Structural studies on Eukaryotic Translation Initiation Factors

and

Enzyme I of the Bacterial Phosphotransferase System

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von

Anselm Erich Oberholzer
von Goldingen/SG und St. Gallen-Tablat/SG

Leiter der Arbeit
Prof. Dr. Ulrich Baumann
Departement für Chemie und Biochemie
Structural studies on Eukaryotic Translation Initiation Factors
and
Enzyme I of the Bacterial Phosphotransferase System

Inauguraldissertation
der Philosophisch-naturwissenschaftlichen Fakultät
der Universität Bern

vorgelegt von
Anselm Erich Oberholzer
von Goldingen/SG und St. Gallen-Tablat/SG

Leiter der Arbeit
Prof. Dr. Ulrich Baumann
Departement für Chemie und Biochemie

Von der Philosophisch-naturwissenschaftlichen Fakultät angenommen.

Bern, den 25. Juni, 2004

Der Dekan
Prof. Dr. G. Jäger
Summary

The thesis is separated into three parts, of which the first two are closely related. The proteins investigated in these first two parts are eukaryotic translation initiation factors or translation initiation factor homologues. The protein investigated in the third part is a key enzyme of the bacterial phosphotransferase system.

The first part deals with two translation initiation factors, involved in the first step of translation initiation called cap recognition. Eukaryotic translation initiation factor 4G (eIF4G) is a scaffold protein with different binding sites for various initiation factors, including eukaryotic translation initiation factor 4A (eIF4A). eIF4A is an RNA-helicase which unwinds secondary structures in the untranslated 5’-end of the messenger RNA to generate a binding site for the ribosome. The interaction of eIF4A and eIF4G is absolutely essential for translation initiation.

Different gene-fragments of both proteins from *Saccharomyces cerevisiae* and *Homo sapiens* were cloned into bacterial expression systems and overexpressed in *Escherichia coli*. The purified fragments were used for binding studies to determine the eIF4A binding domain of eIF4G. The complexed fragments were used for crystallisation in order to solve the three dimensional structure by X-ray crystallography. Diffracting-quality crystals of three different complexes could be obtained, but structure solution was not possible until now.

The second part of the work is concerned with Ypr118w, a protein of initially unknown function from *Saccharomyces cerevisiae*, belonging to the translation initiation Factor 2B (eIF2B) α-, β-, δ-subunit family. It was thought to act as a translation initiation factor by replacing one of the homologous subunits of eIF2B. Ypr118w also shows similarity to
Methythioribose-1-phosphate isomerase from *Bacillus subtilis*, which itself is a member of the eIF2B α-, β-, δ-, subunit family, too. Methythioribose-1-phosphate isomerases are members of the methionine salvage pathway and convert methylthioribose-1-phosphate to methylthioribulose-1-phosphate. The methionine salvage pathway has recently attracted increased attention due to the findings that certain key enzymes of this pathway act as tumorous suppressors. Additionally, this pathway was targeted for anti-malarial drug design.

This work presents the determination of the three-dimensional structure of Ypr118w by X-ray crystallography. The gene encoding Ypr118w was cloned into a bacterial expression vector and overexpressed in *Escherichia coli*. From the purified protein crystals could be obtained which diffracted to 1.75 Å resolution. The structure was determined by the method of multiple wavelength anomalous dispersion employing seleno-methionine labelled protein crystals. The crystal structure contains two physiological dimers in the asymmetric unit. The monomer structure can be divided into two domains. The N-terminal domain forms a three-stranded antiparallel β-sheet followed by 6 α-helices with helices 1, 2, 4 and 5 building the core of this domain. Helix 5 bridges the N- and C-terminal domains. The C-terminal domain contains a central six-stranded, mostly parallel, β-sheet, surrounded by four α-helices and the C-terminal part of helix 6. The core resembles the well-known Rossman-type fold with three βαβ-units. A striking structural similarity to ribose-5-phosphate isomerases is evident and hints at a divergent evolution from an ancestral precursor.

Furthermore, this work presents genetic and enzymatic evidence showing that Ypr118w is the *Saccharomyces cerevisiae* 5-methylthioribose-1-phosphate isomerase.

The protein investigated in the third part is Enzyme I from *Thermoanaerobacter tengcongensis*. Enzyme I is the first enzyme of the phosphoenolpyruvate:sugar phosphotransferase system, which is responsible for the coupled phosphorylation and translocation of numerous sugars across the cytoplasmic membrane.
In this part, the determination of the three dimensional structure of the C-terminal domain of Enzyme I by X-ray crystallography is presented. The C-terminal part of the gene of Enzyme I was cloned into a bacterial expression system and the protein was overexpressed in *Escherichia coli*. From the purified protein crystals could be obtained which diffracted to 1.82 Å resolution. The structure was determined by the method of multiple wavelength anomalous dispersion, using seleno-methionine substituted protein crystals. The crystal structure contains two physiological dimers in the asymmetric unit. The monomer forms an eight-stranded α/β-barrel similar to the well-known TIM-barrel fold.
## Contents

Summary ........................................... 1

Abbreviations .................................... 5

Introduction ...................................... 6

Chapter I  Cloning, purification and preliminary crystallographic studies of complexes of different fragments of eIF4A and eIF4G ........................................... 18

Chapter II  Crystal structure of yeast Ypr118w, a methylthioribose-1-phosphate isomerase related to regulatory eIF2B subunits ........................................... 39

Chapter III  Crystal structure of the phosphoenolpyruvate-binding enzyme I-domain from the *Thermoanaerobacter tengcongensis* PEP: sugar phosphotransferase system (PTS) ........................................... 63

