymidine Kinase: A Calorimetric Study


*Department of Pharmacy, Swiss Federal Institute of Technology (ETH), Winterthurerstr. 190, CH-8057 Zürich, Switzerland and **Department of Biochemistry, University of Zurich, Winterthurerstr. 190, CH-8057 Zürich

submitted
Abbreviations:

ADK: adenylate kinase
ADP: adenosine diphosphate
ADPCP: β,γ-methyleneadenosine 5′-triphosphate
AMP: adenosine monophosphate
AMPCF₂P: β,γ-difluoromethyleneadenosine 5′-triphosphate
ATP: adenosine triphosphate
CD: circular dichroism
dT: thymidine
dTMP: thymidylate, thymidine monophosphate
DTT: DL-dithiothreitol
EDTA: ethylenediamine tetraacetic acid
GSH: glutathione
GST: glutathione S-transferase
HPLC: high performance liquid chromatography
HSV1 TK: herpes simplex virus type 1 thymidine kinase
hTK1: human cellular thymidine kinase
ITC: isothermal titration calorimetry
ΔCp: heat capacity change
ΔG: free energy change
ΔH: enthalpy change
ΔS: entropy change
1.1. Summary

Isothermal titration calorimetry (ITC) has been used to investigate the thermodynamic parameters of the binding of thymidine (dT) and adenosine triphosphate (ATP) to Herpes Simplex Virus Type 1 Thymidine Kinase (HSV1 TK). Binding follows a sequential pathway in which dT binds first and ATP second. The free enzyme does not bind ATP, whose binding site becomes only accessible in the HSV1 TK:dT complex. At pH 7.5 and 25°C, the binding constants are 1.9×10⁵ M⁻¹ for dT and 3.9×10⁶ M⁻¹ for ATP binding to the binary HSV1 TK:dT complex. Binding of both substrates is enthalpy-driven and opposed by a large negative entropy change. The heat capacity change (ΔCp) obtained from ΔH in the range 10-25°C, is -360 cal K⁻¹ mol⁻¹ for dT binding and -140 cal K⁻¹ mol⁻¹ for ATP binding. These large ΔCp-values are incompatible with a rigid body binding model in which the dT and ATP binding sites pre-exist in the free enzyme. Values of ΔCp and TΔS strongly indicate large scale conformational adaptation of the active site in sequential substrate binding. The conformational changes seem to be more pronounced in dT binding than in the subsequent ATP binding. Considering the crystal structure of the ternary HSV1 TK:dT:ATP complex, a large movement in the dT binding domain and a smaller but substantial movement in the LID domain is proposed to take place when the enzyme changes from the substrate-free, presumably more open and less ordered conformation, to the closed and compact conformation of the ternary enzyme-substrate complex.

Keywords: Thymidine kinase, titration calorimetry, heat capacity, conformational entropy, conformational adaptation
1.2. Introduction

Molecular recognition phenomena are at the heart of biological reactions. Key to the understanding of molecular recognition is a comprehensive analysis of the thermodynamics of binding and a meaningful correlation of thermodynamics with structure. A close insight into the thermodynamics of a binding process provides guide marks for structure-based molecular design strategies. The forces that govern a binding reaction are the free energy change ($\Delta G$), the enthalpy change ($\Delta H$), the entropy change ($\Delta S$) and the heat capacity change ($\Delta C_p$). $\Delta C_p$ is an approximate measure of the surface area buried in an association reaction and can be used to predict conformational rearrangements in associating protein molecules. An example of high medicinal interest where such information is essential is thymidine kinase from Herpes Simplex Virus Type-1 (HSV1 TK). The structure of this enzyme is known at high resolution in complex with a series of ligands, including various substrates (natural and non-natural) and inhibitors (Bennett et al., 1999; Brown et al., 1995; Champness et al., 1998; Wild et al., 1995; Wild et al., 1997).

Thymidine kinases (EC 2.7.1.21) catalyze the phosphorylation of thymidine (dT) to thymidine monophosphate (dTMP) in presence of magnesium ions by transferring the $\gamma$-phosphate group of adenosine triphosphate (ATP) to the $5'$-OH group of dT. Herpes viruses encode their own thymidine kinases, which differ considerably from the enzyme of the human cellular host (hTK1). While the human enzyme is highly specific, HSV1 TK is a multifunctional enzyme of broad substrate specificity. It shows deoxycytidine kinase and thymidylate kinase activity (Chen & Prusoff, 1978) and phosphorylates a broad spectrum of pyrimidine as well as purine analogs (Chen et al., 1979; Keller et al., 1981; Larder et al., 1983; Larder & Darby, 1982; Smee et al., 1983; Waldman et al., 1983). Moreover, HSV1 TK displays low stereochemical specificity. The enzyme accepts modified ribose moieties, acyclic side chains and the L-stereoisomer of the deoxyribose of dT (Spadari et al., 1992). The preferred phosphate donor is ATP, yet HSV1 TK also shows high affinity for cytidine triphosphate, uridine triphosphate and guanosine triphosphate and their deoxy analogs.

Therapeutic applications involving HSV1 TK make use of the broad substrate diversity of the viral enzyme in the background of strict substrate selectivity of the host cell enzyme. Therefore, a detailed thermodynamic analysis of substrate binding
**Figure 1:** Ribbon diagram of the symmetric HSV1 TK dimer with bound ADP and dTMP (PDB entry 1VTK; Wild et al., 1997). The domains are defined as for other NMP-kinases (Vonrhein et al., 1995). The CORE domain is depicted in blue, the NMP$_{bind}$ domain in red, and the LID domain in yellow. The additional residues (250-322) are shown in green. Substrate and cofactor are depicted as cyan ball and stick model. The picture was generated with the program MOLSCRIPT (Kraulis, 1991).

to the viral kinase is a prerequisite for the successful design of new therapeutically useful compounds.

HSV1 TK is a homodimer with 376 residues per subunit (Figure 1). The constituent subunits display the general $\alpha/\beta$ folding pattern. A central five stranded parallel $\beta$-sheet is flanked on either side by helices. HSV1 TK is a member of the family of NMP-kinases and contains the classical mononucleotide (NMP) binding fold (Schulz, 1992). In this enzyme family, 3D-structures are known for adenylate kinase (ADK) (Dreusicke et al., 1988; Schlauderer & Schulz, 1996), guanylate kinase (Stehle & Schulz, 1990), uridylate kinase (Muller-Dieckmann & Schulz, 1994), bacteriophage
T4 deoxynucleotide kinase (Teplyakov et al., 1996) and thymidylate kinase (Lavie et al., 1997). The central five β-strand domain is referred to as the CORE domain. Other domains are the LID domain and the NMP\textsubscript{bind} domain (Vonrhein et al., 1995). Main chain superposition of HSV1 TK and ADK reveals substantial similarity in the CORE domain. Major differences exist in the NMP\textsubscript{bind} domain, where HSV1 TK has extensive insertions. The LID domain of HSV1 TK consists of only eight residues, reminiscent of other small variants of NMP-kinases. Further differences to ADK arise from the following unique structural features of HSV1 TK. A 45 residue long amino terminal segment, which is not resolved in the crystal structure and is not necessary for catalytic activity (Halpern & Smiley, 1984) but plays a role in migration within the cell (Degreve et al., 1998); a partially mobile segment of 72 residues between residues 250-322; and an additional 29 C-terminal residues (Wild et al., 1997). Interestingly, many of these differences appear to be located close to the dimer interface.

Adenylate kinases undergo large conformational changes upon substrate binding as shown by crystallography (Muller et al., 1996). The substrate-free enzyme has a more open conformation, and substrate binding leads to a closed conformation. The NMP\textsubscript{bind} domain and the LID domain rearrange upon binding of AMP and ATP, whereas the conformation of the CORE domain remains unchanged (Muller et al., 1996; Vonrhein et al., 1995). The 3D-structures of HSV1 TK known to date have been solved for ternary complexes with natural and non-natural substrates, inhibitors, cofactors, or sulfate ions mimicking the β-phosphate of ATP (Bennett et al., 1999; Brown et al., 1995; Champness et al., 1998; Wild et al., 1995; Wild et al., 1997) and correspond to the closed conformation of NMP-kinases. The structure of the free apo-enzyme is not known. Since structural similarities in the CORE domain and substrate binding pockets are substantial in the NMP-kinase family, one may assume that substrate-free HSV1 TK also exists in an open conformation and that conformational changes take place when HSV1 TK is converted to the closed conformation during substrate binding.

Here, we present a comprehensive thermodynamic analysis of nucleoside (dT) and cofactor (ATP) binding to HSV1 TK. Substrate binding, which was followed by ITC, is shown to be strictly sequential with dT binding first and ATP second. Combining the thermodynamic parameters with the 3D-structure of the ternary HSV1 TK:dT:ATP complex provides insights into the binding mechanism and the structural changes that accompany substrate binding.
complex demonstrates that the sequential binding pathway is accompanied by significant structural rearrangements of the enzyme.

1.3. Results

1.3.1. Experimental Setup

The glutathione S-transferase (GST) fusion protein of HSV1 TK was used in this study to facilitate protein purification and to improve stability during storage. To rule out artifacts caused by the presence of GST in the construct, control experiments were performed with the HSV1 TK obtained after on-column cleavage of the affinity tag. Thermodynamic parameters determined by ITC were identical within error for cleaved HSV1 TK and for the GST-fusion protein (data not shown). The influence of glutathione in concentrations up to 5 mM to HSV1 TK kinetics was investigated as previously published (Fetzer et al., 1994; Gerber & Folkers, 1996). Glutathione in concentration up to 5 mM does not change $K_m$ (data not shown). Since $Mg^{2+}$ is strictly required for the phosphorylation reaction catalyzed by HSV1 TK, all measurements were performed in the presence of 4 mM EDTA to suppress any enzymatic activity in ITC experiments. Since $Mg^{2+}$ could influence the equilibrium binding parameters by either direct participation in the bonding network of the binding pocket or by indirect structural effects, control titrations with the non-cleavable ATP derivative $\beta,\gamma$-methyleneadenosine 5'-triphosphate were performed in the presence and absence of $Mg^{2+}$. No differences were seen (data not shown).

1.3.2. Determination of the Thermodynamic Parameters of dT and ATP Binding by ITC

In the ternary complex TK:dT:ATP, the substrate dT and the cofactor ATP are located in separate and well defined binding pockets of the enzyme. Formation of the ternary enzyme-substrate complex may either proceed through an obligatory sequential pathway or by a random mechanism. Two sequential pathways are possible: $\text{TK} \rightarrow \text{TK:dT} \rightarrow \text{TK:dT:ATP}$ (reactions i and ii of Figure 2), or $\text{TK} \rightarrow \text{TK:ATP} \rightarrow \text{TK:dT:ATP}$ (reactions iii and iv of Figure 2). In a random mechanism, binding of one substrate is not a prerequisite for binding of the other, and all four reactions of Figure 2 can take place.
Figure 2: Formation of the ternary enzyme-substrate complex TK:dT:ATP. The two ordered sequential pathways are i + ii and iii + iv, respectively. In a random binding mechanism, all four reactions take place.

To distinguish between ordered and random binding, HSV1 TK was titrated with dT and with ATP, respectively. Figure 3 shows representative titrations of the substrate-free enzyme with dT (panels A and B) and with ATP (panels C and D). The titration with dT was characterized by a significant exothermic heat effect. In panel B, the integrated heat of each injection (filled squares) is plotted against the molar ratio of dT to enzyme binding sites. The solid line is a nonlinear least squares fit for a single-site binding model with $K_B = 1.9 \times 10^5 \text{ M}^{-1}$ and $\Delta H_{\text{bind}} = -19.1 \text{ kcal mol}^{-1}$ (Table 1). Titration with ATP showed a very different behavior (panels C and D). Only very small heat signals corresponding to nonspecific heat effects were detected. Titrations were repeated at 15°C and at varying concentration ratios of ATP to enzyme.

Figure 3: A,B: Titration of HSV1 TK with dT, corresponding to reaction i of Figure 2. C,D: Titration of TK with ATP, corresponding to reaction iii of Figure 2. E,F: Titration of TK with ATP in the presence of excess dT, corresponding to reaction ii of Figure 2. G,H: Titration of HSV1 TK with 1:1 mixture of dT and ATP. Raw data are shown as differential power signals in panels A, C, E and G. Binding isotherms obtained by integration and normalization of the raw data and by correction for the heat of ligand dilution are shown in panels B, D, F and H. The solid lines represent non-linear best fits for a single-site binding model. Titrations were performed in 50 mM Tris/HCl pH 7.5, 25°C. A,B: 86 µM enzyme was titrated with 2.2 mM dT. C,D: 51 µM enzyme was titrated with 1.3 mM ATP. E,F: 56 µM enzyme in the presence of 1 mM dT was titrated with 1.4 mM ATP. Under these conditions, >99% of the enzyme was present as TK:dT complex. G,H: 62 µM enzyme was titrated with a 1:1 mixture of 1.6 mM dT and ATP.
No heat effect typical of a binding reaction was observed (data not shown). This indicates that empty HSV1 TK does not bind ATP. Binding driven by entropy without heat change could be excluded (see below). In agreement with the inability of the apo-enzyme to bind ATP is the fact that thymidine-depleted HSV1 TK did not bind to ATP-affinity columns, even at very high protein concentration (Fetzer et al., 1994).

To confirm an ordered binding mechanism in which TK:dT precedes the formation of TK:dT:ATP, the experiment shown in panels E and F of Figure 3 was performed. The preformed TK:dT binary complex (1 mM dT) was titrated with ATP (reaction ii of Figure 2). The reaction was again exothermic and yielded values of $K_B = 3.9 \times 10^6 \text{ M}^{-1}$ and $\Delta H_{\text{bind}} = -13.8 \text{ kcal mol}^{-1}$ (Table 1; corrected for protonation effect; see below).

In the cell, HSV1 TK is exposed to both substrates at the same time, which in the scheme of Figure 2 corresponds to a move on the diagonal from TK to TK:dT:ATP. If the binding reaction is indeed sequential as expected from the above experiments, titration of TK with a 1:1 mixture of dT and ATP should yield a heat change corresponding to the sum of the heat changes for reactions i and ii of Figure 2. As shown in panels G and H of Figure 3, the titration isotherm yielded $\Delta H_{\text{bind}} = -33.1 \text{ kcal mol}^{-1}$ (corrected for heat of protonation), identical within error to the sum of the enthalpy changes of reactions i and ii.

The ITC experiment provides the binding constant $K_B$ for a single-site reaction, and $\Delta G$ of reactions i and ii were calculated from $\Delta G = -RT \ln K_B$. $\Delta S$ was obtained from $\Delta G = \Delta H - T \Delta S$. Reactions i and ii were driven by favorable negative changes in binding enthalpy and strongly opposed by unfavorable entropic contributions. Although the reaction with the 1:1 mixture of dT and ATP was more complex, it could still be treated as a single-site reaction if one considered dT+ATP as one ligand. In this case, $K_B$ obtained from ITC equaled $(K_i \times K_{ii})^{1/2}$ where $K_i$ and $K_{ii}$ were the binding constants for reactions i and ii, respectively. Hence, the apparent binding constant for the coupled reactions i and ii was $(K_B)^2$ and $\Delta G$ equaled $-RT \ln (K_B)^2$.

The stoichiometry of binding obtained from the ITC experiment was in the range 0.7 to 0.8 mol ligand per mol of HSV1 TK monomer. Deviation from a value of 1 was due to the presence of inactive protein. By ATP-affinity chromatography in presence of dT we detected 20-30% of inactive protein in all enzyme batches used for ITC. This HSV1 TK fraction does not bind to ATP-chromatography media and shows no secondary structure as determined by CD spectroscopy.
Table 1. Thermodynamic parameters for the binding of thymidine and ATP to HSV1 TK at pH 7.5 and 25°C \(^{a,b}\).

<table>
<thead>
<tr>
<th>Temp. (°C)</th>
<th>TK + dT (reaction i)</th>
<th>TK:dT + ATP (reaction ii)</th>
<th>TK + dT/ATP (^c)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(\Delta H)</td>
<td>(\Delta G)</td>
<td>(T\Delta S)</td>
</tr>
<tr>
<td>10</td>
<td>-13.6</td>
<td>-7.1</td>
<td>-6.5</td>
</tr>
<tr>
<td>20</td>
<td>-17.4</td>
<td>-7.1</td>
<td>-10.2</td>
</tr>
<tr>
<td>(\Delta Cp)</td>
<td>0.36</td>
<td>-0.14</td>
<td>-0.51</td>
</tr>
</tbody>
</table>

\(^{a}\) Values of \(\Delta G\), \(\Delta H\) and \(T\Delta S\) in kcal mol\(^{-1}\), \(\Delta Cp\) in kcal K\(^{-1}\) mol\(^{-1}\). \(^{b}\) Values are the mean of triplicates. \(\Delta G\) was calculated from \(\Delta G = -RT\ln K_B\), where \(K_B\) is the binding constant determined by ITC. Uncertainty of \(\Delta G\) is within ±0.35 kcal mol\(^{-1}\) of the mean. Errors of \(\Delta H\) are about ±5% and mainly reflect the error in ligand concentration. Maximal possible errors of \(T\Delta S\) are 1.5 kcal mol\(^{-1}\). Errors of \(\Delta Cp\) were estimated by reduction of the data set by one data point at a time and were, on average, ±0.02 kcal K\(^{-1}\) mol\(^{-1}\), i.e. within 5-15% of the reported mean. \(^{c}\) 1:1 mixture of dT and ATP. \(^{d}\) Corrected for protonation effects by equation 1. \(^{e}\) Calculated from \(\Delta G = -RT\ln(K_B)^2\).