Danksagung ......................................... 90

Curriculum vitae .................................... 91
Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Å</td>
<td>Ångström</td>
</tr>
<tr>
<td>AMP</td>
<td>Adenosine monophosphate</td>
</tr>
<tr>
<td>BM</td>
<td>Bending magnet</td>
</tr>
<tr>
<td>β-ME</td>
<td>β-mercaptoethanol</td>
</tr>
<tr>
<td>CCD</td>
<td>Charge coupled device</td>
</tr>
<tr>
<td>CCP4</td>
<td>Collaborative computing project 4</td>
</tr>
<tr>
<td>Da</td>
<td>Dalton</td>
</tr>
<tr>
<td>DLS</td>
<td>Dynamic light scattering</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>DESY</td>
<td>Deutsches Elektronen-Synchrotron</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>eIF</td>
<td>Eukaryotic translation initiation factor</td>
</tr>
<tr>
<td>ESRF</td>
<td>European synchrotron radiation facility</td>
</tr>
<tr>
<td>GTP</td>
<td>Guanidine triphosphate</td>
</tr>
<tr>
<td>GST</td>
<td>Glutathion-S-transferase</td>
</tr>
<tr>
<td>HPr</td>
<td>Histidine containing protein</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl-thiogalactopyranoside</td>
</tr>
<tr>
<td>MAD</td>
<td>Multiple anomalous dispersion</td>
</tr>
<tr>
<td>MPD</td>
<td>2-methyl-2,4-pentane-diol</td>
</tr>
<tr>
<td>MTA</td>
<td>Methylthioadenoside</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>OD&lt;sub&gt;x&lt;/sub&gt;</td>
<td>Optical density at x nanometer</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide gelelectrophoresis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PTS</td>
<td>Phosphotransferase system</td>
</tr>
<tr>
<td>RCSB</td>
<td>Research collaboratory for structural bioinformatics</td>
</tr>
<tr>
<td>Rmsd</td>
<td>Root mean square deviation</td>
</tr>
<tr>
<td>RT</td>
<td>Room temperature</td>
</tr>
<tr>
<td>Se-Met</td>
<td>Selenomethionine</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodiumdodecylsulfate</td>
</tr>
<tr>
<td>SLS</td>
<td>Swiss light source</td>
</tr>
<tr>
<td>Tris</td>
<td>Tris-hydroxymethylaminoethane</td>
</tr>
<tr>
<td>XDS</td>
<td>X-ray detector software</td>
</tr>
</tbody>
</table>
Introduction

Three dimensional protein structures are a key for understanding of the function according to the dogma “form follows function”. The aim of this work was the elucidation of the three-dimensional structures of components from the eukaryotic translation initiation machinery. Specifically, the eIF4a-eIF4G complex was targeted as well as a homologue of the regulatory eIF2B subunits. A second project of the thesis was concerned with the determination of the conformation of Enzyme I, the key enzyme of the bacterial phosphoenolpyruvate:sugar phosphotransferase system (PTS). This introduction sheds light on the biological background and the state of research at the beginning of the thesis.

Initiation of translation in eukaryotes

Protein synthesis
Genes drive metabolism, growth and differentiation in living cells. They do this by promoting the synthesis of proteins which in turn catalyze numerous biological reactions. Translation is the process by which a protein is synthesised at the ribosome according to the nucleotide code provided by a messenger RNA.

The multi-step biochemical pathway of translation is conventionally divided into the three main steps of initiation, elongation and termination. The regulation of protein synthesis is the basis of cellular growth and differentiation. Translation regulation contributes significantly to overall regulation of gene expression in cells.
Translation initiation

The initiation is the most extensively studied step of the three main steps of translation and it is normally the overall rate-limiting step [1]. The initiation phase of translation consists of a series of events leading to the recruitment of translation-competent 80 S ribosomes to the initiation codon of an mRNA. The process is well conserved among eukaryotes and requires the participation of several initiation factors (eIFs) [2]. Two different mechanisms of translation initiation are known, cap-dependent initiation and internal initiation.

Cap-dependent initiation (overview)

In the cap-dependent initiation pathway (Figure 1) the initiation factor eIF4G binds eIF4E (the cap-binding protein) [3] and associates with the cap structure of the mRNA and poly(A)-binding protein (PABP), which itself is bound to the poly-(A) tail of the mRNA leading to circularization of the mRNA [4]. Recruitment of eIF4A and eIF4B prepares the 5' end of the mRNA for ribosome binding through ATP-dependent local mRNA unwinding [5]. The 80S ribosome dissociates into the 40S and the 60S subunit. The 40S subunit binds to a protein complex containing eIF1, eIF3, eIF5 and the ternary complex eIF2-GTP-Met-tRNA_i (TC). The resulting 43S preinitiation complex then associates with the mRNA through interaction of eIF3 with eIF4G at the cap structure [6]. From there the initiation complex moves in 5' to 3' direction and recognizes the AUG initiation codon through AUG-Met-tRNA_i base-pairing. This process is called scanning [7]. The factors eIF5 and eIF5B then trigger GTP hydrolysis, eIF2-GDP release, ejection of bound factors and joining of the 60S subunit to the 40S initiation complex to form an 80S ribosome competent for polypeptide elongation. The eIF2•GDP complex is recycled to eIF2•GTP. The GDP/GTP exchange is here catalyzed by eIF2B [8].
**Figure 1**: Model of cap-dependent translation initiation. The scheme was taken from the homepage of Prof. Dr. Hans Trachsel, University of Berne [http://ntbiomol.unibe.ch/].

**Cap recognition**

The first step in cap-dependent initiation is the recognition of the m$\textsuperscript{7}$G cap structure (Figure 2A), present at the 5'-end of all cellular m-RNAs, by eIF4E, which, together with eIF4G, forms the core of the eIF4F complex (Figure 2B). In mammalian cells, eIF4F contains a third subunit, initiation factor eIF4A. In yeast cells, until now, it was not possible to isolate an eIF4G-eIF4E-eIF4A complex. Nevertheless, also in yeast the direct binding of eIF4G-complexed eIF4E to the cap structure requires functional eIF4A [9].
Since the interaction of eIF4A and eIF4G is absolutely essential for translation initiation, we tried to crystallise complexes of different fragments of eIF4A and eIF4G. The knowledge of the three-dimensional structure of the complex would allow us to analyze the interactions between these proteins. This would greatly enhance the understanding of the initiation of translation. Results are summarized in chapter I.

**Ternary complex (TC) formation**

eIF2 can form a binary complex with GTP but not with Met-tRNAi. The eIF2-GTP binary complex is probably formed initially and Met-tRNAi is then bound subsequently. eIF2 consists of three subunits, eIF2α, eIF2β and eIF2γ. The eIF2γ subunit contains the GTP-binding site [10], and the eIF2β and eIF2γ are believed to be responsible for binding Met-
tRNAi [11]. The eIF2-bound GTP is hydrolyzed to GDP when the Met-tRNAi base-pairs with the AUG start codon. This leads to the release of eIF2 with its bound GDP.

*eIF2 recycling*

The released eIF2-GDP complex is bound by the guanine nucleotide exchange factor eIF2B, which exchanges GDP for GTP on eIF2 (Figure 3). The detailed mechanism of the nucleotide exchange reaction is still controversial: mammalian eIF2B was reported to act as an eIF2-GDP dissociation factor [8], whereas yeast eIF2B was claimed to bind GTP and to form a quartenary eIF2-GDP-eIF2B-GTP complex from which eIF2-GTP, eIF2B and GDP may be released [12]. The recycling of eIF2 by eIF2B is regulated by phosphorylation of a specific Ser residue on the α-subunit of eIF2 by a number of specific protein kinases ([13]). Phosphorylation of the eIF2α subunit results in an increased affinity of eIF2 for eIF2B. This renders the phosphorylated eIF2α subunit into a competitive inhibitor of the GDP/GTP exchange reaction. [14, 15].

![Figure 3: Recycling of eIF2](http://ntbiomol.unibe.ch/)

**Ypr118w**

eIF2B is a heteropentameric protein consisting of five subunits, α, β, γ, δ, and ε. The catalytic GDP-GTP exchange activity appears to reside primarily in the ε-subunit, enhanced in yeast by
the γ-subunit, whereas the α-, β-, and δ-subunits possess various regulatory roles [8, 14, 16, 17]. The yeast protein Ypr118w belongs to the eIF2B α-, β-, and δ-subunits family. Due to its homology, it was initially speculated that Ypr118w could replace one of the regulatory subunits of eIF2B or act as a regulator of eIF2 activity. Besides the well-characterized eIF2B subunits, the eIF2B α-, β-, and δ-subunits family also contains eubacterial and archaeal proteins which are also annotated as putative translation initiation factors. In bacteria the regeneration of IF2 seems to occur spontaneously. So there is no need for an IF2 recycling factor in prokaryotic translation. Therefore, other functions associated with those polypeptides might be more reasonable. One of these proteins from Bacillus subtilis has been characterised as a 5-methylthioribose-1-phoshate isomerase (MTNA) [18]. This enzyme participates in the methionine salvage pathway which is outlined below.

Methionine salvage pathway

The amino acid methionine is required for a number of important cellular functions, including the initiation of protein synthesis, the methylation of rRNA and xenobiotics, and the biosynthesis of cysteine, phospholipids and polyamines. The last function is especially important in rapidly proliferating cells, such as most parasites, bacteria, and cancer cells, which synthesize large amounts of polyamines [19]. The production of spermidine from putrescine and of spermine from spermidine consumes Met in the form of decarboxylated S-adenosylmethionine in a one-to-one stoichiometry, with methylthioadenosine (MTA) as a by-product. As the availability of Met is often limited, a unique pathway to recycle the amino acid from methylthioribose (MTR) exists (Figure 4).
Figure 4: The MTR recycling pathway in B. subtilis [20].

5-methylthioribose-1-phosphate isomerase catalyzes the isomerisation of 5-methylthioribose-1-phosphate (MTR-1-P) to 5-methylthioribulo-1-phosphate [20]. Ypr118w shares 37% sequence identity with the established 5-methylthioribose-1-phosphate isomerase from *Bacillus subtilis* and 26-28% identity with the eIF2B α-subunits from *Schizosaccharomyces pombe* and *Saccharomyces cerevisiae*.

Chapter II presents the crystal structure of Ypr118w, the first structure of a member of the eIF2B α-, β-, and δ-subunits family. Furthermore, we present genetic and biochemical evidence which enables to decide that Ypr118w does not function as a translation initiation factor but rather as the yeast *Saccharomyces cerevisiae* 5-methylthioribose-1-phosphate isomerase.
**Phosphoenolpyruvate:sugar Phosphotransferase System (PTS)**

The phosphoenolpyruvate:sugar phosphotransferase system (PTS) is responsible for the coupled phosphorylation and translocation of numerous sugars across the cytoplasmic membrane [21].