1.3.3. Change of Protonation State

Substrate binding may cause the enzyme to take up or release protons, for example through \(pK_a\) changes of side chains accompanying the binding reaction. This will contribute to the overall heat change, \(\Delta H_{obs}\), measured in the ITC experiment. If ionizable groups undergo \(pK_a\) changes on complex formation, protons will be exchanged with the buffer. The heat of protonation/deprotonation depends on the ionization enthalpy of the buffer, \(\Delta H_{ion}\), according to

\[
\Delta H_{obs} = \Delta H_{bind} + n_{H^+} \Delta H_{ion}
\]  

where \(n_{H^+}\) designates the number of protons that are released (\(n_{H^+} > 0\)) or taken up (\(n_{H^+} < 0\)) by the buffer (Cooper & Johnson, 1994a). To study such protonation effects, titration experiments were repeated in various buffers of different \(\Delta H_{ion}\). The intrinsic enthalpy of binding, \(\Delta H_{bind}\), was obtained from the intercept (\(\Delta H_{ion} = 0\)) of a plot.
Figure 4: Experimentally observed enthalpy change, $\Delta H_{\text{obs}}$, as a function of the ionization enthalpy of the buffer, $\Delta H_{\text{ion}}$, at 25°C. ITC experiments were performed at pH 7.5 in PIPES, MOPS, TES and Tris buffers. A: Titration of enzyme with ATP in the presence of excess dT (reaction ii of Figure 2). B: Titration of enzyme with dT (reaction i of Figure 2). C: Titration of enzyme with 1:1 mixture of dT and ATP. Continuous lines are linear least-squares fits according to equation (1).

according to equation 1. The results are shown in Figure 4. Protonation/deprotonation was negligible in the case of dT binding to the free enzyme (Figure 4 line B). Hence $\Delta H_{\text{obs}} = \Delta H_{\text{bind}}$ for reaction i. An uptake of 0.31 protons was observed with ATP binding to the TK:dT complex in reaction ii (Figure 4 line A). Titration with the 1:1 mixture of dT and ATP lead to the uptake of 0.35 protons (Figure 4 line C). It follows that proton uptake occurred with ATP binding but not with dT binding. Heat changes from ITC experiments were corrected accordingly (Table 1).

Phosphate buffer behaved anomalously and influenced the thermodynamic parameters significantly. Therefore, data collected in phosphate buffer were not included in the analysis because of obvious differences in the interaction mechanism.
Figure 5: Temperature dependence of $\Delta H$ (triangles), -$T\Delta S$ (squares) and $\Delta G$ (diamonds). ITC experiments were performed in Tris/HCl, pH 7.5. A: Binding of dT to the free enzyme (reaction i). B: Binding of ATP to the enzyme pre-saturated with dT (reaction ii). C: Binding of a 1:1 mixture of dT and ATP to the free enzyme. $\Delta C_p$ was obtained from the slope of a linear least-squares fit to $\Delta H$ data.
1.3.4. Temperature Dependence of the Thermodynamic Parameters

ITC measurements were performed at 10, 15, 20 and 25°C. The results are presented in Figure 5 and are summarized in Table 1. $\Delta H$ and $T\Delta S$ depended strongly on temperature while $\Delta G$ was almost insensitive to temperature due to enthalpy-entropy compensation. Values of $\Delta C_p$ were calculated from the slopes of the regression lines of $\Delta H_{\text{bind}}$ versus temperature (Figure 5). Binding of dT to the free enzyme (reaction i) was characterized by $\Delta C_p = -360$ cal K$^{-1}$ mol$^{-1}$. $\Delta C_p$ for ATP binding to the TK:dT complex (reaction ii) was measured $-140$ cal K$^{-1}$ mol$^{-1}$, and $\Delta C_p = -510$ cal K$^{-1}$ mol$^{-1}$ for the titration of the enzyme with a 1:1 mixture of dT and ATP (Table 1). The latter value was very close to the sum of $\Delta C_p$'s of reactions i and ii, in agreement with a thermodynamic cycle described by the three reactions.

1.3.5. Decomposition of Experimentally Observed Entropy Changes.

The overall entropy change of a binding reaction, $\Delta S_{\text{tot}}$, can be decomposed into discrete terms (Murphy et al., 1995):

$$\Delta S_{\text{tot}} = \Delta S_{\text{solv}} + \Delta S_{\text{tr/rot}} + \Delta S_{\text{conf}}$$

(2)

The solvation term, $\Delta S_{\text{solv}}$, results primarily from burial of apolar and polar surface area at the complex interface and can be approximated by

$$\Delta S_{\text{solv}} = \Delta C_p n(T/T_s^*)$$

(3)

where $\Delta C_p$ is the experimentally observed heat capacity change and $T_s^* = 385$ K (112°C) is the temperature at which the entropy of polar and apolar hydration is close to zero (Baldwin, 1986; Murphy et al., 1990; Murphy et al., 1994). The translational/rotational contribution to the entropy change, $\Delta S_{\text{tr/rot}}$, accounts for the reduction of the number of particles in solution and for the reduction in the motional degrees of freedom of the molecules. Although estimates of $\Delta S_{\text{tr/rot}}$ vary (Finkelstein & Janin, 1989; Murphy et al., 1993), empirical and theoretical considerations suggest that $\Delta S_{\text{tr/rot}}$ contributes $-4$ to $-10$ cal K$^{-1}$ mol$^{-1}$ to $\Delta S_{\text{tot}}$ of a bimolecular binding reaction (Amzel, 1997; Murphy et al., 1994; Tamura & Privalov, 1997; Yu et al., 1999). This
value is numerically close to the cratic entropy (Kauzmann, 1959), even though there is no obvious physical reason to equate the translational/rotational entropy with the statistical part of the mixing entropy (Gilson et al., 1997; Holtzer, 1995). Finally, the conformational entropy, $\Delta S_{\text{conf}}$, sums up contributions from freezing of bonds in protein side chains and ligand molecules as well as from structural rearrangements in the protein and the ligand. Since $\Delta S_{\text{tot}}$ is experimentally accessible and since $\Delta S_{\text{solv}}$ and $\Delta S_{\text{int}}$ can be estimated, $\Delta S_{\text{conf}}$ can be calculated from equation 2.

The results of this decomposition of $\Delta S_{\text{tot}}$ are summarized in Table 2. The favorable solvent contribution $\Delta S_{\text{solv}}$ is overcompensated by the large unfavorable conformational entropy change $\Delta S_{\text{conf}}$. The contribution by $\Delta S_{\text{int}}$ is unfavorable yet small. In accordance with a thermodynamic cycle, the decomposed entropy changes of reactions i and ii add up to the changes calculated for the titration of the enzyme with the 1:1 mixture of dT and ATP.

### Table 2. Decomposition of entropy changes for substrate binding to HSV1 TK at 25°C.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>$\Delta S_{\text{tot}}^a$</th>
<th>$\Delta S_{\text{solv}}^b$</th>
<th>$\Delta S_{\text{conf}}^c$</th>
<th>$\Delta S_{\text{int}}^d$</th>
</tr>
</thead>
<tbody>
<tr>
<td>TK + dT, reaction i</td>
<td>-39.9</td>
<td>92</td>
<td>-124</td>
<td>-8</td>
</tr>
<tr>
<td>TK:dT + ATP, reaction ii</td>
<td>-16.1</td>
<td>36</td>
<td>-44</td>
<td>-8</td>
</tr>
<tr>
<td>TK + dT/ATP</td>
<td>-54.7</td>
<td>133</td>
<td>-172</td>
<td>-16</td>
</tr>
</tbody>
</table>

$^a$ In cal K$^{-1}$ mol$^{-1}$. $^b$ Calculated from values of $T\Delta S$ in Table 1. $^c$ Calculated from equation 3 using $\Delta C_p$ from Table 1. $^d$ Calculated from equation 2. $^e$ Value for bimolecular and trimolecular reaction, respectively (Kauzmann, 1959).

### 1.3.6. Correlation Between $\Delta C_p$ and Surface Area Buried on Substrate Binding

In protein folding, the changes in enthalpy, entropy and heat capacity can be accounted for in terms of changes in solvent-accessible polar and apolar surface area (Makhatadze & Privalov, 1993; Murphy et al., 1992; Murphy & Freire, 1992; Privalov & Gill, 1988; Privalov & Makhatadze, 1993; Spolar et al., 1989; Spolar et al., 1992; Xie & Freire, 1994). Since changes in the atomic interactions are similar in protein folding and in binding reactions involving proteins (Janin & Chothia, 1990), thermodynamic parameters of protein-protein, peptide-protein and small ligand-protein interactions can be related to changes in solvent-accessible surface in the
same way (Baker & Murphy, 1998; Luque & Freire, 1998). \( \Delta C_p \) and \( \Delta H_{\text{bind}} \) scale with the amount and type of surface buried in the complex according to:

\[
\Delta C_p = \Delta c_{ap} \Delta A S A_{ap} - \Delta c_{pol} \Delta A S A_{pol}
\]

\[
\Delta H_{\text{bind}}(60^\circ\text{C}) = \Delta h_{ap} \Delta A S A_{ap} - \Delta h_{pol} \Delta A S A_{pol}
\]

\( \Delta A S A \) is the apolar (ap) and polar (pol) surface buried in the complex, and \( \Delta c \) and \( \Delta h \) are the elementary contributions per \( \AA^2 \) of apolar and polar surface to the heat capacity and enthalpy changes (Baker & Murphy, 1998). \( \Delta A S A_{ap} \) and \( \Delta A S A_{pol} \) were calculated by simultaneous solution of equations 4 and 5 using the experimentally determined parameters \( \Delta C_p \) and \( \Delta H_{\text{bind}} \) from Table 1. The buried areas thus calculated are shown in Table 3. The total area apparently buried (\( \Delta A S A_{\text{tot}} \) in Table 3) was \(-3250 \AA^2\) for dT binding to the apo-enzyme (reaction i), and about half as much, \(-1700 \AA^2\), for ATP binding to the TK:dT complex (reaction ii). \( \Delta A S A_{\text{tot}} \) for the reaction with a 1:1 mixture of dT and ATP was \(-5100 \AA^2\), the sum of the surface changes for reactions i and ii and in accord with a thermodynamic cycle.

**Table 3**: Changes of solvent accessible surface area (\( \Delta A S A \)) caused by substrate binding to HSV1 TK.

<table>
<thead>
<tr>
<th>Calculation</th>
<th>Reaction</th>
<th>( \Delta A S A_{ap} )</th>
<th>( \Delta A S A_{pol} )</th>
<th>( \Delta A S A_{\text{tot}} )</th>
<th>( \Delta C_p )(^d)</th>
<th>( \Delta C_p )(^e)</th>
</tr>
</thead>
<tbody>
<tr>
<td>rigid body (^b)</td>
<td>TK + dT</td>
<td>-400</td>
<td>-200</td>
<td>-600</td>
<td>-130</td>
<td>-360</td>
</tr>
<tr>
<td></td>
<td>reaction i</td>
<td>-1750</td>
<td>-1500</td>
<td>-3250</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rigid body (^b)</td>
<td>TK:dT + ATP</td>
<td>-300</td>
<td>-450</td>
<td>-750</td>
<td>-20</td>
<td>-140</td>
</tr>
<tr>
<td></td>
<td>reaction ii</td>
<td>-850</td>
<td>-850</td>
<td>-1700</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rigid body (^b)</td>
<td>TK + dT/ATP (^f)</td>
<td>-700</td>
<td>-650</td>
<td>-1350</td>
<td>-150</td>
<td>-510</td>
</tr>
<tr>
<td></td>
<td>TK + dT/ATP (^c)</td>
<td>-2650</td>
<td>-2450</td>
<td>-5100</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) \( \Delta A S A \) in \( \AA^2 \), \( \Delta C_p \) in cal K\(^{-1}\) mol\(^{-1}\). \(^b\) \( \Delta A S A \) calculated by removing dT, ATP, or both, from the crystal structure of the TK:dT:ATP complex (Wild et al., 1995). \(^c\) \( \Delta A S A \) calculated by simultaneously solving equations 4 and 5 using experimental values of \( \Delta C_p \) and \( \Delta H_{\text{bind}} \) from Table 1. \(^d\) Calculated from equation 4 with \( \Delta A S A \) calculated from the crystal structure of TK:dT:ATP (rigid body assumption). \(^e\) From Table 1. \(^f\) Reaction with 1:1 mixture of dT and ATP.
Since there is no structural data available for HSV1 TK in the ligand-free state, \( \Delta \text{ASA}_{\text{ap}} \) and \( \Delta \text{ASA}_{\text{pol}} \) calculated from \( \Delta H_{\text{bind}} \) and \( \Delta C_p \) cannot be tested against actual structural data. As an approximation, one may remove dT and ATP from the crystal structure of the ternary complex to obtain \( \Delta \text{ASA} \). The buried surface area calculated in this way corresponds to a rigid body binding model in which no conformational changes take place when dT and ATP bind. As seen in Table 3, \( \Delta \text{ASA}_{\text{tot}} \) calculated for the rigid body model was very much smaller, of the order of minus 600-800 Å² for both dT binding and ATP binding. Values of \( \Delta C_p \) calculated by equation 4 using the rigid body values of \( \Delta \text{ASA} \) were accordingly smaller than \( \Delta C_p \) determined by experiment (\( \Delta C_p_{\text{calc}} \) and \( \Delta C_p_{\text{exp}} \) in Table 3).

1.4. Discussion

1.4.1. Ordered Binding of Thymidine and ATP to HSV1 TK

The calorimetric titrations presented here provide a comprehensive description of the energetics of substrate binding to HSV1 TK. Substrate binding is strictly ordered. ATP could bind only after the TK:dT complex had been formed. No ATP binding to the apo-enzyme was observed. One might argue that ATP binding was entropy-driven and that \( \Delta H \) for ATP binding was so small that it escaped detection by ITC. This possibility could be ruled out because the entire set of thermodynamic parameters satisfied the cycle shown in Figure 6. Figure 6 summarizes the energetics of binding of dT and ATP at 25°C and pH 7.5. Summation of the parameters for reactions i and ii equaled almost exactly the parameters determined independently for the single reaction of the apo-enzyme with dT and ATP together. The free energy change for dT binding was smaller than for ATP binding, -7.2 versus -9.0 kcal mol\(^{-1}\), but the enthalpy, entropy and heat capacity changes were significantly larger for the initial dT binding. Comparing absolute values one notes that \( \Delta H \) was 1.4 times larger, and \( T \Delta S \) and \( \Delta C_p \) were 2.5 times larger for dT binding than for the subsequent ATP binding. This means that dT binding, which induced the formation of a tight ATP binding site, was driven by a large enthalpy change but at a high cost in entropy.
1.4.2. Unfavorable Entropy Change Originates from Unfavorable Conformational Component

The total entropy change can be decomposed into three terms accounting for solvent reorganization at the molecular surface (ΔS_{solv}), for changes in rotational and translational degrees of freedom of the reacting molecules (ΔS_{rot}), and for conformational events accompanying the formation of the enzyme-substrate complexes (ΔS_{conf}) (Table 2). Solvent reorganization provided a substantial gain in binding entropy due to water release, about 90 cal K^{-1} mol^{-1} for dT binding and about 35 cal K^{-1} mol^{-1} for ATP binding. Since kinases transfer phosphoryl groups from a donor (mostly ATP) to an acceptor OH-group, hydrolysis by competing water must be avoided. Therefore, the release of water from the active site is expected and in agreement with considerable ΔS_{solv}. After subtracting from ΔS_{tot} the favorable ΔS_{solv} term and the small unfavorable cratic ΔS_{rot}, there remained a large unfavorable entropic contribution ΔS_{conf}, which had to be accounted for by conformational changes.

ΔS_{conf} may be thought to have two main origins. One is "freezing" of bond rotations. Amino acid side chains directly contacting the bound substrate are less mobile. Also bonds that are not directly contacting the substrate may become less mobile when
substrate binding induces a more closed or a more compact conformation of the enzyme. The other contribution to $\Delta S_{\text{conf}}$ is partial folding or tightening of domains, particularly in the area of substrate binding. Interdomain movements within the enzyme dimer also can contribute to $\Delta S_{\text{conf}}$ if the HSV1 TK dimer becomes more closed in the substrate-bound form.

We tried to estimate the contribution to $\Delta S_{\text{conf}}$ by side chain immobilization using an algorithm based on the parameterization of a large set of protein structures (Baker & Murphy, 1998). The calculation sums up the loss in side chain entropy for every type of side chain between a freely mobile peptide and a tightly folded protein. When applied to the side chains contacting dT and ATP in the TK:dT:ATP complex, one estimates that side chain immobilization contributed approximately 5% to $\Delta S_{\text{conf}}$. Although this calculation is very approximate, it would still seem that the major contribution to $\Delta S_{\text{conf}}$ originated from conformational domain movements and partial folding, or refolding, of domains.

1.4.3. Substrate-Induced Conformational Changes Deduced from Large $\Delta C_p$-Values

In the relatively narrow temperature range studied, $\Delta H$ became more favorable and $T\Delta S$ less favorable with increasing temperature. As a result, $\Delta G$ remained remarkably insensitive to temperature through entropy/enthalpy compensation. This is an ubiquitous phenomenon seen in many association reactions and is thought to be directly related to the role of solvent water molecules in the association process (Lumry & Rajender, 1970). According to the laws of thermodynamics, the temperature dependence of $\Delta H$ and $\Delta S$ results from substantial changes in heat capacity. In almost all association processes with proteins, $\Delta C_p$ has a negative sign if the free components are the reference state (Baker & Murphy, 1998). In the present case, $\Delta C_p$ of the overall reaction was -510 cal K$^{-1}$ mol$^{-1}$. Binding of dT contributed 70% and ATP binding 30% to this large negative heat capacity change.

Since, in general, $\Delta C_p$ correlates well with the amount of surface area buried at the complex interface (Baker & Murphy, 1997; Connelly & Thomson, 1992; McNemar et al., 1997; Murphy et al., 1993), we have tried to estimate the surface area of HSV1 TK buried by dT and ATP, respectively, and by both substrates together. The
calculation (equations 4 and 5) is based on a large body of structural and thermodynamic data for protein folding, protein-protein association and protein-ligand binding. The calculated buried surface area was -3250 Å² for dT and -1700 Å² for ATP (Table 3). These figures would seem to be too large for the binding of the two small substrate molecules. Indeed, the roughly -5100 Å² of surface area that was apparently buried by both substrates together, is comparable to the surface buried on folding of a small globular protein of 50 to 60 residues.

Since the 3D-structure of substrate-free HSV1 TK is not known, we could not relate the calculated surface changes with actual structural data. However, from the crystal structure of the TK:dT:ATP complex one can calculate the buried area for the case of a rigid body binding model in which the dT and ATP binding sites pre-exist and no conformational changes take place. Obviously, this assumption was unjustified in view of the ordered sequential binding mechanism. Indeed, calculation of ΔASA for a rigid binding model yielded buried surface areas that were 5 times smaller for dT binding and 2 times smaller for ATP binding (Table 3). The heat capacity changes calculated for the rigid body binding model were accordingly much smaller.

One has to remember that surface burial calculated according to equations 4 and 5 takes into account mainly hydration effects, which are not the only contributions to ΔCp. Changes in molecular vibration modi are influenced by ligand binding in a deep hydrophobic binding pocket, as seen in the crystal structure of the TK:dT:ATP complex in which the binding sites are close to the interface between the tightly associated monomers of the functional dimer. Hence, vibrational effects may have contributed to the large negative ΔCp's in addition to hydration effects, but were not considered in the semi-empirical analysis.