The primary phosphoryldonor is phosphoenolpyruvate (PEP). Phosphorylgroups are not directly transferred from PEP to the transported sugars but through a cascade of several phosphoproteins. These proteins together with the transporters constitute the core of the PTS.

The PTS further comprises proteins involved in gene regulation (antiterminators and transcription regulators) [22] and a variant of a dihydroxyacetone kinase [23]. The regulatory proteins control the expression and activity of catabolic enzymes in response to the availability of different carbon sources.

The PTS occurs in many facultative and obligatory anaerobic bacteria but not in eukaryotes. The size of the PTS varies for each species, between three and more than twenty proteins. The smallest PTS consist of Enzyme I (EI), the phosphorylcarrier protein HPr and at least one IIA subunit. EI and HPr are general compounds used by all phosphotransferase systems of all the different sugars. EI catalyses the phosphoryltransfer from PEP to the phosphorylcarrier protein HPr. EI is thus the source of all PTS activities. HPr subsequently transfers the phosphorylgroup to the different IIA proteins. The latter form a structurally and functionally heterogeneous group of proteins. Some are subunits of transporters, some are components of signal transduction, others are both. Larger PTS comprise in addition to EI and HPr one or several carbohydrate specific transporters (EII). EII always consist of three functional units termed IIA, IIB and IIC. IIA and IIB relay phosphorylgroups from HPr to the sugar, IIC are multispanning membrane permeases. They contain 6 membrane-spanning α-helices which form a channel through which the sugar transport takes place.
Figure 5: Phosphate transfer from phosphoenolpyruvate to glucose through components of the PTS. EI and HPr are general compounds used by all phosphotransferase systems, IIA, IIB and IIC are specific for the individual sugar compound.

EI is one of the most highly conserved proteins in bacteria with no counterpart in eukaryotes and therefore it has been considered as potential target for novel anti infectives (Mukhija et al., 2002). EI plays a pivotal role for all PTS activities and it was the objective of recent kinetic, mechanistic, enzyme inhibition [24] and virulence studies [25]. The amino-terminal domain (EIN, 1-250) harbours the active side His-189, which is transiently phosphorylated by PEP [26, 27]. The tree-dimensional structure of the N-terminal domain has been elucidated by X-ray crystallography [28] and its mode of interaction with HPr has been characterised by NMR spectroscopy [29]. It is composed of an HPr binding α-helical subdomain and an α/β subdomain which is structurally similar to the phosphor-histidine swivel domain of pyruvate phosphate dikinase PPDK [30]. The carboxy-terminal domain (EIC) contains the PEP binding site [31] and mediates the phosphorylation of His-189. EIC further mediates dimerisation [32] and confers species and subunit specificity during the phosphoryl transfer to HPr [33]. Its overall amino acid sequence is homologous to the pyruvate/PEP-binding domain of Pyruvate Phosphate Dikinase from Clostridium symbiosum and other PEP-binding enzymes [34]. EIC of E. coli is proteolytically unstable and flexible. Therefore, thermophilic organisms were thought to be more suitable sources for structural analysis. In chapter III we present the
crystal structure of the C-terminal domain of EI from *T. tengcongesis* and discuss its properties and implications for function.

References


**Chapter I**

**Cloning, purification and preliminary crystallographic studies of complexes of different fragments of eIF4A and eIF4G**

**Introduction**

The initiation phase of translation consists of a series of events leading to the recruitment of translation-competent 80 S ribosomes to the initiation codon of an mRNA. The process is well conserved among eukaryotes and requires the participation of several initiation factors (eIF) [1]. In the cap-dependent initiation pathway, the first step is the recognition of the m⁷G cap structure present at the 5'-end of all cellular mRNAs by eukaryotic initiation factor 4E (the cap-binding protein) which also interacts with the scaffold protein eIF4G [2]. This large protein in turn binds to the initiation factor 4A [3] and the poly(A)-binding protein, which binds the poly(A) tail at the 3'-end of most cellular mRNAs. This leads to mRNA circularisation [4]. RNA secondary structures in the 5'-untranslated region of the mRNA are then melted by the ATP-depended RNA helicase eIF4A in conjunction with eIF4B [5]. This creates a binding site on the mRNA for the 43 S preinitiation complex, which recognizes the mRNA through interaction of ribosome-bound eIF3 with mRNA bound eIF4G [6]. The preinitiation complex scans the 5'-untranslated region for the translation initiation codon [7].

eIF4G is a modular adapter protein that plays a pivotal role in coordinating the assembly of translation factors and the small ribosomal subunit during the rate-limiting initiation stage of protein synthesis [2]. Two related eIF4G proteins (eIF4GI and eIF4GII) encoded by two different genes exist in plants [8], mammals [9] and the yeast *Saccharomyces cerevisiae* [10]. This central component of the translation machinery contains a poly(A) binding protein
(PABP) interacting site [11, 12], a Tyr-X₄-Leu-Φ (Φ denotes a hydrophobic amino acid) eIF4E-recognition motif [13] and a phylogenetically conserved middle segment responsible for eIF4A binding [3].

The three dimensional structure of the middle domain of human eIF4GII has been elucidated by X-ray crystallography [14]. The crescent-shaped molecule consists of ten α helices, which form a right-handed solenoid. The structure shows similarities to various HEAT repeat proteins. HEAT repeat proteins participate in a wide variety of cellular processes that are dependent on assembling large multiprotein complexes [15].