Even so the calculated surface area changes may be deficient (possibly too large), the significant discrepancy to a rigid binding model is an indisputable sign for significant conformational rearrangements accompanying substrate binding to HSV1 TK. Indeed, in the crystal structure of the ternary enzyme-substrate complex, dT is deeply buried and completely caved due to closing of the LID domain (Figure 1). The ATP binding site is more surface exposed with ATP located like a plug just in front of dT (Wild et al., 1995). Thus, a large contribution by conformational rearrangements would seem to be justified. Reorganization of the dimer interface concomitant to substrate binding would also seem to be feasible. The dT binding site in the ternary
complex of HSV1 TK is localized in a large pocket of the NMP\text{bind} domain that is lined up by helices $\alpha3$, $\alpha4$ and $\alpha5$. These elements also belong to the dimer interface, and one might assume that the association of dT and the subsequent formation of the ATP binding site could be communicated to the dimer interface. In addition, parts of the CORE domain and of the mobile segment between residues 250 and 322 contribute to the NMP\text{bind} domain. While these considerations remain speculative before the structure of substrate-free HSV1 TK is known, recent crystallographic studies of adenylate kinases, to which the viral thymidine kinase is homologous, have indeed demonstrated major changes induced by substrate binding in the NMP\text{bind} domains and the LID domains, as well as small changes in the CORE domains (Muller et al., 1996; Vonrhein et al., 1995).

1.5. Conclusion

This is the first report providing a comprehensive thermodynamic description of substrate and cofactor binding to HSV TK1, a representative of the large family of nucleotide and nucleoside kinases. The results obtained by titration microcalorimetry reveal an extreme case of positive heterotropic interaction. Formation of a binary complex of thymidine with HSV1 TK is a stringent prerequisite for ATP binding. Since the ATP binding site is in fact generated by thymidine binding, one expects the enzyme to undergo considerable conformational rearrangements. This has been supported by a semi-empirical analysis of the observed heat capacity and entropy changes, which were large and negative and indicated burial of molecular surface to an extent much larger than expected if the substrate binding sites would pre-exist on the apo-enzyme and no rearrangement would occur (rigid body binding model). The favorable gain in entropy from water release from buried surface was overcompensated by a large decrease in conformational entropy. The findings support the view that substrate binding to HSV1 TK leads to a conformational closing of the substrate binding sites to bring thymidine and ATP into an orientation appropriate for catalysis. The details of the predicted rearrangements have to await a firm structural foundation. Work is in progress in this laboratory to solve the crystal structure of the substrate-free apo-form of HSV1 TK.
1.6. Material and Methods

1.6.1. Materials

All chemicals were of analytical grade and were used without further purification. Thymidine, ATP, glutathione, buffer reagents and glutathione agarose (SH-coupled via 12 C-spacer) were from Sigma; DTT and EDTA were from Fluka.

1.6.2. Expression and Purification

The bacterial expression vector pGEX2T-TK was constructed as described previously (Michael et al., 1997). HSV1 TK was expressed as glutathione S-transferase fusion protein (GST-TK) in E. coli strain BL21 and purified by glutathione affinity chromatography (Fetzer et al., 1994). After isolation of the fusion protein from the crude extract by glutathione-sepharose chromatography, the protein was directly (on-column) exchanged into the experimental buffer by thoroughly rinsing the column with excess of buffer. The protein was eluted by addition of 5 mM glutathione and was directly used for titration experiments. Enzyme concentration was determined using a dye-binding assay (Bradford, 1976).

1.6.3. Isothermal Titration Calorimetry

ITC experiments were carried out using an OMEGA titration microcalorimeter (Microcal Inc., Northampton, MA) equipped with a nanovolt preamplifier to reduce electrical noise (Wiseman et al., 1989). The reference cell was filled with water containing 0.01% sodium azide, and the calorimeter was calibrated using standard electrical pulses as recommended by the manufacturer. All solutions were degassed for 10 min with gentle stirring under vacuum. Solutions of the fusion protein were filled in the sample cell (1.34 ml volume) and titrated with dT or ATP. Substrate solutions were prepared in the buffer from the final step of protein purification. The substrate concentration in the injection syringe was usually 25 times higher than the concentration of protein binding sites. A typical experiment consisted of a first control injection of 1 μl followed by 19 injections, each of 4 μl and 15 seconds duration, with 4 minutes interval in between.

ITC measurements were routinely performed in 50 mM Tris, pH 7.5, 4 mM EDTA (to suppress enzymatic activity), 5 mM glutathione, 1 mM DTT. Heat contributions due to
coupled protonation events upon binding were evaluated by calorimetric experiments in various buffers of different ionization enthalpies under otherwise the same conditions. The buffers and their ionization enthalpies (in kcal mol\(^{-1}\) at 25°C) were as follows: PIPES (2.7) MOPS (4.9), TES (7.7), Tris (11.34) (Cooper & Johnson, 1994b). The pH of the buffer was adjusted at the experimental temperature. Buffer concentrations were 50 mM and the ionic strength where the same for all buffers. In control experiments, the ligand was injected into buffer. The observed heat effects were concentration-independent and were identical to the heat signals detected after complete saturation of the protein. Therefore, the non-specific background was usually estimated by averaging the small heats at the end of the titration. Raw data were collected, corrected for ligand heats of dilution and integrated using the Microcal Origin software supplied with the instrument. Since protein concentration was expressed on a subunit basis, a single-site binding model was fit to the data by a non-linear regression analysis to yield binding constants (K\(_b\)), enthalpies of binding (\(\Delta H\)) and stoichiometry of binding.

1.6.4. HPLC Assay

High performance liquid chromatography was used for concentration determination of thymidine and ATP in the final ligand solutions, and to monitor potential phosphorylation products during calorimetric experiments using a modified protocol of a previously published method (Masson et al., 1993). Nucleotides were determined by reverse-phase ion-paired chromatography using a C\(_{18}\) column (LiChrospher 100 RP-18, 5 \(\mu\)m, 250x4 mm, Merck) in 0.2 M NaH\(_2\)PO\(_4\), 25 mM tetrabutylammonium, 3% (v/v) methanol at 1.0 ml/min and detection at 254 nm. Ligand concentrations were calculated by means of calibration curves from standard solutions showing linearity in the range of <0.04 to at least 2 mM. The detection limit for dT, dTMP and ATP was < 20 nmol (Pilger et al., 1999).

1.6.5. Calculation of Solvent Accessible Surface Area

The program NACCESS (Hubbard & Thornton, 1996), an implementation of the Lee and Richards solvent-accessibility algorithm (Lee & Richards, 1971), was used with a probe radius of 1.4 Å and a slice width of 0.25 Å. Calculations from experimental values of \(\Delta C_p\) and \(\Delta H_{bind}\): \(\Delta S_{A,p}^a\) and \(\Delta S_{A,p}^p\) were calculated by simultaneous
solution of equations 4 and 5, using $\Delta c_{ap} = 0.43$, $\Delta c_{pol} = -0.26$, $\Delta h_{ap} = -7.26$ and $\Delta h_{pol} = 29.14$ [all in units of cal K$^{-1}$ (Å$^2$ mol)$^{-1}$] (Baker & Murphy, 1998). Calculations from crystal structure: $\Delta \text{ASA}_{ap}$ and $\Delta \text{ASA}_{pol}$ were calculated using the coordinates of the ternary complex TK:dT:ATP (Wild et al., 1995) and removing either dT or ATP or both.

The contribution of side-chain freezing was calculated according to Baker & Murphy (1998).

1.6.6. Acknowledgements

This work was supported in part by grant 31-45556.95 (to H. R. B.), Kanton of Zurich.

1.7. References


Spolar, R. S., Livingstone, J. R. & Record, M. T., Jr. (1992). Use of liquid hydrocarbon and amide transfer data to estimate contributions to


Seite Leer / Blank leaf
Substrate Diversity of Herpes Simplex Virus Thymidine Kinase - Impact of the Kinematics of the Enzyme


*From the Department of Pharmacy, Swiss Federal Institute of Technology (ETH), Winterthurerstr. 190, CH-8057 Zürich, Switzerland and the #INFM, Instituto Nazionale per la Fisica della Materia and ##SISSA/ISAS, International School for Advanced Studies, Via Beirut 4, 34014 Trieste, Italy

Abbreviations

HSV 1, herpes simplex virus type 1
TK, thymidine kinase
dT, thymidine
dC, deoxycytidine
dCK, deoxycytidine kinase
dTMP, thymidine monophosphate
TmpK, thymidylate kinase
ACV, acyclovir; GCV, ganciclovir
PCR, polymerase chain reaction
GST, glutathione S-transferase
DTT, DL-dithiothreitol
IPTG, isopropyl-ß-D-thiogalactopyranoside
PMSF, phenylmethylsulfonyl fluoride
PAGE, polyacrylamide gel electrophoresis
HPLC, high performance liquid chromatography
AZT, 3'-azido-deoxythymidine
DFT, density functional theory
CD, circular dichroism
2.1. Summary
Herpes simplex virus type 1 thymidine kinase (HSV 1 TK) exhibits an extensive substrate diversity for nucleobases and sugar moieties in contrast to other TKs. This substrate diversity is the crucial molecular basis of selective antiviral and suicide gene therapy. The mechanisms of substrate binding of HSV 1 TK were studied by means of site-directed mutagenesis combined with isothermal calorimetric measurements and guided by theoretical calculations and sequence comparison. The results show the link between the exceptionally broad substrate diversity of HSV 1 TK and the presence of structural features such as the residue triad H58/M128/Y172. The mutation of Met 128 into a Phe as well as the double mutant M128F Y172F result in mutants that have lost their activity. However, by exchanging His to form the triple mutant H58L/M128F/Y172F the enzyme regains activity. Strikingly, this triple mutant becomes resistant towards acyclovir. Furthermore, we give evidence for the importance of Glu 225 of the flexible LID region for the catalytic reaction. The presented data give new insights to understand mechanisms ruling substrate diversity and thus are crucial for both the development of new antiviral drugs and engineering of mutant TKs apt to accept novel substrate analogs for gene therapeutic approaches.

2.2. Introduction
Herpes Simplex virus type 1 thymidine kinase (HSV 1 TK) is a multifunctional enzyme that possesses kinase activities normally performed by three separate cellular enzymes. It phosphorylates thymidine (dT), which is then transformed by cellular kinases to the triphosphorylated DNA building block, and deoxyuridine (dU), both reactions comparable with the function of human cellular TK (hTK). Further, it converts deoxycytidine (dC) to dCMP as does human deoxycytidine kinase (dCK) and phosphorylates thymidylate (dTMP) as does human TMP kinase (hTmpK) (Chen & Prusoff, 1978; Chen et al., 1979; Koonin & Senkevich, 1992). Moreover, unlike its cellular counterpart hTK, HSV 1 TK is able to phosphorylate pyrimidine as well as purine analogs and discloses low stereochemical demands for the ribose moiety as it also accepts acyclic side chains as phosphorylgroup-acceptors e.g. (Cheng et al., 1983; Elion et al., 1977; Keller et al., 1981). These differences in substrate diversity
are the crucial molecular basis for the selective treatment of viral infections. Nowadays, the most widely used therapeutic compounds to interfere with a severe HSV 1 infection are the purine analogs acyclovir (ACV), penciclovir and their prodrugs valaciclovir and famciclovir, respectively. They require HSV 1 TK to be efficiently activated in order to block virus proliferation by inhibition of viral DNA polymerase. HSV 1 TK is the key enzyme in this antiviral strategy. In gene therapy of cancer (Culver et al., 1994; Tong et al., 1997) and AIDS (Caruso & Bank, 1997) HSV 1 TK is used as suicide enzyme in combination with the purine analog ganciclovir (GCV). Another important application is the employ of HSV 1 TK as a rescue system in allogeneic bone marrow transplantation (BMT) induced graft versus host disease (Bonini et al., 1997). In addition to the significance from a therapeutic point of view, HSV 1 TK seems to be important for the reactivation of the virus from lifelong latent infection in neuronal ganglia (Coen et al., 1989; Efstathiou et al., 1989; Jacobson et al., 1998). However, there is evidence that human TK can functionally replace viral TK in terms of reactivation of the virus from latency (Chen et al., 1998).

There are no recognizable sequence similarities between HSV 1 TK and hTK (Bradshaw & Deininger, 1984). Rather, sequence alignments have detected similarities between herpesvirus TKs and hdCK (Harrison et al., 1991) and to a lesser extent cellular TMPK (Robertson & Whalley, 1988). Despite the limited sequence homology with enzymes of the nucleotide kinases (NK) family, HSV 1 TK shares structural features comprising a parallel five stranded β sheet and a glycine rich loop common to all NK. In the crystal structure, HSV 1 TK is a homodimeric enzyme with 376 amino acids per subunit (Brown et al., 1995; Wild et al., 1995; Wild et al., 1997). The two subunits are related by C2-symmetry. The active site is formed by an ATP- and a nucleoside-binding region. The visual representation of the thymidine binding site is depicted in Fig. 1, featuring a complex hydrogen-bond network within the active site. The thymine ring makes pairwise hydrogen-bond interaction via its 4-carbonyl and 3-NH group with the amide group of the highly conserved Gln 125 and is hydrogen-bonding with Arg 176 by means of two ordered water molecules. Moreover, the pyrimidine ring of thymidine is fixed between Met 128 and Tyr 172 forming a sandwich-like complex. His 58 and Arg 163 both interact with the hydroxyl group of Tyr 172 sealing the position of tyrosine. The deoxyribose makes hydrogen-bond interaction via its 3'-OH with Tyr 101 and the highly conserved Glu 225 and via its 5'-OH with Glu 83. Glu 225 belongs to the “LID” domain, a region rich in lysine and
Fig. 1. **Representation of a portion of the active site of HSV 1 TK with bound thymidine.** The position and geometry of dT and the amino acids that are directly involved in substrate binding are shown as capped sticks and are labeled (Wild et al., 1995). The secondary structure of the protein is displayed as tubes. The hydrogen-bond mediating water molecules are presented as small spheres and hydrogen-bonds are displayed as dashed lines. The LID region and P-loop (glycine-loop) are indicated. The figure has been prepared using the program SYBYL V.6.3 (Tripos Associates). The coordinates in PDB format are indexed as 2VTK in the Brookhaven protein structure database.

arginine residues which appears to be able to form a flap that encloses the active site. This LID region is expected to undergo conformational changes upon substrate binding and therefore influencing the catalytic phosphorylation rate, similar to ADK, with which TK shares similar three-dimensional features (Muller et al., 1996).

Up to now, various studies e.g. (Balasubramaniam et al., 1990; Liu & Summers, 1988; Munir et al., 1992; Munir et al., 1994) tried to elucidate the role and functionality of the amino acid residues in HSV TK and with resolving of the crystal structure (Brown et al., 1995; Wild et al., 1995; Wild et al., 1997) pivotal supplementary information became accessible. For example, these structural information allowed to render some mechanisms responsible for development of herpesviral resistance, an increasing problem in clinic in the treatment of
immunocompromised patients (Darby et al., 1981; Englund et al., 1990; Gaudreau et al., 1998), comprehensible (Kussmann-Gerber et al., 1998). Despite the increased structural knowledge, the basis of the molecular difference in substrate and drug specificity of HSV TK and the particular role of the LID region still remain unclear. This work reports the study on the nature of mechanism of binding of HSV 1 TK by means of site-directed mutagenesis combined with isothermal calorimetric measurements, and guided by ab initio calculations and sequence comparison. It shows the link between the broad substrate diversity of HSV 1 TK and the presence of structural features such as the residue triad H58/M128/Y172 which is thought to confer distinctive binding of an exceptionally large variety of substrates to the HSV 1 TK and to guide the catalytic properties. Furthermore, we give evidence for the importance of the flexible LID domain for enzyme function.

2.3. Experimental Procedures

2.3.1. Materials

(Methyl-1\(^\text{1}\), 2\(^\text{1}\)-\(^\text{3}\)H) thymidine (3 TBq/mmol) was obtained from Amersham Life Science and (sidechain-2\(^\text{1}\)-\(^\text{3}\)H) acyclovir (1.2 TBq/mmol) from NEN DuPont. (Methyl-\(^\text{3}\)H)-3\(^\text{1}\)-Azido-deoxythymidine (0.2 TBq/mmol) was purchased from Moravek Biochemicals. Nucleotides and AmpliTaq Gold\(^\text{TM}\) polymerase were bought 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\(^\text{R}\)endA1 recA1 rel A1 gyrA96 thi-1 hsdR17 supE44 lacZ M15 F\(^\text{−}\)λ\(^\text{−}\)) (Clontech) was used for all cloning steps. Strain BL21 (ompT, \(\text{F}^\text{−}\), hsdS (rB\(^\text{−}\),mB\(^\text{−}\)), gal) (Pharmacia) served as host for expression. The plasmid pGEX-2T was purchased from Pharmacia. The plasmid pBR322-TK containing the gene for HSV 1 strain F TK was a gift from S. McKnight. The expression vector pGEX-2T-TK was constructed as described earlier (Michael et al., 1997).

2.3.2. Mutagenesis

Site-directed mutagenesis was performed by using oligonucleotide-directed polymerase chain reaction based on a three primer method (Barettino et al., 1994; Steinberg & Gorman, 1994). The primers were ordered, synthesized and purified at
Microsynth (Balgach, CH). Briefly, in the first PCR reaction, bacteriophage M13mp18 containing the BamHI-KpnI fragment of HSV 1 TK was amplified using the respective antiparallel mutagenic primer (H58L: 5'-GAC GGT CCC CTC GGG ATG GG-3', M128I: 5'-CAG ATA ACA ATC GGC ATG CC-3', M128A: 5'-GCG CCC AGA TAA CAG CGG GCA TGC CTT ATG C-3', M128F: 5'-GCG CCC AGA TAA CAT TCG GCA TGC CTT ATG C-3', Y172F: 5'-CTG TGC TTC CCG GCC G-3') and the M13mp universal primer (5'-GCT ATG ACC ATG TTA CG-3'). The resulting amplification products were gel-purified and subsequently used as megaprimer. In the second PCR reaction, the isolated megaprimer was hybridized to pGEX-2T-TK and extended within a single PCR cycle. Then, the flanking M13mp universal primer and pGEX-2T universal primer (5'-GGG CTG GCA AGC CAC GTTTGG TG-3') were added to the PCR tubes and 30 PCR cycles were performed. In a last step, each of the fragments containing the desired mutation was cloned into the expression vector pGEX-2T-TK by digestion with the restriction enzymes BamHI and KpnI, and subsequent gel purification and ligation. Additional restriction steps with BamHI/SacI and BamHI/AccI respectively, were necessary to obtain the mutants M128FY172F, H58LM128FY172F and H58LM128F.