![Figure 1: Ribbon drawing of the conserved central region of eIF4GII viewed along the cylindrical axes of the α helices [14].](image)

eIF4A is a member of the DEAD-box protein family (were DEAD corresponds to Asp-Glu-Ala-Asp). Yeast cells contain two and mammalian cells contain three closely related eIF4A proteins. These proteins share sequence elements and function in a large number of biochemical reactions [16]. eIF4A was shown to have RNA unwinding (RNA helicase) activity in the presence of eIF4B *in vitro* [17]. Since mRNA translation is inhibited by RNA secondary structures in the 5′ untranslated region [5] it is assumed that eIF4A and eIF4B play
a crucial role in RNA secondary-structure melting during the scanning process. Besides translation initiation, the DEAD-box family and related DExH- and DEAH-box proteins are involved in a variety of cellular processes including splicing, ribosome biogenesis and RNA degradation [18-20]. eIF4A consists of two domains, a ATPase domain and a C-terminal domain, which are connected with a extended 11-residue linker [21]. The C-terminal domain alone is sufficient to bind eIF4G.

The three dimensional structure of both domains has been elucidated by X-ray crystallography (Figure 2). The ATPase domain, containing the Walker A and B motif, has an αβα fold. The protein contains a central, seven stranded, twisted β sheet flanked by five helices on one side and four helices on the other side of the sheet. The C-terminal domain, containing the RNA binding motif, has a parallel α–β structure with the same topology as the equivalent domain of other helicases [22].

**Figure 2:** A: Structure of the ATPase domain of eIF4A. Ribbon diagram of the structure. The blue arrows represent β strands and the red ribbons represent α helices [23]. B: Structure of the carboxyl-terminal domain of eIF4A. Ribbon drawing of the structure. Conserved motifs are coloured as follows: motif IV, VIFCNTRR, residues 263-270, green; "conserved R" motif, residue Arg-298, purple; motif V, RGID, residues 321-324, magenta; motif VI, HRIGRGGR, residues 345-352, cyan. The strands of the β-sheet are labelled sequentially [21].
The interaction of eIF4A and eIF4G is absolutely essential for translation. Knowledge of the three-dimensional structure of the complex between eIF4A and eIF4G would allow to analyze the interactions between these proteins a deeper understanding of the molecular details of eukaryotic translation initiation could be reached.

**Materials and Methods**

**Cloning, Expression and Purification**

All constructs of eIF4A and eIF4G used were amplified by PCR from a *Saccharomyces cerevisiae* genomic DNA library or a *Homo sapiens* cDNA library with primers containing an *NdeI* restriction site at the 5'-end of the sense primer and an *XhoI* restriction site at the 5'-end of the antisense primer. The PCR-products were purified with a PCR-purification kit from Qiagen or Sigma and ligated into a pET28a or a pET22b Vector (Novagen). Competent XL1 BLUE (Stratagene) cells were transformed with the ligation mix and plated on LB-kanamycin agar plates for plasmid selection. DNA was isolated from several colonies and checked for the correct insert size by restriction analysis. Plasmid containing the eIF4A constructs was transformed into BL21 Star™ (DE3) (Invitrogen) cells, positive clones of the eIF4G constructs into BL21-CodonPlus™(DE3)–RIL (Stratagene) cells for protein expression. The cells were grown at 37° C to an A₆₀₀ of 1 and then induced with 1mM isopropyl-thiogalactopyranoside (IPTG) for 12-20 hours at 20° C. Typical yields were 40 mg/l pure protein for the eIF4A constructs and 10 mg/l for the eIF4G constructs. The expressed proteins were isolated by His-tag-chromatography on 10 ml of Ni-NTA SUPERFLOW (Qiagen) according to the manufacturer’s instructions. The (His)₆-tag of the pET28a constructs was cleaved off by thrombin digestion (Amersham Bioscience). The single components of the
Complexes were first purified by size exclusion chromatography (Superdex 75, Pharmacia, 20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.01 % NaN₃) or ion exchange chromatography (Resource Q, Pharmacia, 20 mM Tris-HCl, pH 8.0, linear NaCl gradient: 0-500 mM). This preliminary purification of the individual components greatly enhances the monodispersity of the complex formed later. (SeMet)-labelled proteins were prepared by the method of methionine biosynthesis inhibition [24] and purified as described above. All fragments were checked by mass spectroscopy or sequencing of the clone or by both methods, in order to insure no mutations were introduced by PCR. Mass spectrometric analysis (MS) were carried out by the analytical research services of the Chemistry Department, University of Berne.

**Complex Formation**

The single components were mixed together and concentrated at 4°C with a Centricon-10 (Millipore) concentrator. A second size exclusion chromatography was performed with the complex. All fractions were examined by Dynamic light scattering (DLS) prior to crystallisation attempts with a protein concentration of 1-2 mg/ml in GPC-buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 1 mM EDTA, 0.02 % NaN₃). Scattered intensities were recorded at 293 K using a DynaPro-MS molecular-sizing instrument equipped with a microsampler MS200/12 (Protein Solutions). Fractions with a Cp/Rh <20 % were pooled and used for crystallisation, whereas Rh is the hydrodynamic radius and Cp the standard deviation of Rh in nm.

**Binding Studies**

The known eIF4A binding site of eIF4G from *S. cerevisiae* (AA 542-853) was N-terminally shortened in steps of 10 AA and complexed with full length eIF4A. Binding studies
preformed with the eIF4A fragments were complexed with eIF4G(572-853)(K591R). All formed complexes were analysed by size exclusion chromatography, SDS-PAGE and DLS.