For the mutant E225L we used a four primer based PCR method described by Higuchi (Innis et al., 1990). Basically, two primary PCR reactions are performed separately. The mutation is introduced as part of the respective mutagenic inside-primers (forw. 5'-CCC GGG CCT GCG GCT GGA CC-3'; rev. 5'-GGT CCA GCC GCA GGC CGG G-3') each of which is amplified with a suitable outside primer. The two products overlap in sequence; both contain the same mutation. After gel purification, these overlapping primary products were denatured and allowed to reanneal together, producing two possible heteroduplex products. The subsequent reamplification of one of these products with only the right- and leftmost ("outside-") primers resulted in the enrichment of the full-length, secondary product which was then introduced into the vector pGEX-2T-TK replacing the respective wildtype fragment by Smal/SacI restriction.

2.3.3. Sequence Verification

Competent E. coli DH5α were transfected with the mutated pGEX-2T-TK DNA. After DNA isolation of several clones, we sequenced the entire TK gene of the respective
mutant progeny, using the dye terminator method (ABI PRISM™ 310) to verify that the targeted mutation and no frameshift or additional mutation had occurred.

2.3.4. Expression and Purification of the Mutant HSV1 TKs

Competent *E. coli* BL21 were transformed with the vector pGEX-2T-TK containing the respective mutated full-length *tk* gene as GST-fusion protein. Protein expression was induced by the addition of 0.2 mM IPTG (isopropyl-β-D-thiogalactopyranoside). After 20 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 DNasel (10 mM MgCl₂, 1 mM MnCl₂ and 10 mM EDTA for inactivation of DNasel afterwards) for 30 min at 4°C and by additional sonication for 3 min. The lysate was clarified by centrifugation at 12,000xg for 20 min and applied to a one step glutathione agarose purification procedure and subsequent on column thrombin cleavage as described (Fetzer et al., 1994). Purification was monitored by SDS-PAGE and led to a > 90% pure thrombin-cleaved protein, which was directly used for kinetic studies. Total protein concentration was measured using the Bio-Rad Protein Assay (Bio-Rad).

2.3.5. Thymidine, Acyclovir and AZT Kinetics

Kinetic studies measuring the conversion of labeled substrate to substrate monophosphate were performed using the DEAE-cellulose method as described earlier (Furlong, 1963; Gerber & Folkers, 1996). Reactions were carried out in a final volume of 30 μl containing 50 mM Tris pH 7.2, 5 mM ATP, 5 mM MgCl₂, 1.5 mg/ml BSA. The amount of enzyme and concentrations of (³H)-substrate were chosen in consideration of Michaelis Menten conditions for initial velocity measurements. The *Kₘ* and *Vₘₐₓ* values have been determined by non-linear fit of the raw data to the Michaelis Menten equation using the Microcal Origin software. *kₗₐₜ* values were determined by dividing *Vₘₐₓ* by the enzyme concentration. These values were measured based on at least three independent assays.
2.3.6. Spectrophotometric Assay for Thymidine Kinase Activity

An UV-spectrophotometric test was employed to monitor ADP formation during substrate phosphorylation. Enzyme activity was measured using a lactate dehydrogenase - pyruvate kinase coupled assay (Keller et al., 1981). The change in $A_{340}$ was recorded over time by analyzing mutant TKs and different substrates.

2.3.7. HPLC Assay

High performance liquid chromatography was applied to monitor ADP and dTMP, ACVMP or AZTMP formation during substrate phosphorylation with ion-pair chromatography using a modified version of the previously published protocol (Masson et al., 1993) (Column: RP-18; Solvent: 0.2 M NaH$_2$PO$_4$, 25 mM Tetrabutylammoniumhydrogensulfate, 3% Methanol; Flow: 1.1 ml/min; Detection: UV 254 nm). This method was applied to check those mutants that could not be measured either by the radioactive or UV spectrometric approach. Reactions were carried out in a final volume of 75 µl containing 50 mM Tris pH 7.2, 5 mM ATP, 5 mM MgCl$_2$, 2 mM thymidine (ACV/AZT) and 1 µg to 5 µg thymidine kinase. The reaction was stopped after one hour at 37°C by a ten fold dilution in water and freezing at -20°C prior to injection. The formation of the nucleotide monophosphate was monitored qualitatively. Two different blank reactions (no enzyme or no substrate) were run concurrently to account for the occurring minimal reaction independent ATP hydrolysis. The detection limit for phosphorylated substrate lies under 20 nmol making this method even more sensitive than the UV-assay.

2.3.8. Titration Calorimetry

To evaluate the binding affinity of the less active mutants, isothermal titration microcalorimetry (ITC) was carried out. For stability reasons all measurements were performed with the GST-fusionprotein since the biochemical properties are identical to thymidine kinase (Fetzer et al., 1994). After isolation of the fusionprotein from the crude extract onto the glutathione sepharose, the protein was directly (on-column) exchanged into the measuring buffer (50 mM Tris/HCl, pH 7.5, 4 mM EDTA, 5 mM glutathione, 1 mM DTT and 1 mM ATP) by thoroughly rinsing the column with excess amount of buffer. The purified protein was eluted by addition of 5 mM glutathione into
the buffer and was directly used for titration experiments. The enzyme concentration was determined using Bio-Rad Protein Assay (Bio-Rad) and was corrected for impurities detected by SDS-PAGE and quantified by gel densitometry. This purification protocol usually let to a purity of 70-80% of the fusion protein.

Isothermal titration microcalorimetry was performed employing an OMEGA Microcalorimeter from Microcal, Inc. (Northampton, MA), with a cell volume of 1.3338 ml and using a 100 µl microsyringe while stirring at 375 rpm. The calorimeter and the equations used to fit calorimetric data have been described in detail previously (Wiseman et al., 1989). The reference cell was filled with water containing 0.01% sodium azide, and the instrument was calibrated with standard electrical pulses. All ITC measurements were performed at 25°C in 50 mM Tris/HCl, pH 7.5 (at 25°C), 4 mM EDTA, 5 mM glutathione, 1 mM DTT and 1 mM ATP. Prior to loading into the microcalorimeter, all solutions were degassed for 10 min with gentle swirling under vacuum. Solutions of the fusion protein were filled in the sample cell and titrated with thymidine with a first control injection of 1 µl followed by 29 identical injections of 4 µl. Thymidine solutions were prepared by dissolving in the same buffer to concentrations generally 25 times higher than the protein solution. The titration experiment was designed to ensure complete saturation of the enzyme before the final injection. The heat of dilution for the ligand was concentration-independent and corresponded very well to the heat observed from the last injections after the protein was saturated. Therefore, the baseline of the titrations could usually be well estimated from the last injections of the titration. No interference of spontaneous DTT oxidation with measurements was observed. Data were collected, corrected for ligand heats of dilution and deconvoluted using the Microcal Origin software supplied with the instrument to yield binding constants (Kₐ) and enthalpies of binding (ΔH). The thermodynamic parameters were calculated from the basic equations of thermodynamics: ΔG = ΔH - TΔS = -RTlnKₐ, where ΔG, ΔH and ΔS are the changes in free energy, enthalpy and entropy of binding, respectively.

2.3.9. Sequence Alignments

Sequences were taken from the EMBL data base and were aligned with the aid of the program multAlin (Corpet, 1988).
2.4. Results

2.4.1. Sequence Alignments

Two sequence motifs (GXXGXGK; (F)DRH) are highly conserved among thymidine kinases. The region GXXGXGK corresponds to the glycine-rich loop (Liu & Summers, 1988), which accommodates the ATP-phosphate and the (F)DRH motif (Folkers & Trumpp, 1987) is located in the five stranded beta sheet core of the protein (Brown et al., 1995; Wild et al., 1995; Wild et al., 1997). Based on the sequence alignment in Fig. 2 we analyzed the residues which are involved in thymidine fixation and in deoxyribose orientation (Glu 225) (see Fig. 1). Gln 125, Arg 163 and Glu 225 are highly conserved in all TKs. The succession of His 58, Met 128 and Tyr 172 in TKs is so far known for four viral strains, namely HSV 1, HSV 2, MHV (Fig. 2) and BHV. Instead, in most other TKs the combination X58/F128/F172 is found where X can be any hydrophobic amino acid and not histidine. In all the studied TKs, a Tyr at position 172 is never combined with a Phe at position 128 (HSV 1 TK numbering).

![Fig. 2. Multiple alignment of amino acid sequences of HSV 1 and related virus strains thymidine kinases. Residues of interest are marked using numbering of HSV 1 TK. Selected excerpts of TKs from herpes simplex virus type 1.2 (HSV 1.2), marmoset herpes virus (MHV), equine herpes virus type 4 (EHV), varicella-zoster virus (VZV) and Epstein-Barr virus (EBV) are displayed. Asterisks denote positions that appear to be completely conserved in all herpesviral TK sequences. Sequences were taken from EMBL data base.](image-url)
2.4.2. Substrate Diversity and Induced Fit

To address mechanisms guiding substrate specificity and the role of amino acids in substrate fixation within the active site of HSV 1 TK, we studied eight different mutants of HSV 1 TK. Moreover, the role of the negatively charged Glu 225 sitting within the supposedly moving, otherwise positively charged LID region, was studied by means of an additional mutant. The purification and expression scheme allowed rapid isolation of milligram amounts of wildtype and mutant HSV 1 TK enzyme. The two additional unspecific cleavage sites for thrombin within TK led to a truncated TK, lacking the N-terminal 33 amino acids that are not essential for catalytic activity (Halpern & Smiley, 1984).

Previously performed ab initio calculations, using the Carr-Parinello approach, clearly indicated that the molecular orbitals of Met 128 and Tyr 172 do not overlap with those of the substrate, nor are \( \pi-\pi \) interactions between the Tyr 172 ring and the substrate present (Alber et al., 1998). Interestingly, no polarization effect was found on the Met-sulfur atom. This indicated that sulfur has only a hydrophobic effect although it is a highly polarizable element. Instead, strong polarization effects were located on the substrate, especially on the O and N atoms. Electrostatic interactions between tyrosine and thymine can therefore play an important role in substrate fixation. In order to verify the results and predictions of the ab initio calculations and the alignments, site-directed mutagenesis at position 58, 128 and 172 respectively has been performed. The results of these experiments are summarized in Table I.

**M128A**—The replacement of methionine at position 128 by an alanine led to a loss of enzyme activity, pointing out the important role of a rather bulky residue at position 128 for substrate fixation.

**M128I**—Interestingly, the exchange of methionine by isoleucine resulted in the same activity as the wild type enzyme.

**M128F**—With phenylalanine at position 128 a completely inactive enzyme emerged. Even with the HPLC system no phosphorylated product was observed.

**Y172F**—Exchange of tyrosine 172 to phenylalanine which entails the loss of hydrogen-bonds to His 58 and Arg 163 (see Fig. 1), led to an enzyme with an altered activity profile. Under standard kinetic conditions, no explicit progressive thymidine phosphorylation was detectable but by increasing the pH of the reaction mixture, the
Table I. Catalytic properties of mutant HSV 1 TKs on Position 58, 128 and 172. The leftmost column indicates the amino acid change of each mutant in single letter code. The kinetic constants (± standard errors) derived in this study are presented in the next three columns. The rightmost column indicates the pH value above which full enzyme activity is attained. The values have been determined by non-linear fit of the raw data to the Michaelis Menten equation using the Microcal Origin software. The co-substrate ATP was kept in 10-100 fold excess with regard to its $K_m$ value ($K_m$ ATP 13 $\mu$M). The DEAE paper method (Gerber & Folkers, 1996) was applied for the kinetic measurements for wt, M128I, H58L, Y172F and the triple mutant for the latter additionally the UV-spectrometric assay. '+' indicates that less than 3% of the activity with regard to the wt enzyme was detectable, but formation of dTMP was detectable in HPLC. '-' means that no activity and no formation of dTMP could be detected with neither method.

<table>
<thead>
<tr>
<th>HSV 1 TK</th>
<th>$K_m$, dT, $\mu$M</th>
<th>$k_{cat}$ dT, s$^{-1}$</th>
<th>$K_m$, ACV, mM</th>
<th>pH starting max. activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>wild type</td>
<td>0.20 ± 0.05</td>
<td>0.35 ± 0.014</td>
<td>0.20 ± 0.05</td>
<td>6.5</td>
</tr>
<tr>
<td>M128A</td>
<td>+</td>
<td>-</td>
<td>n.d.</td>
<td>n.d.</td>
</tr>
<tr>
<td>M128I</td>
<td>0.36 ± 0.07</td>
<td>0.51 ± 0.039</td>
<td>0.51 ± 0.10</td>
<td>6.5</td>
</tr>
<tr>
<td>M128F</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Y172F</td>
<td>0.51 ± 0.23</td>
<td>0.46 ± 0.08</td>
<td>+</td>
<td>8.5</td>
</tr>
<tr>
<td>H58L</td>
<td>112 ± 21</td>
<td>6E-3 ± 2E-3</td>
<td>-</td>
<td>8.5</td>
</tr>
<tr>
<td>M128FY172F</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>H58LM128FY172F</td>
<td>183 ± 42</td>
<td>0.04 ± 5E-3</td>
<td>-</td>
<td>8.5</td>
</tr>
</tbody>
</table>

kinetics were measurable, indicating a different pH-sensitivity profile of this mutant in comparison with the wildtype enzyme. However, under the adapted conditions the $K_m$ of the mutant towards thymidine remained within the same order of magnitude compared to the wildtype.

**H58L**—To elucidate the role of His within the context of the stringent triad His/Met/Tyr, we established the single mutant H58L. Surprisingly, the $K_m$ for thymidine was largely increased (~600-fold) and the reduction in $k_{cat}$ was about 60-fold. With the UV- and HPLC-assay the phosphorylation of neither deoxycytidine nor acyclovir was detectable.

**M128FY172F**—Combination of the inactive mutant M128F with Y172F creating the "double-F sandwich" enclosing thymidine did not show any activity, nor was any phosphorylation detected with the HPLC assay. To assure that no hydrophobic collapse was initiated by the mutations, CD spectra were recorded. However, comparison of mutant CD spectra with the wild type spectrum revealed no discrepancy. This particular Phe/Phe-combination is encountered in TKs (Fig. 2) but
Table II. Kinetics of thymidine and AZT phosphorylation of mutant E225L and wildtype HSV 1 TK. The kinetic constants (± standard errors) derived in this study are presented in the second and third column. The values have been determined by non-linear fit of the raw data to the Michaelis Menten equation using the Microcal Origin software. The DEAE paper method (Gerber & Folkers, 1996) was applied for the kinetic measurements. For AZT kinetics an additional washing step of the paper discs with ethanol 100% was necessary to remove unphosphorylated educt.

<table>
<thead>
<tr>
<th></th>
<th>$K_m$, μM</th>
<th>$k_{cat}$, s$^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Thymidine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSV 1 TK</td>
<td>0.2 ± 0.05</td>
<td>0.35 ± 0.014</td>
</tr>
<tr>
<td>E225L HSV 1 TK</td>
<td>12.3 ± 1.16</td>
<td>0.016 ± 3E-4</td>
</tr>
<tr>
<td>AZT</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HSV 1 TK</td>
<td>5.2 ± 1.70</td>
<td>0.056 ± 0.013</td>
</tr>
<tr>
<td>E225L HSV 1 TK</td>
<td>17.0 ± 1.23</td>
<td>0.032 ± 4E-3</td>
</tr>
</tbody>
</table>

with the difference that no His is found at position 58 (HSV 1 TK numbering). 

**H58LM128FY172F**—Following the indications of the alignment studies, the His at position 58 in the completely inactive mutant M128FY172F was removed by subcloning to form the triple mutant H58LM128FY172F. Indeed, the triple mutant turned out to regain activity and ability to phosphorylate dT with about 600-fold increased $K_m$, yet high enough to be determined. However, the phosphorylation rate was only reduced tenfold with regard to the wild type enzyme and significantly increased compared to single mutant H58L. Attempts to phosphorylate the guanine nucleoside analog ACV that is a prototype of many other anti-herpes drugs or deoxycytidine failed even with high amounts of enzyme using the HPLC assay.

**E225L**—In order to explore the role of electrostatic influence to enhance or alter substrate binding or the catalytic rate, we replaced the Glu 225 by the neutral Leu. Glu 225 forms a hydrogen-bond with the 3'-OH of the deoxyribose moiety of dT which is lost by the mutation. Besides studying the consequence on affinity and velocity of thymidine kinetics, we analyzed the kinetics with AZT of wildtype and mutant enzyme. The bulky, electron-rich azidogroup in 3'-position of AZT offers the possibility to reconnoiter the electrostatic proportions between active site and the LID region. The resulting effects are summarized in Table II. The decrease in $k_{cat}$ and increase in $K_m$ of the mutant towards dT compared to the wildtype is more than one order of magnitude. However, E225L reveals the same $K_m$ for both, AZT and dT whereas the catalytic rate of AZT conversion of E225L is even increased compared
to dT phosphorylation. The wildtype shows an increased $K_m$ and decreased $k_{cat}$ for AZT phosphorylation while in contrast, the mutant is even capable of increasing the $k_{cat}$ for AZT which is the opposite of the wild type's behavior. $K_m$ for AZT remains within the same order of magnitude in both enzymes, indicating that the binding mode for the substrate analog seems not to be altered by the mutation.