**Crystallisation and Data Collection**

All complexes were crystallised by vapour diffusion in sitting drops. The droplets consisted of 1 µl protein solution (10-20 mg/ml, 5 mM Tris, pH 7.5) mixed with an equal volume of reservoir solution. The following sparse matrix screens were carried out for each complex at 4°C and at 18°C: Crystal Screen I and II, Natrix- and PEG/ION-Screen from Hampton Research. For complexes eIF4G(572-853)(K591R)/eIF4A(9-395) and eIF4G(572-853)/eIF4A(226-395) Index Screen I and II from Hampton Research and Wizard Screen I and II from Emerald BioStructures were carried out additionally. Promising crystallisation conditions were optimised by fine grid screenings of pH, drop ratio, protein- precipitant- and salt concentrations as well as by Additive Screen I-III and Detergent Screen I-II from Hampton Research.

Diffraction data were collected in-house at 291 K and 110 K on a RaxisIV image plate area detector mounted on a Rigaku RU300 rotating-anode generator operating at 100 mA and 50 kV and at the Swiss Light Source at 100 K on beamline X06SA with a MAR CCD detector (Marresearch GmbH, Hamburg, Germany). Data were processed using XDS [25-27].

**Structure solution**

The molecular replacement method was used for structure determination. The N-terminal domain (RCSB PDB code 1QDE) [23] and the C-terminal domain (RCSB PDB code 1FUK) [28] of eIF4A from *S. cerevisiae* and the middle domain (RCSB PDB code 1HU3) [14] of eIF4G from *Homo sapiens* (33 % sequence identity) were used as search models. The molecular replacement was performed employing the program MOLREP [29] Version 7.3 as
implemented in the CCP4 interface. The solutions were inspected using the graphics program O [30].

Results and Discussion

Table 1 gives an overview of the cloned and expressed fragments of eIF4A and eIF4G. All eIF4G fragments are eIF4GI. The proteins were purified to homogeneity as judged from Coomassie-stained SDS-PAGE analysis.

Table 1: Overview of the cloned eIF4G- and eIF4A-fragments

<table>
<thead>
<tr>
<th>Species</th>
<th>Protein</th>
<th>Fragment size (AA)</th>
<th>Plasmid</th>
</tr>
</thead>
<tbody>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>542-883</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>552-883</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>562-883</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>572-883</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>542-862</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>552-862</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>562-862</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>572-862</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>542-853</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>552-853</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>562-853</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>572-853</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4G</td>
<td>582-853</td>
<td>pET28a</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>eIF4G</td>
<td>676-970</td>
<td>pGEX6P3</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4A</td>
<td>1-395</td>
<td>pET28a/pET22b</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4A</td>
<td>9-395</td>
<td>pET28a</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4A</td>
<td>226-395</td>
<td>pET28a/pET22b</td>
</tr>
<tr>
<td>S. cerevisiae</td>
<td>eIF4A</td>
<td>230-395</td>
<td>pET28a</td>
</tr>
<tr>
<td>Homo sapiens</td>
<td>eIF4A</td>
<td>1-406</td>
<td>pET28a</td>
</tr>
</tbody>
</table>

*aCloned and expressed by Siamak Djafarzadeh

*bPointmutation (K591A) (Figure 5a)


**Binding studies**

Former binding experiments performed by Diana Dominguez [31] suggested that the eIF4A-binding domain is located between residues 542 and 853. Figure 3 shows the results of the binding studies between full length eIF4A and different N-terminal shortened eIF4G-fragments. The smallest fragment with the full binding affinity reaches from residues 572 to 853. No stable complex could be purified with the eIF4G(582-853) construct.

The C-terminal domain of eIF4A alone is sufficient to bind eIF4G as mentioned above. eIF4A(226-395) shows almost similar binding affinity as full length eIF4A. eIF4A(230-395) shows a drastically reduced binding affinity. The four residues lacking in the eIF4A(230-395) construct are located within the flexible, 11-residue linker [21]. These residues are involved in eIF4G binding.

![Figure 3: Results of the binding studies of N-terminal shortened eIF4G-fragments.](image)

**Crystallisation**

Table 2 gives an overview over the crystallised complexes of eIF4G and eIF4A. Complexes crystallised under different conditions. From 4 complexes measurable crystals could be obtained (marked with a °).
### Table 2: Overview over the crystallised complexes of eIF4G/eIF4A

<table>
<thead>
<tr>
<th>eIF4G</th>
<th>eIF4A</th>
<th>DLS</th>
<th>Crystallisation</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>542-883</td>
<td>1-395</td>
<td>monodispers</td>
<td>needles</td>
<td>diffraction</td>
</tr>
<tr>
<td>542-883</td>
<td>226-395</td>
<td>monodispers</td>
<td>needles</td>
<td></td>
</tr>
<tr>
<td>552-883</td>
<td>1-395</td>
<td>monodispers</td>
<td>needles</td>
<td></td>
</tr>
<tr>
<td>572-883</td>
<td>1-395</td>
<td>monodispers</td>
<td>no crystals</td>
<td></td>
</tr>
<tr>
<td>572-883</td>
<td>226-395</td>
<td>monodispers</td>
<td>no crystals</td>
<td></td>
</tr>
<tr>
<td>552-862</td>
<td>1-395</td>
<td>monodispers</td>
<td>needles</td>
<td></td>
</tr>
<tr>
<td>552-853</td>
<td>1-395</td>
<td>monodispers</td>
<td>no crystals</td>
<td></td>
</tr>
<tr>
<td>552-853</td>
<td>226-395</td>
<td>monodispers</td>
<td>no crystals</td>
<td></td>
</tr>
<tr>
<td>562-853</td>
<td>1-395</td>
<td>monodispers</td>
<td>no crystals</td>
<td></td>
</tr>
<tr>
<td>562-853</td>
<td>226-395</td>
<td>monodispers</td>
<td>no crystals/sperulites</td>
<td></td>
</tr>
<tr>
<td>572-853</td>
<td>1-395</td>
<td>monodispers</td>
<td>crystals</td>
<td>diffraction</td>
</tr>
<tr>
<td>572-853</td>
<td>9-395</td>
<td>monodispers</td>
<td>plates</td>
<td>diffraction</td>
</tr>
<tr>
<td>572-853</td>
<td>226-395</td>
<td>monodispers</td>
<td>crystals</td>
<td>diffraction</td>
</tr>
<tr>
<td>676-970</td>
<td>1-406</td>
<td>polydispers</td>
<td>no crystals</td>
<td>aggregate peak in gel-filtration</td>
</tr>
</tbody>
</table>