2.4.3. Titration Calorimetry

We performed titration experiments with the three mutants comprising the phenylalanine in position 128 to gain detailed insights into the reasons for the decreased binding affinities. The results are reported in Table III. The single replacement of Met 128 by Phe (M128F) resulted in significantly increased entropy of the system whereas the enthalpy contribution was diminished. By introduction of the second Phe the entropy becomes even more favorable, yet is responsible for establishment of binding. Since both mutants show a fairly decreased binding enthalpy, it can be suggested that the development of hydrogen-bonds for thymidine is hampered and therefore both mutants remain very weak binders. However, by complete transposition of the triad X58/F128/F172, the entropy contribution adapts to the wild type value again which enables thymidine to reappoint the correct and therefore productive hydrogen-bonding. However, the binding affinity of the triple mutant remains reduced by two orders of magnitude. These measurements completely agree with the findings revealed by the kinetical characterization. Namely,

<table>
<thead>
<tr>
<th>HSV 1 TK</th>
<th>$\Delta H$, kcal/mol</th>
<th>$K_a$, $M^{-1}$</th>
<th>$\Delta G$, kcal/mol</th>
<th>$T\Delta S$, kcal/mol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>-26.35 ± 0.47</td>
<td>2.29E+7 ± 1.46E+7</td>
<td>-9.95 ± 0.40</td>
<td>-16.4 ± 0.85</td>
</tr>
<tr>
<td>M128F</td>
<td>-8.14 ± 0.20</td>
<td>1.04E+5 ± 1.00E+4</td>
<td>-6.84 ± 0.08</td>
<td>-1.30 ± 0.20</td>
</tr>
<tr>
<td>M128FY172F</td>
<td>-3.71 ± 0.08</td>
<td>3.42E+4 ± 1.85E+4</td>
<td>-6.14 ± 0.34</td>
<td>2.43 ± 0.42</td>
</tr>
<tr>
<td>H58LM128FY172F</td>
<td>-19.82 ± 0.13</td>
<td>2.61E+5 ± 6.36E+3</td>
<td>-7.38 ± 0.01</td>
<td>-12.44 ± 0.14</td>
</tr>
</tbody>
</table>
that mutants M128F and M128FY172F are barely able to bind thymidine and thus not succeed in phosphorylating dT, whereas the triple mutant most remarkably regains binding and phosphorylation ability. For the “inactive” mutants this method provides even more accurate information as they were not distinguishable from each other with the kinetic measurements.

2.5. Discussion

To study binding properties and mechanisms, the possibility of purifying mutants with impaired binding affinities was an important issue. Our one step purification protocol addressed this issue properly and was a very convenient method allowing the purification of TK mutants not amenable to the conventional thymidine affinity resins. The subsequent thrombin cleavage led to a > 90% pure thrombin-cleaved protein.

Sequence alignments have detected similarities between herpesvirus TKs and hdCK (Harrison et al., 1991) and to a lesser extent cellular TmpK (Robertson & Whalley, 1988). Such similarity suggests a common ancestry of HSV TK and cellular dCK or cellular dTmpK, although the mechanisms by which herpesviruses have acquired these genes can only be speculated upon. The superimposition of over 100 amino acids of HSV TK with the recently published crystal structures of E. coli and yeast dTmpK (Lavie et al., 1998b; Lavie et al., 1997) reveals striking structural similarity, the latter divided into the same structural type dTmpK as cellular dTmpKs (Lavie et al., 1998a). The root mean square (rms) deviation values are markedly low with a value of 2.5 Å when the glycine-rich loop and the five stranded β-sheets of yeast dTmpK and HSV 1 TK were superimposed. Moreover, by this superimposition the overall 3D-structure of HSV 1 TK with the exception of the region between amino acid 250 and 320 was also nicely fitting to the shorter dTmpK. A closer look at yeast and E. coli dTmpK revealed not unexpectedly a similar binding mode for the nucleobase compared to HSV 1 TK, namely Phe 69 and the Cβ of serine 98 sandwiching thymine (yeast numbering). Such meaningful structural similarity suggests even more convincing than sequence alignments the evolutionary relationship among herpesviral TKs and dTmpK. Unfortunately, to date no such material is available for dCK.

Despite the progress made in understanding detailed aspects of HSV 1 TK ligand binding, several key questions regarding the broad substrate diversity remain
unresolved. In our attempt to rationalize this property we have designed a series of HSV 1 TK mutants at the positions 58, 128 and 172 and characterized their biochemical and physicochemical properties. For the comparison of the binding behavior of the different mutants, the Michaelis constant $K_m$ has been used as binding constant because it has been previously shown that $K_m$ of dT ($K_m = 0.2 \, \mu M$) and ACV ($K_m = 200 \, \mu M$) correspond to the dissociation constant of dT ($K_2 = 0.139 \, \mu M$) and ACV ($K_5 = 162 \, \mu M$) (Kussmann-Gerber et al., 1999). Furthermore, for ACV the $K_m$ value of 0.2 mM (Table I) corresponds with the $K_i$ values ranging from 100 to 200 $\mu M$ reported in literature (Fyfe et al., 1983; Larder et al., 1983a; Larder et al., 1983b). This represents a peculiarity of HSV 1 TK by which dT is binding prior to ATP and the rate constant of disintegration of the intermediate complex (ES) is negligible as compared to the corresponding dissociation rate constant (Kussmann-Gerber et al., 1999).

In our study it is noteworthy that no mutation affecting an amino acid hydrogen-bonding thymidine was introduced. The combination of His 58, Tyr 172 and Met 128 in TK is hitherto only found in four viral strains, namely HSV 1, HSV 2, MHV and BHV, causing an extensive broad substrate acceptance towards both, the sugar and base moiety of the nucleoside. Instead, in TKs that are more limited in phosphorylation with respect to the base and/or sugar moiety (Abele et al., 1991; Gustafson et al., 1998; Roberts et al., 1993), the combination X58/F128/F172 is found where X can be Tyr, Phe, Ile, Met or Pro which are all hydrophobic. Our results emphasize the extreme sensitivity of substrate affinity on mutational changes at the positions 58, 128 and 172 of HSV 1 TK.

At position 128, we have shown the role of sulfur to be purely hydrophobic. The M128I mutant constitutes a bio-isosterical modification of the wild type exhibiting almost identical affinity for thymidine and substrate analog drug ACV. This is in full agreement with previous theoretical investigations (Alber et al., 1998). We conclude that a modulation of residue size in the hydrophobic pocket at position 128 has a direct impact on binding affinity. As expected, the mutant M128A loses biological activity, which is an indication that the small alanine sidechain is not sufficient to stabilize the thymine within the active site. The analysis of the structure suggests that the retained activity minimum is probably due to a partial compensation of the missing methylgroup (C=C) of Met 128 by C=C of Ile 97 through dynamic
rearrangements. Instead, the loss in biological activity of the M128F mutant is due to unproductive orientation of weakly bound thymidine as our ITC measurements clearly indicate. The nature of the sandwich-like complex is modified by the introduction of the bulky Phe 128 into the available space provided by surrounding hydrophobic amino acids (namely W 88, Thr 96 (Cγ), Ile 97 and Ile 100). This finding is further supported by the fact that the combination of 128F/172Y (HSV 1 TK numbering) sandwiching thymidine has not been found so far in any TK sequence. Additionally, the mutant L58/F128/Y172, another in nature nonexistent combination, shows only barely detectable dT phosphorylation. The alteration of the flexibility of the system represented by the entropy difference between mutant and wild type may explain experimental findings showing that the mutant M128F is not able to phosphorylate thymidine. Further evidence is added by ITC measurements of this mutant in absence of ATP. The resulting binding enthalpies are the same as when titrated in the presence of ATP. In contrary, a substantial difference in binding affinities can be measured with the wild type enzyme under the same experimental conditions suggesting an extended induced fit. These results imply an altered or impaired binding site for ATP in the mutant.

Similarly the double mutant M128FY172F is not able to enforce catalytic turnover. The adaptation of a second Phe in the mutant M128FY172F even changed the sign of the entropy term, giving a hint of a favorably preformed binding pocket and of a reduction of the induced fit movement. However, the establishment of hydrogen-bonds for the substrate seems to be severely impaired by this arrangement with a more than ten-fold reduction in binding enthalpy. It can be learned from the sequence alignments that the Phe-Phe combination indeed exists, but not with a mutual His in the P loop region. In the above discussed mutants, the possibility for a reorganization upon thymidine binding is greatly impaired indicated by the already advantageous entropy of binding. Rather, removal of His is prerequisite to productive binding. The mode of binding seems to be enthalpy forced for wildtype and triple mutant TK and entropy driven with M128FY172F.

Most striking, an additional mutation H58L, thus introducing a hydrophobic leucine at the histidine position recovers catalytic activity in the M128FY172F mutant. Interestingly, the exchange of the H58/M128/Y172 to the L58/F128/F172 results in a mutant enzyme whose specificity is mainly based on a general loss in affinity. We observe catalytic activity only towards the natural substrate, thus developing
resistance towards purine nucleoside analogs. The rational for the severely reduced affinity must lie in a different orientation of the base as all the amino acids that form hydrogen-bonds with dT or guanine (20) are still present. This is corroborated by our ITC measurements, revealing a $\Delta S$ similar to wild type which indicates a restored flexibility of the enzyme. From the results, we believe the H58/M128/Y172 triad to be responsible for a better hydrophobic fit to natural substrates allowing the occurrence of the successive movement for completing the catalytic cycle. As we have shown, the residue 58 plays a central role in the formation of a hydrophobic pocket in a catalytically active mutant enzyme. However, the understanding of the functional role of His 58 in the binding process is not fully settled yet.

Nonetheless, on the basis of our study, we may postulate some functional roles. As can be seen in Fig. 1, the exchange of histidine 58 will presumably affect the orientation of the ribose part of the substrate. A structural variation at position 58 could indeed allow a reorientation of the ribose part, being responsible for the recovered catalytic activity in the triple mutant. His 58, positioned between the two Glu 225 and 83, respectively, may also serve as transmitter of electron density and therefore play a central role in catalysis (together with E225 and 83) and electrostatics (hydrogen-bond with Tyr 172) within the active site. However, the structural data and ab initio calculation that are so far available do not yet provide sufficient information on the role of this residue.

The mutant E225L, belonging to the LID domain gives further support for the involvement of movement and electrostatic interaction. In the structure Glu 225 is together with Glu 83 the only negatively charged amino acid within a cluster of positively charged residues. Glu 225 forms a hydrogen-bond with the 3'-OH of the deoxyribose moiety which is broken by the mutation. However, the hydrogen-bond of Tyr 101 to the 3'-OH is preserved (Fig. 1). The loss in affinity for dT might be explained by the missing hydrogen-bond, but not the decrease in velocity. The catalytic rate is severely reduced although all amino acids, besides E225, involved in nucleobase binding, and all the amino acids (E83, R163) apparently involved in catalysis, are disposable. Strikingly, $k_{\text{cat}}$ of AZT phosphorylation is faster than the $k_{\text{cat}}$ for dT whereas for the wildtype the opposite situation has been noticed. In TKs the $\gamma$-phosphate of ATP and 5'-OH of deoxyribose need to be activated for the catalytic reaction. This is achieved by clusters of positive charges from the LID domain (Arg and Lys) and the Mg$^{2+}$, turning the phosphorus atom amenable for an nucleophilic
attack of the polarized 5'-O which is positioned between E225 and E83 (Wild et al., 1997) and needs to be negatively charged. During AZT phosphorylation, E225 of wild type HSV 1 TK is displaced by the bulky 3'-azidogroup of AZT (Christians et al., 1999), which leads to a reduction of polarization of the 5'-O with a consequent decrease in velocity. In the E225L mutant the decline in velocity lies within the same order of magnitude. This finding suggests that either displacement or removal of the negative charge (Glu 225) results in an equal effect. By searching for similar features in dTmpK, we found Asp 14 in yeast and Glu 12 in E. coli dTmpK that might take over the function of both, Glu 225 and Glu 83 in polarizing 3'-OH of the deoxyribose and of one oxygen of the α-phosphate of dTMP.

It is noteworthy that our mutagenesis study on the triad involved residues without direct hydrogen-bond contact with the substrate underscoring the capability of hydrophobic contacts and electrostatics. The residues maintaining direct hydrogen-bonds rather guide resistance patterns (Kussmann-Gerber et al., 1998). Our results emphasize the extreme sensitivity of substrate affinity and diversity on mutational changes at the HSV 1 TK positions 58, 128 and 172. This finding is in complete agreement with our sequence alignment study, indicating the residue triad His 58, Tyr 172, and Met 128 to be a common motif in thymidine kinases with broad substrate diversity whereas variations at these positions, which correspond to the X58/F128/F172 HSV 1 TK mutations, are a common feature for enzymes with restricted substrate acceptance. Our findings indicate that the existence of a structurally flexible sandwich complex and the maintenance of balanced electrostatics are crucial for substrate diversity in HSV 1 TK and a plausible evolutionary pattern. Therefore, we add a new piece of information for the design of new antiviral drugs and modified TKs for gene therapy of cancer and AIDS.

Acknowledgments
We thank Dr. I. Jelesarov for constructive discussion and technical assistance with the ITC measurements, U. Kessler for aid with the HPLC, F. Seegy for assistance with the figure and P. Pospisil for support with the structural comparison of dTmpK.
2.6. References


bisubstrate inhibitor P1-(5'-adenosyl) P5-(5'-thymidyl) pentaphosphate (TP5A) at 2.0 Å resolution: implications for catalysis and AZT activation. *Biochemistry* **37**(11), 3677-86.


The Periplasmic Domain of the Histidine Autokinase CitA Functions as a Highly Specific Citrate Receptor

Sibylle Kaspar\textsuperscript{1,3}, Remo Perozzo\textsuperscript{2}, Stefan Reinelt\textsuperscript{2}, Margareta Meyer\textsuperscript{1}, Karin Pfister\textsuperscript{1}, Leonardo Scapozza\textsuperscript{2}, and Michael Bott\textsuperscript{1,3}

\textsuperscript{1}\textsuperscript{}Mikrobiologisches Institut, Eidgenössische Technische Hochschule, Schmelzbergstrasse 7, CH-8092 Zürich, Switzerland; \textsuperscript{2}\textsuperscript{}Departement Pharmazie, Eidgenössische Technische Hochschule, Winterthurerstr. 190, CH-8057 Zürich, Switzerland; \textsuperscript{3}\textsuperscript{}Present address: Institut für Biotechnologie I, Forschungszentrum Jülich, D-52425 Jülich, Germany

Dedicated to Professor Rudolf K. Thauer on the occasion of his 60th birthday.

Molecular Microbiology 33(4), 858-872 (1999)
3.1. Summary

The two-component regulatory system CitA/CitB is essential for induction of the citrate fermentation genes in *Klebsiella pneumoniae*. CitA represents a membrane-bound sensor kinase consisting of a periplasmic domain flanked by two transmembrane helices, a linker domain, and the conserved kinase or transmitter domain. A fusion protein (MalE-CitAC) composed of the maltose binding protein and the CitA kinase domain (amino acids 327-547) showed constitutive autokinase activity and transferred the γ-phosphoryl group of ATP to its cognate response regulator CitB. The autokinase activity of CitA was abolished by an H350L exchange, and phosphorylation of CitB was inhibited by a D56N exchange, indicating that H350 and D56 represent the phosphorylation sites of CitA and CitB, respectively. In the presence of ATP, CitB-D56N formed a stable complex with MalE-CitAC. To analyze the sensory properties of CitA, the periplasmic domain (amino acids 45 to 176) was overproduced as a soluble, cytoplasmic protein with a C-terminally attached histidine tag (CitAPHis). Purified CitAPHis bound citrate, but none of the other tri- and dicarboxylates tested, with high affinity (K_D ~ 5 μM at pH 7) in a 1:1 stoichiometry. As shown by isothermal titration calorimetry, the binding reaction was driven by the enthalpy change (∆H = -76.3 kJ/mol), whereas the entropy change was opposed (-T ∆S = +46.3 kJ/mol). The pH-dependency of the binding reaction indicated that the dianionic form H-citrate$^{2-}$ is the citrate species recognized by CitAPHis. In the presence of Mg$^{2+}$ ions the dissociation constant significantly increased, suggesting that the Mg-citrate complex is not bound by CitAPHis. This work defines the periplasmic domain of CitA as a highly specific citrate receptor and elucidates the binding characteristics of CitAPHis.

3.2. Introduction

Regulation of bacterial gene expression in response to environmental changes is often controlled by two-component regulatory systems, consisting of a sensor kinase and a response regulator. The sensor kinase responds to a certain stimulus by autophosphorylation of a conserved histidine residue, and the phosphoryl group is subsequently transferred to a conserved aspartate residue of the response regulator, which then mediates changes in gene expression or cell behavior (for reviews see
In the majority of cases, sensor kinases are transmembrane proteins with an extracellular N-terminal sensor domain (typically 100 to 200 amino acids in length) flanked by two transmembrane helices and an intracellular C-terminal autokinase domain (~220 amino acids) that is connected to the second transmembrane helix via a linker region of variable length. This architecture enables the proteins to sense external stimuli and transduce information to the cytoplasm. Results obtained with several sensor kinases have shown that these proteins function as dimers and that the autophosphorylation occurs in trans (for review see Stock et al., 1995). Although many two-component systems have been analyzed in the past years, only in a few cases the primary signals recognized by a given sensor kinase have been identified biochemically. In part, this lack of information is caused by the difficulties in purifying membrane-bound sensor kinases. A strategy to overcome this problem is to express the extracytoplasmic domain as a separate, soluble protein, which then allows to study its properties in vitro. Successful examples of this approach represent the Mg^{2+}/Ca^{2+}-responsive sensor kinase PhoQ (Vescovi et al., 1997), or the aspartate chemoreceptor (Milligan & D.E. Koshland, 1993), whose modular structure resembles that of the sensor kinases described above.

In this work, we have analyzed the periplasmic and the kinase domain of the sensor CitA from Klebsiella pneumoniae. This protein, in cooperation with its cognate response regulator CitB, is essential for the expression of the citrate fermentation genes, as shown by the fact that citA and citB mutants were no longer able to grow under anoxic conditions with citrate as sole carbon and energy source (Bott, 1997; Bott et al., 1995). The citrate fermentation genes form a cluster on the chromosome, which is composed of two divergently transcribed units, the citC operon and the citS operon. The citC operon encodes citrate lyase ligase (citC), the \( \gamma \) (citD), \( \beta \) (citE), and \( \alpha \)-subunit (citF) of citrate lyase, and a protein (citG) presumably involved in the biosynthesis of the 2'-5''-phosphoribosyl)-3'-dephospho-CoA prosthetic group of citrate lyase (Bott & Dimroth, 1994). The citS operon encodes the Na\(^+\)-dependent citrate carrier CitS (citS), and the \( \gamma \) (oadG), \( \alpha \) (oadA), and \( \beta \)-subunit (oadB) of the Na\(^+\)-pump oxaloacetate decarboxylase. The citAB genes are the promoter-distal genes of the citS operon and are positively autoregulated (Bott et al., 1995). The sensor kinase CitA shows the typical modular composition described above, i.e. it
Fig. 1. Domain organization of the sensor kinase CitA from *K. pneumoniae*. The two transmembrane helices (TMH1 and TMH2), two putative Q-linkers (Q1 and Q2; Wootton and Drummond, 1989), and the five signature amino acid segments of the kinase domain (named H-, N-, G1-, F-, and G2-box according to the most characteristic amino acids; Parkinson and Kofoid, 1992) are shown. The position of the different features within the primary structure is indicated by the numbers. In the lower part, the proteins CitAPHis and MalE-CitAC used in this study are shown schematically.

consists of a periplasmic domain flanked by two transmembrane helices, a linker region, and the kinase domain composed of the phosphorylation subdomain with the conserved H-box and the ATP-binding subdomain with the N-, G1-, F-, and G2-boxes (Fig. 1). The periplasmic location of the N-terminal domain predicted from sequence analysis was confirmed accidentally in the course of a topological analysis of the \( \alpha \)-subunit of oxaloacetate decarboxylase (Di Berardino *et al.*, unpublished) with the TnphoA transposon (Manoil & Beckwith, 1985). In this study, a target plasmid was used that contained besides the oadB gene also the 5'-terminal part of citA up to codon 121. Analysis of transposon insertions resulting in strong alkaline phosphatase activity revealed that the majority of them were located in the citA part of the plasmid. In one case, the exact fusion site was determined by DNA sequencing and found to be after Thr-47, located immediately behind the first transmembrane helix of CitA. The transcriptional activator CitB is composed of the conserved N-terminal receiver domain and a C-terminal domain with a helix-turn-helix motif. Purified, unphosphorylated CitB binds specifically to the *citC-citS*
intergenic region, but with low affinity. *In vitro* phosphorylation of CitB elicits structural changes, resulting in an increase of its binding affinity by a factor of \( \sim 50 \) (Meyer et al., 1997).

Expression of the citrate fermentation genes has to be carefully regulated, since the synthesis of both citrate lyase and oxaloacetate decarboxylase under inappropriate conditions could severely affect the function of the citric acid cycle, either by triggering a futile cycle of citrate synthesis and cleavage or by deprivation of oxaloacetate. In addition, oxaloacetate decarboxylase could perturb the \( \text{Na}^+ \) balance across the cytoplasmic membrane. Previous studies showed that expression of the *citC* operon and the *citS* operon is dependent on citrate, \( \text{Na}^+ \) ions, and micro- or anaerobic conditions (Bott et al., 1995). Therefore, it was proposed that CitA may function as a citrate sensor. In this work, we tested this proposal by analyzing the periplasmic domain as a separate protein. Moreover, the properties of the CitA autokinase domain and its interaction with CitB were analyzed.

### 3.3. Results

#### 3.3.1. Isolation and Characterization of the CitA Kinase Domain

As part of our efforts to understand the mode of action of the two-component sensor protein CitA from *Klebsiella pneumoniae*, the CitA kinase domain (amino acids 327 - 547) was fused to the C-terminus of the maltose binding protein lacking its signal peptide. The resulting fusion protein MalE-CitAC (66.8 kDa) was overproduced in *Escherichia coli* and purified by affinity chromatography using amylose resin. As shown in Fig. 2, the fraction obtained after elution with maltose-containing buffer contained predominantly a protein of the expected size of 66 kDa, but also several degradation products in the range between 65 and 40 kDa and another one of about 27 kDa. Whereas the latter corresponds to the CitA part of MalE-CitAC, the former presumably consist of the entire MalE (~40 kDa) plus different proportions of the CitA kinase domain. Further purification of MalE-CitAC by gel filtration on a Superdex-200 column (Pharmacia) revealed that generally about 80-90% of the protein was present in a multimeric state of >600 kDa, whereas 10-20% was present in a dimeric state with a size of 130 kDa (data not shown). Upon storage of the dimeric form at 4°C, larger aggregates formed again as shown by native PAGE (Fig. 2, lane 7).
Fig. 2. Overproduction, purification and functional analysis of MalE-CitAC. Lanes 1 and 4, protein standards. Lane 2 and 3, whole cell lysates of E. coli DH5*/pMalE-CitAC before and 3 h after IPTG induction, respectively. Lane 5, MalE-CitAC (15 µg) obtained after affinity chromatography of cell extract on amylose resin. Lane 1 to 5 show Coomassie-stained SDS-polyacrylamide gels. Lane 6, autophosphorylation of MalE-CitAC (0.5 µg) after 10 min incubation at room temperature with 20 µM $\gamma$-[32P]-ATP (30000 dpm/pmol) as revealed by SDS-PAGE and subsequent Phospholmager analysis. Lane 7, native PAGE (Laemmli system) showing Coomassie-stained multimeric forms of MalE-CitAC (~5 µg) obtained after storage of the dimeric form at 4°C for one week.

Incubation of the protein obtained after amylose chromatography with $\gamma$-[32P]-ATP led to autophosphorylation as shown by SDS-PAGE and subsequent Phospholmager analysis (Fig. 2, lane 6). Several degradation products including that of 27 kDa also became phosphorylated. Analysis of the MalE-CitAC complexes obtained after gel filtration revealed that the large multimer (>600 kDa) usually had a significantly weaker autokinase activity compared to the dimeric form (data not shown).

According to sequence alignments with other histidine protein kinases, His-350 was proposed to be the autophosphorylated residue of CitA (Bott et al., 1995). By site-directed mutagenesis, the corresponding residue in MalE-CitAC was replaced by a leucine residue. The resulting protein MalE-CitAC-H350L was overproduced and purified as described for MalE-CitAC. Gel filtration of the protein obtained after amylose chromatography showed the same result as described for MalE-CitAC, i.e. the major part of the protein formed a multimer of >600 kDa and a minor part was present in the dimeric form. MalE-CitAC-H350L was not autophosphorylated in the
presence of $\gamma^{[32P]}$-ATP, providing strong support for the assumption that His-350 represents the phosphorylation site of CitA.

### 3.3.2. CitB Phosphorylation by MalE-CitAC

The ability of the CitA kinase domain to phosphorylate CitB was tested in vitro with MalE-CitAC and CitB$_{His}$, a CitB derivative containing eight additional amino acids at its C-terminus including a (His)$_5$ tag (Meyer et al., 1997). As shown in Fig. 3, MalE-CitAC transferred the $\gamma$-phosphate group of ATP efficiently to CitB$_{His}$. In the absence of MalE-CitAC, no phosphorylation of CitB$_{His}$ with $\gamma^{[32P]}$-ATP was observed. The velocity of CitB$_{His}$ phosphorylation increased with increasing MalE-CitAC concentrations (data not shown). Besides CitB$_{His}$ also CitB$_{NHis}$ which lacks the C-terminal DNA-binding part of CitB (amino acids 139-242) was phosphorylated by MalE-CitAC. This result shows that the kinase domain of CitA (amino acids 327-547) and the receiver domain of CitB (amino acids 1-138) carry all the structural determinants required for the specific protein-protein interaction and phosphoryl transfer. Exchange of the Asp-56 residue in CitB$_{His}$ against asparagine led to CitB$_{His}^{-}$D56N, which could no longer be phosphorylated in vitro by MalE-CitAC (Fig. 3). This result strongly supports the prediction based on sequence alignments (Bott et al., 1995) that Asp-56 represents the phosphorylation site of CitB. Interestingly, the presence of CitB$_{His}^{-}$D56N usually stimulated the extent (1.5- to 5-fold; Fig. 3) and the rate (about 3-fold; data not shown) of MalE-CitAC autophosphorylation. This effect could only be explained by an interaction of the two proteins. Analysis by native PAGE revealed that CitB$_{His}^{-}$D56N was able to form a stable complex with MalE-CitAC, especially in the presence of ATP (Fig. 4, lanes 5 and 6). CitB$_{His}$ also formed a complex with MalE-CitAC, but only to a minor extent and apparently not influenced by the presence of ATP (Fig. 4, lanes 3 and 4).

### 3.3.3. Overproduction and Purification of the Periplasmic Domain of CitA

The CitA/CitB two-component regulatory system is essential for the induction of genes specifically required for the anaerobic catabolism of citrate in *K. pneumoniae*. Therefore, citrate was a likely target to be recognized by the sensor kinase CitA, or more specifically, by the periplasmic part of this protein (Bott et al., 1995). To verify this hypothesis, we tested whether this domain is capable of binding citrate. For this
Fig. 3. Phosphorylation of MalE-CitAC in the absence and in the presence of CitB\(_{\text{His}}\) or CitB\(_{\text{His}}\)-D56N. The three assay mixtures (total volume 50 \(\mu\)l) contained 0.5 \(\mu\)M MalE-CitAC (dimer), 0.5 mM \(\gamma^{32}\)P-ATP (~1000 dpm/pmol), and either no other protein (lanes 1-3), or 9.4 \(\mu\)M CitB\(_{\text{His}}\) (lanes 4-6), or 9.4 \(\mu\)M CitB\(_{\text{His}}\)-D56N (lanes 7-9). After start of the reaction by ATP addition, the mixtures were incubated at room temperature. 15 \(\mu\)l aliquots were withdrawn after 5 min (lanes 1, 4, and 7), 10 min (lanes 2, 5, and 8), and 15 min (lanes 3, 6, and 9), mixed immediately with an equal volume of 2-fold concentrated SDS loading buffer, and stored on ice. Subsequently, 10 \(\mu\)l of each sample was subjected to SDS-PAGE and the dried gels were analyzed with a Phospholmager.

Fig. 4. Complex formation between MalE-CitAC and CitB\(_{\text{His}}\) or CitB\(_{\text{His}}\)-D56N. A, Coomassie-stained native polyacrylamide gel (5\%). B, Western blot with CitB\(_{\text{His}}\) antiserum of an identical gel. Lane 1, MalE-CitAC; lane 2, MalE-CitAC preincubated with ATP; lane 3, MalE-CitAC and CitB\(_{\text{His}}\); lane 4, MalE-CitAC and CitB\(_{\text{His}}\) preincubated with ATP; lane 5, MalE-CitAC and CitB\(_{\text{His}}\)-D56N; lane 6, MalE-CitAC and CitB\(_{\text{His}}\)-D56N preincubated with ATP; lane 7, CitB\(_{\text{His}}\); lane 8, CitB\(_{\text{His}}\)-D56N. The concentrations in the samples were 2.7 \(\mu\)M for MalE-CitAC (dimeric form), 16 \(\mu\)M for CitB\(_{\text{His}}\), 14.5 \(\mu\)M for CitB\(_{\text{His}}\)-D56N, and 7.5 mM for ATP. After 20 min at room temperature, aliquots were mixed with sample buffer and subjected to native PAGE at 4\(^{\circ}\)C.
purpose, the expression plasmid pET24-CitAP was constructed which allows the overproduction of the periplasmic part of CitA (amino acids 45-176), supplemented with an N-terminal methionine and a C-terminal (His)$_6$ tag, as a soluble, cytoplasmic protein (CitAP$_{His}$). After induction of expression using *E. coli* BL21(DE3) as host, a polypeptide of ~15 kDa was formed which constituted the predominant protein of the cells (Fig. 5). The apparent size of this protein corresponded to that predicted for CitAP$_{His}$ (15.3 kDa). After cell disruption and ultracentrifugation, the overproduced protein was found in the supernatant, showing that CitAP$_{His}$ was present in a soluble state. The six consecutive histidine residues at the C-terminus provided the means for purification of CitAP$_{His}$ by Ni$^{2+}$ chelate affinity chromatography, resulting in a highly pure (>98%) protein (Fig. 5). The yield was usually in the range of 10-20 mg per g cells (wet weight). Gel filtration of CitAP$_{His}$ indicated that ~95% of the protein was present in a monomeric state with an apparent mass of 13 kDa and ~5% in a dimeric form with an apparent mass of 26 kDa (data not shown).

![Fig. 5. Coomassie-stained SDS-polyacrylamide gel showing overproduction and purification of CitAP$_{His}$. Lane 1, protein standard; lane 2 and 3, whole cell lysates of *E. coli* BL21(DE3)/pET24-CitA before and 3 h after IPTG addition, respectively; lane 5, CitAP$_{His}$ (~3 µg) obtained after Ni$^{2+}$ chelate affinity chromatography.](image-url)
3.3.4. Binding Studies with 1,5-[\textsuperscript{\textit{14}}C]-Citrate

Binding assays with 1,5-[\textsuperscript{\textit{14}}C]-citrate revealed that purified CitAP\textsubscript{His} was able to bind citrate (Fig. 6). A non-linear fit of log [free citrate] versus [bound citrate] to the Hill equation (see Experimental Procedures) yielded a Hill coefficient $n_H$ of 1.0, indicating that binding is non-cooperative (Fig. 6A). The calculated $K_D$ value was 34 $\mu$M and the calculated maximal concentration of binding sites ($B_{\text{max}}$) was 193 $\mu$M. This latter value corresponded to 77% of the monomeric CitAP\textsubscript{His} concentration in the assay as determined by the bicinchoninic acid protein assay and thus clearly indicated a single citrate binding site per CitAP\textsubscript{His} monomer. A Scatchard plot (Fig. 6B) of the binding data gave a straight line, indicative of non-cooperative binding, a $K_D$ value of 39 $\mu$M and a maximal concentration of binding sites of 200 $\mu$M. These values are in fairly good agreement with those obtained by the fit to the Hill equation. Gel filtration of CitAP\textsubscript{His} in the presence of 10 mM citrate led to an elution profile identical to that obtained in the absence of citrate, showing that citrate binding did not cause a monomer-dimer transition.

The data described above were obtained with TKMD as assay buffer, which contains 5 mM MgCl\textsubscript{2}. Since citrate is known to form complexes with Mg\textsuperscript{2+} ions and other divalent cations, binding assays were performed in several other buffers that did not contain Mg\textsuperscript{2+} ions. Evaluation of the corresponding data yielded significantly lower $K_D$ values of about 5 $\mu$M at pH 7, indicating that Mg\textsuperscript{2+} ions inhibited citrate binding to CitAP\textsubscript{His}.

The ligand specificity of CitAP\textsubscript{His} was studied by testing the ability of other tri- and dicarboxylic acids, i.e. isocitrate, tricarballylate, succinate, fumarate, malate, or tartrate, to displace CitAP\textsubscript{His}-bound [\textsuperscript{\textit{14}}C]-citrate. A slight displacement was only observed with tricarballylate and fumarate, and the $K_D$ values calculated from these data according to Clarke and Koshland (1979) were more than three orders of magnitude higher than the $K_D$ value for citrate. Consequently, the CitAP\textsubscript{His} protein functions as a highly specific citrate receptor.
Fig. 6. Citrate binding to the periplasmic domain of the sensor kinase CitA. Binding was measured with \([^{14}C]\)-citrate as described under "Experimental procedures". The CitAPH concentration in the assay as determined with the bicinchoninic acid assay was 250 µM. A. Plot of [bound citrate] versus log [free citrate]. The curve is a nonlinear fit to the Hill equation (see Experimental Procedures). B. Scatchard analysis of the same data showing the resulting linear function, which indicates that binding is non-cooperative.
3.3.5. Binding Studies Using Isothermal Titration Calorimetry

As an alternative and independent method to measure the binding properties of CitAP\textsubscript{His}, we used isothermal titration calorimetry (ITC; (Ladbury & Chowdhry, 1996; Wiseman \textit{et al.}, 1989)). ITC measures binding interactions by detecting the heat absorbed or released during a binding event. Deconvolution of the binding isotherm yields the number of binding sites (n), the enthalpy of binding (\(\Delta H\)), and the equilibrium association constant (\(K_A\)). From these parameters, the Gibbs free energy (\(\Delta G\)) and the entropy of binding (\(\Delta S\)) can be calculated. A representative microcalorimetric experiment is depicted in Figure 7. From the raw data shown in panel A, trace I, which were obtained by 30 injections of citrate (1.43 mM) into a solution of CitAP\textsubscript{His} (61.57 \(\mu\text{M}\)) at pH 7.0, it is evident that citrate binding to CitAP\textsubscript{His} is an exothermic process. In panel B, the integrated heat of each injection (filled squares) is plotted against the molar ratio of citrate to CitAP\textsubscript{His}. The solid line represents a nonlinear least squares fit according to a single-site binding model, which yields a \(K_D\) value (\(K_D = 1/K_A\)) of 5.5 \(\mu\text{M}\), an observed enthalpy change (\(\Delta H_{\text{obs}}\)) of -76.4 kJ/mol, and a stoichiometry (n) of 0.9. In contrast to citrate, neither isocitrate nor tricarballylate exhibited demonstrable binding to CitAP\textsubscript{His} under the same microcalorimetric conditions (Fig. 7A, traces II and III, respectively), which is in agreement with the binding assays described above. Additional displacement experiments were necessary in order to exclude complete entropically driven binding which does not release significant binding heat. The results of these experiments showed that the presence of isocitrate and tricarballylate does not alter the binding behavior of citrate towards CitAP\textsubscript{His} (data not shown). Therefore we conclude that both isocitrate and tricarballylate do not bind to the citrate binding site.

Depending on pH, citrate exists in four different species, i.e. \(\text{H}_3\text{-citrate}\), \(\text{H}_2\text{-citrate}\), \(\text{H}\text{-citrate}\), and \(\text{citrate}\) with \(pK_1 = 3.13\), \(pK_2 = 4.76\), and \(pK_3 = 6.40\) (Sillén & Martell, 1964). In order to determine, which of these species is the preferred ligand of CitAP\textsubscript{His}, the pH-dependency of the citrate association constant (\(K_A\)) was determined. Experiments were performed in duplicates from pH 4.0 to 9.0 in pH intervals of one and the data derived from these calorimetric titrations are summarized in Table 1.

The curve obtained from a plot of log \(K_A\) versus pH (not shown) had a maximum at about pH 5.7 and decreased above and below this value, indicating that citrate binding is linked to ionization events (Bradshaw & Waksman, 1998). As outlined in
Fig 7. Binding of citrate and analogues to CitAP_his measured by isothermal titration calorimetry. The titrations were performed in 50 mM sodium phosphate buffer pH 7.0 at 25°C. (A) Raw data for titration of 61.57 μM CitAP_his with 1.43 mM citrate (I), 98.49 μM CitAP_his with 1.24 mM isocitrate (II), and 112.09 μM CitAP_his with 2.25 mM tricarballylate (III) (B) Titration plot derived from the integrated raw data (corrected for the heat of dilution) and the [citrate]/[CitAP_his] ratio. The solid line represents a best fit for citrate binding according to a single-site model, which yields a K_D value of 5.5 μM, a binding enthalpy of ~76.2 kJ/mol and a stoichiometry of 0.9 per monomer (15.3 kDa). No heat release was observed upon titration of isocitrate and tricarballylate.
the Discussion, this result can be explained by the assumption that H-citrate$^{2-}$ is the citrate species recognized by CitAP$_{His}$. The $\Delta H_{obs}$ and $\Delta S$ values were negative over the pH range studied, showing that binding of citrate to CitAP$_{His}$ was driven by the enthalpy change, whereas the entropy change was opposed. The mean value of the binding stoichiometry (n) derived from iterative fittings was 0.87 ±0.06, and all data fitted well to a single-site binding model.