^a*S. cerevisiae*

^b*Homo sapiens*

^cDiscussed below

**eIF4G(542-883)/eIF4A(1-395)**

The crystals obtained from the complex eIF4G(542-883)/eIF4A(1-395) are shown in Figure 4A. The crystallisation conditions are reported in Table 3. The initially obtained microneedles could be improved by the addition of 2 % (v/v) MPD (Figure 4B).
Figure 4: Crystals of eIF4G/eIF4A complexes. (A, B) Crystals of eIF4G\textsubscript{(542-883)}/eIF4A\textsubscript{(1-395)}. (C, D) Crystals of eIF4G\textsubscript{(572-853)(K591R)}/eIF4A\textsubscript{(1-395)}. (E, F) Crystals of eIF4G\textsubscript{(572-853)(K591R)}/eIF4A\textsubscript{(9-395)}. (G, H) Crystals of eIF4G\textsubscript{(572-883)(K591R)}/eIF4A\textsubscript{(226-395)}. Dimensions are indicated by bars. Crystallisation conditions are listed in Table 3.
Table 3: Crystallisation conditions of eIF4G/eIF4A complexes

<table>
<thead>
<tr>
<th>Crystals</th>
<th>Complex</th>
<th>Method</th>
<th>Crystallisation buffer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crystals A</td>
<td>eIF4G(542-883)/eIF4A(1-395)</td>
<td>Hanging drop, 4°C</td>
<td>0.2 M Calcium Chloride, pH 5.1 20 % (w/v) PEG 3350</td>
</tr>
<tr>
<td>Crystals B</td>
<td>eIF4G(542-883)/eIF4A(1-395)</td>
<td>Hanging drop, 4°C</td>
<td>0.2 M Calcium Chloride, pH 5.1 14 % (w/v) PEG 3350 2 % (v/v) MPD</td>
</tr>
<tr>
<td>Crystals C</td>
<td>eIF4G(572-853)(K591R)/eIF4A(1-395)</td>
<td>Sitting drop, 4°C</td>
<td>0.2 M Potassium Sodium Tartrate tetrahydrate, pH 7.2 20 % (w/v) PEG 3350</td>
</tr>
<tr>
<td>Crystals D</td>
<td>eIF4G(572-853)(K591R)/eIF4A(1-395)</td>
<td>Hanging drop, 4°C</td>
<td>0.2 M Potassium Sodium Tartrate tetrahydrate, pH 7.2 20 % (w/v) PEG 3350 10 mM AMP</td>
</tr>
<tr>
<td>Crystals E</td>
<td>eIF4G(572-853)(K591R)/eIF4A(9-395)</td>
<td>Sitting drop, 4°C</td>
<td>0.2 M Potassium Sodium Tartrate tetrahydrate, pH 7.2 20 % (w/v) PEG 3350</td>
</tr>
<tr>
<td>Crystals F</td>
<td>eIF4G(572-853)(K591R)/eIF4A(9-395)</td>
<td>Sitting drop, 4°C</td>
<td>0.2 M Potassium Sodium Tartrate tetrahydrate, pH 7.2 20 % (w/v) PEG 3350 3 % (v/v) MPD</td>
</tr>
<tr>
<td>Crystals G</td>
<td>eIF4G(572-853)(K591R)/eIF4A(226-395)</td>
<td>Sitting drop, 4°C</td>
<td>0.2 M tri-Ammonium Citrate, pH 7.0 20 % (w/v) PEG 3350 50 mM β-ME</td>
</tr>
<tr>
<td>Crystals H</td>
<td>eIF4G(572-853)(K591R)/eIF4A(226-395)</td>
<td>Hanging drop, 4°C</td>
<td>0.2 M tri-Ammonium Citrate, pH 7.0 20 % (w/v) PEG 3350 50 mM β-ME</td>
</tr>
</tbody>
</table>

Needle shaped crystals (100×4×4 µm³) (Figure 4B) were capillary-mounted and measured in-house at 291 K. It could be excluded that these needles are salt crystals but no diffraction pattern was observed, indicating that the crystal lattice is not well ordered or the crystals are too small.

*eIF4G(572-853)(K591R)/eIF4A(1-395)*

The molecular mass of eIF4G(572-853) was shown to differ by 34 Da from the calculated mass (32511.3 Da) by MS. The sequencing of the clone affirmed that Lys^{591} was changed to an arginine (Figure 5) (mass-difference: 28 Da). A sequence alignment with different eIF4G and related proteins shows that Lys^{591} is not conserved through the different species. It was decided
to continue with this clone, also with respect to the reduced flexibility of the Arg side-chain compared with Lys. This was thought to be helpful for crystallisation.