The inhibitory influence of Mg$^{2+}$ ions on citrate binding observed before could be confirmed by ITC. As shown by the titration experiments presented in Table 2, increasing Mg$^{2+}$ concentrations resulted in an exponential decrease of the binding affinity. At 20 mM Mg$^{2+}$, a 10-fold reduction of $K_A$ was measured.

3.4. Discussion

3.4.1. The CitA Kinase Domain and its Interaction with CitB

In order to understand the mode of action of membrane-bound sensor kinases we have characterized two domains of the CitA protein which controls the phosphorylation status of the response regulator CitB and thus the expression of the citrate fermentation genes in *Klebsiella pneumoniae*. The kinase or transmitter domain, representing the characteristic feature of this protein family, was analyzed in fusion with the maltose binding protein (MalE) from *Escherichia coli*. In contrast to constructs that contained a His-tag instead of the MalE-tag (data not shown), MalE-CitAC was produced as a soluble protein and thus could be isolated without previous denaturation. Purification and functional analysis of MalE-CitAC revealed a number of important properties: (i) Despite the use of protease inhibitors, a significant proportion of the CitA part of MalE-CitAC obtained after affinity chromatography was degraded, indicating the existence of flexible regions that are highly susceptible to proteolysis. The structural data obtained recently with EnvZ (Tanaka *et al.*, 1998) and CheA (Bilwes *et al.*, 1999) confirm the presence of at least one such region within the ATP-binding subdomain of these proteins. (ii) Several of the C-terminally truncated MalE-CitAC degradation products became phosphorylated upon incubation with γ-$[^{32}P]$-ATP. This can be explained by the formation of dimers between a full-length MalE-CitAC protein and a C-terminally truncated form. Since phosphorylation occurs in *trans*, such a heterodimer would allow the phosphorylation of the monomer lacking
the ATP-binding site. As outlined below, the phosphorylation subdomain rather than the ATP-binding subdomain is presumably responsible for dimerization. (iii) The majority of MalE-CitAC was present in a multimeric state of >600 kDa and a minor part in a dimeric state of 130 kDa. Since the maltose binding protein has not been reported to form dimers or multimers, the dimerization and subsequent multimerization of MalE-CitAC must be caused by the CitA part of the fusion protein. In the case of EnvZ it was shown that the subdomain containing the phosphorylation site formed a stable dimer, whereas the ATP-binding subdomain was present as a monomer (Park et al., 1998). The latter finding was confirmed by the structural analysis of this domain (Tanaka et al., 1998). Thus, one may propose that the phosphorylation subdomain is responsible for the dimerization of class I histidine kinases, where it is located immediately in front of the ATP-binding subdomain. The multimerization of MalE-CitAC might be an artificial effect resulting from the deletion of the N-terminal CitA part and/or the unphysiological high concentration of the purified protein. (iv) In accordance with the proposal that H350 of CitA is the phosphorylation site, a MalE-CitAC derivative with a H350L exchange was no longer capable of autophosphorylation.

The phosphorylation cascade from CitA to CitB could be reconstituted in vitro with MalE-CitAC and CitB_{His} (Fig. 3). We have shown previously that phosphorylation of CitB_{His} enhances its DNA-binding affinity by a factor of ~50 (Meyer et al., 1997). A CitB_{His} derivative with an D56N exchange could no longer be phosphorylated by MalE-CitAC, supporting that D56 is the phosphorylation site. Interestingly, the presence of CitB_{His}-D56N usually stimulated the rate and the extent of MalE-CitAC autophosphorylation (Fig. 3). This effect can be explained by the assumption that CitB_{His}-D56N forms a complex with phosphorylated MalE-CitAC and thereby increases the half-life of phosphorylated H350. Indeed, such a complex could be demonstrated by native PAGE and its formation was significantly stimulated by ATP (Fig. 4). This result fits into the following simplified picture of the phosphotransfer: First, signal recognition triggers a conformational change of the kinase domains allowing interaction of the ATP-binding subdomains with the phosphorylation subdomains. Following autophosphorylation, the two subdomains dissociate and now the receiver domain of the response regulator can bind to the phosphorylation subdomain. After the His-Asp phosphotransfer, the receiver domain dissociates and a new cycle can start. If the phosphotransfer cannot take place, however, because
the aspartate residue has been exchanged, then the dissociation is inhibited and the complex remains stable. Such a complex might even offer a chance for structural analysis, which certainly could provide fascinating molecular details of the phosphotransfer reaction.

3.4.2. Characterization of the Periplasmic CitA Domain as a Highly Specific Citrate Receptor

In a previous study we proposed that citrate is the favorite signal recognized by CitA (Bott et al., 1995). Attempts to demonstrate this function with an *E. coli* in vivo system were not successful (data not shown). Therefore, we analyzed the periplasmic domain as a separate, soluble protein supplemented with a C-terminal His-tag (CitAP_{His}). With two independent methods, i.e. a classical binding assay with [14C]-citrate and isothermal titration calorimetry, we obtained congruent results showing that CitAP_{His}, which is a monomeric protein, binds citrate in a 1:1 stoichiometry with a K_D of ~5 μM at pH 7.0. None of the other tri- and dicarboxylic acids, including isocitrate and tricarballylate, were bound. Thus, CitAP_{His} functions as a highly specific citrate receptor, which indicates that the conformation of the protein is highly similar to that of the native periplasmic CitA domain.

The affinity of citrate binding to CitAP_{His} at different pH values closely reflected the appearance of the H-citrate^2^ form. As deduced from Table 1, the maximal affinity was found at pH 5.7 and decreased above and below this value, similar to the occurrence of the divalent citrate form (not shown). An obvious explanation of this result is that CitAP_{His} binds specifically the H-citrate^2^ species and therefore proton

<table>
<thead>
<tr>
<th>pH</th>
<th>n</th>
<th>K_A (10^3 M^-1)</th>
<th>K_D (μM)</th>
<th>ΔH_{ads} (kJ/mol)</th>
<th>ΔG (kJ/mol)</th>
<th>TΔS (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.0</td>
<td>0.98±0.01</td>
<td>6.26±1.33</td>
<td>16.7±3.6</td>
<td>-103.42±0.17</td>
<td>-27.33±0.54</td>
<td>-76.09±0.71</td>
</tr>
<tr>
<td>5.0</td>
<td>0.85±0.01</td>
<td>38.55±1.25</td>
<td>2.6±0.1</td>
<td>-97.81±1.67</td>
<td>-31.89±0.08</td>
<td>-65.92±1.72</td>
</tr>
<tr>
<td>6.0</td>
<td>0.80±0.01</td>
<td>51.65±2.95</td>
<td>1.9±0.1</td>
<td>-86.22±0.80</td>
<td>-32.60±0.13</td>
<td>-53.61±0.63</td>
</tr>
<tr>
<td>7.0</td>
<td>0.89±0.01</td>
<td>18.25±0.75</td>
<td>5.5±0.2</td>
<td>-76.43±0.21</td>
<td>-30.01±0.08</td>
<td>-46.37±0.08</td>
</tr>
<tr>
<td>8.0</td>
<td>0.85±0.01</td>
<td>9.05±0.09</td>
<td>11.1±0.1</td>
<td>-73.33±0.75</td>
<td>-28.29±0.04</td>
<td>-45.03±0.71</td>
</tr>
<tr>
<td>9.0</td>
<td>0.83±0.01</td>
<td>6.47±0.02</td>
<td>15.5±0.1</td>
<td>-74.88±0.17</td>
<td>-27.46±0.04</td>
<td>-47.42±0.17</td>
</tr>
</tbody>
</table>
release has to occur at pH values below pH 5.7, where the H$_2$-citrate$^-$ species becomes dominant, and proton uptake must take place at pH values above pH 5.7, where the citrate$^3$ form becomes prevalent. Taking into account that the H-citrate$^2$ species is recognized by CitAP$_{His}$, four functional groups are available for binding interactions, i.e. one uncharged and two charged carboxyl groups and the hydroxyl group. Current crystallographic data of macromolecules involved in binding of uncomplexed citrate clearly indicate that citrate binding is preferred in the extended conformation with important interactions of all four functional groups (Glusker, 1992; Remington et al., 1982; Russell et al., 1997). The fact that neither isocitrate nor tricarballylate were bound by CitAP$_{His}$ shows that the hydroxyl group at C3 is essential for binding. In this context, we can also explain the reduced binding affinity of citrate in the presence of Mg$^{2+}$ ions. The H-citrate$^2$ distribution calculated from pH profile changes dramatically in presence of Mg$^{2+}$ (Van der Rest et al., 1991). Using a stability constant of the Mg-citrate$^-$ complex of 1585 M$^{-1}$ (Sillén & Martell, 1964), the preferred H-citrate$^2$ species is reduced to less than 0.01% at pH 7.0 and a Mg$^{2+}$ concentration of 20 mM and nearly all of the citrate (96.2%) is complexed as Mg-citrate$^-$. Crystallographic data of metal citrate complexes generally show the hydroxyl group and either one or two of the carboxyl groups being involved in metal binding (Glusker, 1980). This means that the Mg-citrate$^-$ complex has to dissociate before the functional groups can interact with CitAP$_{His}$. This behavior is reflected by the occurrence of additional binding enthalpy in the presence of Mg$^{2+}$. The additional amount of binding enthalpy ($\Delta H_{obs}$) and of entropy ($\Delta S$) corresponds very well to the dissociation heat and entropy of the Mg-citrate$^-$ complex as measured by ITC (data not shown). In summary, the pH and Mg$^{2+}$ dependence of citrate binding to

Table 2. Thermodynamic parameters of citrate binding to CitAP$_{His}$ at 25°C in 50 mM sodium phosphate buffer pH 7.0 including different Mg$^{2+}$ concentrations as determined by ITC. The data were obtained from a single experiment for each Mg$^{2+}$ concentration and the errors quoted for n, K$_A$ and • H are those for the least-squares fit to the binding isotherms in the individual titrations.

<table>
<thead>
<tr>
<th>MgCl$_2$ (mM)</th>
<th>n</th>
<th>K$_A$ (10$^4$ M$^{-1}$)</th>
<th>K$_D$ (µM)</th>
<th>$\Delta H_{obs}$ (kJ/mol)</th>
<th>$\Delta G$ (kJ/mol)</th>
<th>$\Delta S$ (kJ/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.89 ± 0.01</td>
<td>18.25 ± 0.75</td>
<td>5.5</td>
<td>-76.43 ± 0.21</td>
<td>-30.01</td>
<td>-46.37</td>
</tr>
<tr>
<td>2</td>
<td>0.83 ± 0.01</td>
<td>7.78 ± 0.17</td>
<td>12.9</td>
<td>-79.02 ± 0.08</td>
<td>-27.92</td>
<td>-51.10</td>
</tr>
<tr>
<td>10</td>
<td>0.85 ± 0.05</td>
<td>2.47 ± 0.20</td>
<td>40.5</td>
<td>-89.15 ± 0.84</td>
<td>-25.07</td>
<td>-64.08</td>
</tr>
<tr>
<td>20</td>
<td>0.89 ± 0.03</td>
<td>1.74 ± 0.11</td>
<td>57.5</td>
<td>-89.57 ± 0.46</td>
<td>-24.19</td>
<td>-65.38</td>
</tr>
</tbody>
</table>
CitAP_His indicate that H-citrate$^2-$ is the preferred binding species in a non-restricted, extended conformation and that the inhibitory effect of Mg$^{2+}$ is not due to an interaction with CitAP_His, but only due to complex formation with citrate. Analysis of all ITC experiments for the binding stoichiometry (n) revealed values that varied between 0.85 and 0.95. In the binding assays with $[^{14}C]$-citrate, values of n were typically between 0.6 and 0.8. These data clearly show that one molecule of citrate binds per CitAP_His molecule, which is >95% monomeric, independent of whether citrate is bound or not. Deviations from the integer value are most likely due to inaccurate protein determination and/or to missfolded protein.

3.4.3. Comparison of the Citrate Binding Properties of CitAP_His and CitS

The CitA/CitB system induces the synthesis of a group of proteins involved in the anaerobic catabolism of citrate, including the citrate carrier CitS which is responsible for the uptake of the substrate. A comparison of the citrate binding properties of CitAP_His with those of CitS revealed a number of similarities: (i) Like CitAP_His, CitS is highly specific for citrate (Bandell et al., 1997). (ii) The apparent $K_M$ value of purified CitS was determined to be ~13 $\mu$M at pH 6 (Pos & Dimroth, 1996) and thus is in the same range as the $K_D$ value of the citrate/CitAP_His complex. (iii) Evidence has been provided that H-citrate$^2-$ is the citrate species recognized by CitS (Van der Rest et al., 1992a). (iv) Mg$^{2+}$ ions inhibited citrate transport by CitS, showing that the Mg-citrate$^-$ complex is not a substrate for the carrier (Van der Rest et al., 1992b). These similarities show that the citrate sensor CitA is very well adapted to the citrate transporter CitS.

3.4.4. A family of Two-Component Signal Transduction Systems Related to CitA/CitB

Database searches with CitA led to the identification of a group of sensor histidine kinases which presumably share a similar topology and modular structure and displayed sequence similarity not only within the conserved kinase domain, but also within the periplasmic and the linker domain (Fig. 8).

The protein most closely related to CitA from K. pneumoniae is the CitA protein from Escherichia coli. There is evidence that the CitA/CitB system of E. coli regulates the expression of the citCDEFXYZ gene cluster, encoding citrate lyase ligase (citC), the
three subunits of citrate lyase (citDEF), two proteins probably involved in the biosynthesis of the prosthetic group of citrate lyase (citXG), and a transporter (citT) which presumably catalyzes a citrate/succinate antiport (Pos et al., 1998). These proteins are required for citrate fermentation by E. coli and therefore a function of CitA as citrate sensor is obvious. Preliminary studies with the periplasmic domain of the E. coli CitA protein confirm its ability to bind citrate (Kaspar and Bott, unpublished). Besides the regulation of the citrate fermentation genes, the CitA/CitB two-component system of E. coli has also been implicated in the regulation of plasmid inheritance (Ingmer et al., 1998). In these studies, overexpression of CitB was found to destabilize the inheritance of intron-containing plasmids such as pSC101 and therefore CitB was named DpiA and the cognate sensor DpiB. However, we suggest to adhere to the CitA/CitB designation.

A second protein of E. coli showing significant similarity to K. pneumoniae CitA is the sensor kinase DcuS. This protein, together with its cognate response regulator DcuR, is responsible for the C4-dicarboxylate dependent induction of several proteins involved in fumarate respiration, including the dicarboxylate transporter DcuB and fumarate reductase (Golby et al., 1999; Zientz et al., 1998). Several results indicate that the periplasmic domain of DcuS functions as a receptor for a variety of C4-dicarboxylates, including succinate, fumarate, malate, tartrate, aspartate, and maleate.

Besides CitA and DcuS from E. coli, four other histidine kinases exhibited sequence similarity to K. pneumoniae CitA, i.e. CitS, YdbF, and Yufl from Bacillus subtilis (Kunst et al., 1997) and a Streptomyces coelicolor protein derived from gene 19c on cosmid 5B8 ((Redenbach et al., 1996); accession number AL022374). Neither the signals recognized by these proteins nor their target genes have been identified. However, a remarkable feature observed upon inspection of the vicinity of the chromosomal loci in B. subtilis was the presence of genes encoding proteins which are presumably involved in the transport of di- or tricarboxylic acids: (i) Immediately upstream of citS, the yflS gene is located encoding a protein with 35% sequence identity to the citrate/succinate antiporter CitT of E. coli (Pos et al., 1998). In addition, the third gene downstream of citS, designated citM encodes a secondary transporter for the Mg-citrate complex (Boorsma et al., 1996). (ii) Immediately upstream of B. subtilis ydbF, the ydbE gene is located encoding a protein with 41% sequence identity to the periplasmic C4-dicarboxylate binding protein DctP from Rhodobacter
Fig. 8. Sequence alignment of the CitA subfamily of histidine kinases. The transmembrane helices of CitA from *K. pneumoniae* and the characteristic H-, N-, G1-, F-, and G2-boxes (Parkinson & Kofoid, 1992) are indicated. The phosphorylated histidine residue within the H-box is marked by an asterisk. *Kp*, *Klebsiella pneumoniae*; *Ec*, *Escherichia coli*; *Bs*, *Bacillus subtilis*; *Sc*, *Streptomyces coelicolor*.
capsulatus (Shaw et al., 1991). Downstream of the ydbG gene for the cognate response regulator of YdbF, the ydbH gene is found encoding a protein with >50% sequence identity to the DctA C4-dicarboxylate transporters found in a variety of bacterial species. (iii) Five genes downstream of yufL, the yufR gene is located encoding a protein with >30% sequence identity to the Na⁺-dependent citrate carrier CitS from K. pneumoniae (Van der Rest et al., 1992b) and related proteins. In view of these data it is tempting to speculate that the CitA-related histidine kinases of B. subtilis serve as sensors for tri- or dicarboxylic acids and regulate the expression of genes encoding the corresponding transporters. In the case of the S. coelicolor citA paralog, a gene encoding a transporter of the sodium:solute symporter family is found upstream, but no hints exist on the substrate of this protein.

It is noteworthy that the periplasmic domains of the CitA-related sensor kinases have no apparent sequence similarity to the corresponding domains of DctB from Rhizobium leguminosarum (Ronson et al., 1987) and R. meliloti (Jiang et al., 1989) and DctS from Rhodobacter capsulatus (Hamblin et al., 1993). Both histidine kinases are involved in the transcriptional regulation of the DctA C4-dicarboxylate transporter.

**Fig. 9.** Sequence alignment of the CitB-subfamily of response regulators. The highly conserved residues within the N-terminal receiver domain (Volz, 1995) including the phosphorylation site (marked by an asterisk) and the proposed helix-turn-helix motif are indicated. The C-terminal DNA-binding domain encompasses approximately the lower half of the alignment. Kp: Klebsiella pneumoniae, Ec, Escherichia coli; Bs, Bacillus subtilis; Sc, Streptomyces coelicolor.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Kp-CitB</th>
<th>Ec-CitB</th>
<th>Bs-CitB</th>
<th>Sc-CitB</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seq1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seq2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seq3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seq4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seq5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Helix-Turn-Helix
and have been proposed to function as dicarboxylate sensors. In addition, there is also no sequence similarity to the periplasmic domain of the chemoreceptor Tcp from Salmonella typhimurium, which mediates chemotaxis towards citrate (Yamamoto & Imae, 1993).

All of the citA-like genes were found to be linked to genes encoding response regulators which showed significant similarity to *K. pneumoniae* CitB not only in the conserved N-terminal receiver domain, but also in the C-terminal output domain (Fig. 9). In particular, the helix-turn-helix motifs presumably involved in DNA-binding showed a high degree of conservation. This suggests that the CitA- and the CitB-related proteins co-evolved from a common ancestral two-component system and form a distinct subfamily within these universal signal transduction systems. Studies are in progress to identify the ligands of the CitA-related sensor kinases and the target genes of their cognate response regulators.

### 3.5. Experimental Procedures

#### 3.5.1. Bacterial Strains

*E. coli* DH5α (Bethesda Research Laboratories) was used as host for all cloning procedures and for overproduction of MalE-CitAC. *E. coli* XL1-Blue (Stratagene) was used for overproduction of MalE-CitAC-H350L. *E. coli* BL21(DE3), which contains the phage T7 polymerase gene under the control of the *lacUV5* promoter (Studier & Moffatt, 1986), served as host for overproduction of CitAPHis and the different CitB derivatives from the respective pET-based expression plasmids. The *E. coli* strains were routinely grown at 37°C in Luria-Bertani (LB) medium (Sambrook *et al.*, 1989).

#### 3.5.2. Construction of Expression Plasmids

For overproduction and purification of the kinase domain of the *K. pneumoniae* sensor kinase CitA, the expression plasmid pMalE-CitAC was constructed. To this end, a 2.18 kb *BamH1-HindIII* fragment from pGW234 (Woehlke *et al.*, 1992) containing the 3'-end of *oadB*, *citA*, and the 5'-part of *citB* was cloned into pBluescript KS+ (Stratagene), resulting in plasmid pKP3. Subsequently, a 0.76 kb *EcoRV-SalI* fragment of pKP3 was cloned into pMal-c (New England Biolabs) cut with *SstI* and *SalI*. The pMal-c vector contains a modified *E. coli* malE gene encoding maltose
binding protein without signal peptide (amino acids 2-26). The cloning strategy resulted in a protein in which the kinase domain of CitA (amino acids 327-547) was fused to the carboxy-terminus of MalE. The synthesis of this MalE-CitAC protein (604 amino acids, 66.8 kDa) was controlled by the tac promoter upstream of malE.

To introduce a mutation in MalE-CitAC which results in the replacement of the histidine residue at position 350 of CitA (position 407 in MalE-CitAC) by leucine, the QuickChange Site-Directed Mutagenesis Kit from Stratagene was used. The reaction was performed with the oligonucleotides mut1-f (5′-GCGCGCCGCTACAGCTCGAGCATCGCAATCTG-3′) and mut1-r (5′-CAGATTGCGATGCTCGAGCTGTACGGCGCGC-3′) and pMalE-CitAC as template (altered codon underlined). The desired plasmid pMalE-CitAC-H350L was identified by the XhoI restriction site introduced by the mutation and verified by sequencing. The corresponding protein was designated MalE-CitAC-H350L.

For overproduction and purification of the periplasmic domain of the K. pneumoniae sensor kinase CitA, the expression plasmid pET-CitAP was constructed. The region of the citA gene encoding the periplasmic domain (amino acids 45-176) was amplified from plasmid pKP12 (Bott et al., 1995) using primers that introduced an NcoI restriction site in front of codon 45 (pCitA-fo: 5′-CCGCCATGGACATTACCGAGGAGCGTCTG-3′) and an XhoI restriction site after codon 176 (pCitA-re: 5′-CGCCTCGAGTTGCTCAATGGTATAGCCTAC-3′). The polymerase chain reaction (PCR) was performed with Vent DNA polymerase (New England Biolabs). After digestion with NcoI and XhoI, the 0.45 kb PCR product was cloned into the expression vector pET24d (Novagen). The PCR-derived part of the resulting pET-CitAP plasmid and the ligation sites were controlled by DNA sequencing. The cloning strategy resulted in a protein designated CitAP_His (139 amino acids, 15.3 kDa) that contained besides amino acids 45-176 of CitA a methionine residue at the amino-terminus and six histidine residues at the carboxy-terminus. Synthesis of CitAP_His was controlled by a T7 promoter and the translation initiation signals were those of the gene 10 protein of T7.

For overproduction and purification of CitB and its individual domains, the expression plasmids described previously (Meyer et al., 1997) were used. To introduce a mutation into CitB_His which results in the replacement of the aspartate residue at position 56 by an asparagine, the Sculptor in vitro mutagenesis system (Amersham) was used. To obtain the single-stranded citB target DNA required by this system, a
1.8 kb HindIII-ClaI fragment from pKP12 (Bott et al., 1995) was cloned into the phage vector M13mp19 cut with HindIII and AccI, resulting in plasmid M13mp19-cit7. The mutagenic oligonucleotide D56N56 (5' - CTGTATCAACCGCAGCTCATCCTGCTGAATAATTTTTACCCGACGGC-3', altered bases underlined) was designed to replace the aspartate codon (GAT) by an asparagine codon (AAT) and to delete a PvuII restriction site (CAGCTG → CACCTC) simultaneously. The procedure was performed as described by the manufacturer. The desired mutant plasmids were pooled and a 303 bp citB-internal HindIII-SalI fragment including the mutation was used to replace the corresponding fragment of the plasmid pET-CitB, resulting in pET-CitB-D56N.

3.5.3. Overproduction and Purification of CitA and CitB Derivatives

For overproduction of the CitA and CitB derivatives, E. coli cells transformed with one of the expression plasmids were grown in LB medium containing either 200 μg/ml ampicillin (pMal-c derivatives) or 50 μg/ml kanamycin (pET24 derivatives). The cultures were incubated at 37°C and 180 rpm until the optical density at 600 nm (OD600) reached a value between 0.6 and 0.8. Then, expression of the desired genes was induced by the addition of 0.5 mM (pMal-c derivatives) or 1 mM (pET24 derivatives) isopropyl-β-D-thiogalactopyranoside (IPTG) and the cultures were incubated for another three hours at 30°C. Subsequently, cells were harvested by centrifugation, washed once in TKMD buffer (pMal-c derivatives) or in TN5 buffer (pET24 derivatives), and stored at -20°C. TKMD buffer consists of 50 mM Tris-HCl pH 7.5, 200 mM KCl, 5 mM MgCl₂, and 5 mM DTT. TN5 buffer (the number indicates the millimolar imidazole concentration) consists of 20 mM Tris-HCl pH 7.9, 500 mM NaCl, and 5 mM imidazole. For disruption, the cells were resuspended in the same buffer (4 ml/g wet weight) supplemented with 0.25 mg/ml DNase I and a protease inhibitor cocktail (Complete, EDTA-free) at the concentration recommended by the supplier (Boehringer Mannheim). After passing the cell suspension twice through a French pressure cell (SLM Aminco) at 108 MPa, intact cells and cell debris were removed by centrifugation (30 min at 27 000 x g, 4°C). The cell-free extract was subjected to ultracentrifugation (1 h at 150 000 x g, 4°C) to sediment the membranes. The supernatant was passed through a 0.2 μm filter and used for the isolation of the desired proteins.
MalE-CitAC and MalE-CitAC-H350L were purified by affinity chromatography on amylose resin (New England Biolabs). The soluble protein fraction derived from about 2 g cells (wet weight) was loaded onto a column containing 5 ml bed volume of amylose resin, which had been equilibrated with TKMD buffer. After washing with 40 ml of this buffer, MalE-CitAC or MalE-CitAC-H350L was eluted with TKMD buffer containing 10 mM maltose. In most experiments, the proteins obtained after affinity chromatography were further purified by gel filtration using a Superdex-200 HR 10/30 column integrated into an FPLC system (Pharmacia). The column was equilibrated with TKMD buffer and chromatography was performed at a flow rate of 0.4 ml/min.

Purification of CitA\textsubscript{His}, CitB\textsubscript{His}, CitB\textsubscript{His}-D56N, CitB\textsubscript{His}, and CitB\textsubscript{His}-D56N was performed by Ni\textsuperscript{2+} chelate affinity chromatography essentially as described previously for CitB\textsubscript{His} (Meyer \textit{et al}., 1997). Briefly, the soluble fraction obtained from 1-2 g cells (wet weight) was loaded onto a column with 2 ml (bed volume) His-Bind resin (Novagen) that had been equilibrated with TNI5 buffer. Weakly bound proteins were removed by washing with five bed volumes of either TNI30 buffer (CitA\textsubscript{His}), or TNI60 buffer (CitB\textsubscript{His}, CitB\textsubscript{His}-D56N), or TNI80 buffer (CitB\textsubscript{His}-D56N), or TNI100 buffer (CitB\textsubscript{His}). Elution was performed either with TNI200 buffer (CitB\textsubscript{His}-D56N, CitB\textsubscript{His}, CitB\textsubscript{His}-D56N) or with TNI400 buffer (CitA\textsubscript{His}, CitB\textsubscript{His}). Subsequent buffer exchanges were performed by gel filtration with Sephadex G-25 (PD-10 columns, Pharmacia). The purification of all proteins was monitored by SDS-PAGE (Laemmli, 1970) and subsequent staining with Coomassie brilliant blue.

3.5.4. Determination of Protein Concentrations

Protein concentrations of the CitB derivatives, MalE-CitAC and MalE-CitAC-H350L were determined according to Bradford (1976) using the Bio-Rad protein assay and ovalbumin as standard. CitA\textsubscript{His} concentrations were determined with the bicinechoninic acid method (Smith \textit{et al}., 1985) using the BCA protein assay reagent from Pierce. Alternatively, CitA\textsubscript{His} concentrations were estimated spectrophotometrically either at 280 nm using an extinction coefficient ($\varepsilon$) of 6.4 mM$^{-1}$ cm$^{-1}$ or at 276 nm using $\varepsilon = 7.25$ mM$^{-1}$ cm$^{-1}$ (Gill & von Hippel, 1989).
3.5.5. Phosphorylation Assays

Purified proteins (0.5-25 μM) were incubated at room temperature for 10 min and phosphorylation was started by the addition of γ[^32P]-ATP (1000 dpm/pmol) to a final concentration of 0.05 - 0.5 mM. In distinct time intervals, 5 -10 μl aliquots were removed, mixed with an equal volume of 2 x SDS loading buffer (124 mM Tris-HCl pH 6.8, 20% glycerol, 4.6% SDS, 1.4 M -mercaptoethanol, 0.01% bromphenolblue) and kept on ice until all samples had been processed. Subsequently, without prior heating, 5 - 10 μl of the samples were subjected to SDS-PAGE (12% separating gel). After drying the gel was analyzed with a Phospholmager (Molecular Dynamics).

3.5.6. Binding Assay with [14C]-Citrate

Citrate binding was measured essentially as described by Milligan and Koshland (1993). Briefly, purified CitAP_His (25 - 300 μM) was incubated on ice with varying concentrations of unlabelled sodium citrate (1 - 1000 μM) and a constant concentration (usually 6 μM) of 1,5-[14C]-citrate (83 mCi/mmol, NEN Life Science Products) in a total volume of 200 μl. After 10 min, 95-μl aliquots were added to tubes containing either 2 μl of a 1.0 M sodium citrate solution or 2 μl H2O. Subsequently, 80 μl aliquots of the two mixtures were subjected to ultracentrifugation using Microcon-3 (3 000 dalton nominal molecular weight cutoff) filter units (Millipore). In this way, free citrate (present in the filtrate) was separated from CitAP_His-bound citrate (present in the retentate). Free citrate was determined by liquid scintillation counting of 10 μl aliquots of the filtrate supplemented with 200 μl H2O and 5 ml scintillation fluid. The concentration of unbound citrate was calculated from the radioactivity measured in the filtrate of the +H2O tubes, and the concentration of citrate specifically bound by CitAP_His was calculated from the difference between the +1 M citrate and +H2O tubes. The data were analyzed by fitting the curve of [bound ligand] vs log [free ligand] to the Hill equation in the form $B = B_{\text{max}} F^n/(K_D + F^n)$, where $B$ is the concentration of bound ligand, $B_{\text{max}}$ the maximum number of binding sites, $F$ the concentration of free ligand and $n$ is the Hill coefficient. In addition, the binding data were analyzed according to Scatchard (Scatchard, 1949). To test the ligand specificity of CitAP_His, the 1 M citrate solution was replaced by 1 M solutions of
isocitrate, fumarate, succinate, tartrate, malate, or tricarballylate, all adjusted to pH 7. The data were evaluated as described by Clarke and Koshland (1979).

3.5.7. Isothermal Titration Calorimetry

All ITC measurements were performed using an OMEGA Microcalorimeter from Microcal, Inc. (Northampton, MA) with a cell volume of 1.3338 ml at a temperature of 25°C. Injections were made by use of a 250 µl microsyringe while stirring at 375 rpm. The calorimeter and the equations used to fit calorimetric data have been described in detail previously (Wiseman et al., 1989). The reference cell was filled with water containing 0.01% sodium azide and the instrument was calibrated using standard electrical pulses. Prior to loading into the microcalorimeter, all solutions were degassed for 10 min with gentle swirling under vacuum.

In order to minimize enthalpic contributions from proton transfer contributions due to ligand binding, all measurements were performed in 50 mM sodium phosphate buffer. Purified CitAP_His was dialyzed 36 h at 4°C against phosphate buffer of different pH (pH adjusted with NaOH). Solutions of the protein (62-180 µM) were filled into the sample cell and titrated with either citrate, isocitrate, or tricarballylate with a first control injection of 1 µl followed by 29 (citrate, tricarballylate) or 19 (isocitrate) identical injections of 5 µl. The ligand solutions were prepared by dissolving the compounds in the same buffer to concentrations 10-25 times higher than the protein solution.

The titration experiment was designed to ensure complete saturation of the enzyme before the final injection. The heat of dilution for all ligands was concentration-independent and corresponded very well to the heat observed from the last injections after the protein was saturated. Therefore, the baseline of the titrations of CitAP_His with ligands could usually be well estimated from the last injections of the titration. Data were collected, corrected for ligand heats of dilution and deconvoluted using the Microcal Origin software supplied with the instrument to yield binding constants (K_a), enthalpies of binding (ΔH), and stoichiometries of binding (n). The thermodynamic parameters were calculated from the basic equations of thermodynamics, i.e. ΔG = ΔH − TΔS = -R·T·ln K_a, where ΔG, ΔH and ΔS are the changes in free energy, enthalpy and entropy of binding, respectively, R the gas constant, and T the absolute temperature.
Acknowledgements
The authors would like to thank Prof. P. Dimroth and Prof. H.R. Bosshard for continuous support and Dr. I. Jelesarov for valuable suggestions throughout this work. This work was supported by a grant from the Swiss National Foundation for Scientific Research to M.B.

3.6. References


PART 4

Appendix
Appendix

1.1. List of Publications


1.2. Posters

Scapozza, L., Perozzo, R., Folkers, G., Structural Rearrangement of HSV1- Thymidine Kinase During Substrate Binding, 13th Protein Society Meeting (7/17- 23/99), Boston, USA
1.3. Oral Presentations

Kinematics of HSV1-Thymidine Kinase During Substrate Binding: A Thermodynamic and Crystallographic Study, 12/4/1999, Texas A&M University, College Station, Texas, USA

HSV1 Thymidine Kinase Substrate Binding Pathway: A Matter of Structural Rearrangement, 28-31/3/1999, ABC99 2nd International Conference on Applications of Biocalorimetry, Martin-Luther-University, Halle/S., Germany


New Insights into Ligand/Protein-Interactions by Isothermal Titration Calorimetry, 4/1/98, Campus University Irchel, Zurich, Switzerland

Characterization of Binding Interactions by Isothermal Titration Calorimetry, 10/15/1997, First Czech-Swiss Ph.D. Students Conference on Life Sciences, ICT Prague
1.4. Curriculum Vitae

1968 born on November 19th in Wetzikon, Switzerland

9/1989 High school diploma (Matura Type D), Kantonsschule Zürcher Oberland, Wetzikon

1990 Military Service (Education in medical service and supplies)

1990-1995 Study of pharmacy at the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland

1992-1993 12 months compulsory practical training at the pharmacy Hof-Apotheke, Rapperswil, Switzerland

12/1995 Graduation as federal qualified pharmacist (Eidg. Staatsexamen)

12/1995-4/1996 Chief Pharmacist at Regenbogen-Apotheke, Zurich

5/1996-9/1999 Ph.D. student at the Department of Pharmacy, Swiss Federal Institute of Technology (ETH) Zurich, Department of Pharmacy, section Pharmaceutical Chemistry, Prof. Dr. Gerd Folkers

7/1996-1/1999 Teaching assistant in pharmaceutical chemistry, Swiss Federal Institute of Technology (ETH), Zurich

10/1997 Marriage with Sandra Monsch

9/1999 Final examination to obtain the degree of Doctor of Natural Sciences, Swiss Federal Institute of Technology (ETH), Zurich
Vielen Dank!

Nachdem es nun endlich vollbracht ist, einen Teil der Arbeit von 3 Jahren und 5 Monaten Forschung auf ca. 200 Seiten Papier zu pressen, sollen an dieser Stelle endlich einmal die Menschen genannt werden, ohne die wohl nie etwas gegangen wäre....

An erster Stelle danke ich meinem Doktorvater Prof. Dr. Gerd Folkers, der es mir erst möglich gemacht hat, in seiner Forschungsgruppe mitzuarbeiten. Sein offenes Ohr für Probleme aller Art und seine Grosszügigkeit haben mir auch in den wildesten Zeiten immer weitergeholfen. Ebenfalls gebührt ein herzliches Dankeschön meinem direkten Betreuer Dr. Leonardo Scapozza. Er ist die Geduld in Person, ging nie einer wissenschaftlichen Diskussion aus dem Weg (trotz chronischem Zeittmangel seinerseits), und seine nahezu hellseherischen Fähigkeiten suchen seinesgleichen!

Vielen Dank auch an alle meine Kolleginnen und Kollegen vom M-Stock und an das Team von der Chemikalienausgabe.


Schliesslich möchte ich ein liebes Dankeschön an meine Familie richten, die letztendlich am meisten unter meinem Drang zur Forschung zu leiden hat.

Sandra, Alina und Marco, ich liebe euch!