![Figure 5: a) Sequence alignment of the eIF4G sequence from the Swiss-Prot-Database (P39935) and the sequence of the clone. The different codons (AAG) (AGG) are marked with a black box. b) Multiple sequence alignment of eIF4G1 from S. cerevisiae (Swiss-Prot entry P39935, eIF4GII from S. cerevisiae (Swiss-Prot entry P39936, eIF4G1 from Homo sapiens (Swiss-Prot entry Q04637), eIF4G1 from O. cuninculus (Swiss-Prot entry P41110), eIF4GII from Homo sapiens (Swiss-Prot entry Q43432), eIF4G1 from A. thaliana (Swiss-Prot entry Q9LKQ7), eIF4GII from Mus musculus (Swiss-Prot entry Q62448. The mutated residue is marked with a green box. Residues corresponding to K591 are marked with a yellow box. The alignments were calculated with the program MultAlign [32] and the figure created with the program ESPript [33].

The crystals obtained from the complex eIF4G<sub>(572-853)</sub>(K591R)/eIF4A<sub>(1-395)</sub> are shown in Figure 4C. The crystallisation conditions are reported in Table 3. The initially obtained crystals could be improved by the addition of 10 mM AMP (Figure 4D).

Prior to freezing in liquid nitrogen, crystals were soaked in mother liquor without AMP containing 20 % Glycerol. Only one crystal could be frozen without any serious damages (cracks, disorder of the crystal lattice). Several other cryoprotectants were tested, namely
Glycerol (5-20 %, direct and serial transfer), Glycerol with 10 mM AMP (5-20 %, direct and serial transfer), MPD with 10 mM AMP (5-35 %, direct and serial transfer), PEG 400 with 10 mM AMP (5-40 %, direct and serial transfer).

The crystal was measured in-house at 110 K. The space group was determined as P2 or P2(1) respectively with unit cell parameters of a=68.84 Å, b=111.13 Å, c=112.04 Å, α=γ=90°, β=99.55°. The data collection statistics are reported in Table 4.

### Table 4: X-ray diffracton data statistics.

<table>
<thead>
<tr>
<th></th>
<th>eIF4G(572-853(K591R)/eIF4A(1-395))</th>
<th>eIF4G(572-853(K591R)/eIF4A(9-395))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P2(1)</td>
<td>P4(1)22 or P4(3)22</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td>a=68.84 Å, b=111.13 Å, c=112.04 Å, β=99.55°</td>
<td>a=b=67.70 Å, c=316.18 Å</td>
</tr>
<tr>
<td>Beamline</td>
<td>In-house</td>
<td>X06SA</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.5418</td>
<td>0.9999</td>
</tr>
<tr>
<td>Resolution rangea (Å)</td>
<td>20-2.90 (2.95-2.90)</td>
<td>20-3.77 (3.99-3.77)</td>
</tr>
<tr>
<td>No. measurements</td>
<td>128000</td>
<td>46675</td>
</tr>
<tr>
<td>Unique reflections</td>
<td>36641</td>
<td>13220</td>
</tr>
<tr>
<td>Rsymb (%)</td>
<td>10.2 (52.2)</td>
<td>9.9 (31.9)</td>
</tr>
<tr>
<td>I/σ(I)</td>
<td>9.15 (2.46)</td>
<td>11.4 (4.6)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.6 (100)</td>
<td>92.3 (74.5)</td>
</tr>
</tbody>
</table>

* The values in parentheses of resolution range, Rsym, I/σ(I) and completeness correspond to the outermost resolution shell.

* Rsym = Σhhkl∑j(I(hkl;j) - 〈I(hkl)〉)/((ΣhhklΣj(I(hkl)) where I(hkl;j) is the jth measurement of the intensity of the unique reflection (hkl) and 〈I(hkl)〉 is the mean over all symmetry-related measurements.

The molecular replacement was performed with the program MOLREP [29] using data in the resolution range 19.44 – 3 Å. The N-terminal domain [23] and the C-terminal domain [28] of eIF4A from S. cerevisiae and the middle domain [14] of eIF4G from Homo sapiens (33 % sequence identity) were used as search models. Only the N-terminal domain of eIF4A (> 30% of the scattering matter) could be located. A solution with two molecules per asymmetric unit was obtained in space group P2(1), corresponding to a Matthew’s coefficient of 2.7 and a solvent content of 54.2 % (Table 5). The correlation coefficient and the R factor of the solution were 55. % and 53.9 %, respectively (Table 6). A large psuedo-origin peak in the
native Patterson function at (0.4946, 0.5, 0.2495) indicated the presence of a local twofold axis parallel to the crystallographic one. This is in accord with the solution.

SDS-PAGE confirmed that all components of the complex are present in the crystal (Figure 6). In the full length structure of eIF4A [28] the N-terminal domain of eIF4A forms a crystal lattice by itself whereas the C-terminal domain is disordered and only a partial polypeptide backbone for one of the two molecules per asymmetric unit could be determined.

Inspection of the molecular replacement solution of the eIF4G(572-853)(K591R)/eIF4A(1-395) complex using the graphics program O [30] indicated that with two molecules per asymmetric unit the N-terminal domain of eIF4A could not form a well-packed crystal lattice by itself. Therefore the other components must also be involved in the crystal lattice formation and can not be flexible parts.

Other reasons besides flexibility for the failure to localise the missing domains might be that the complex formation could induce a conformational change in the C-terminal domain of eIF4A or in the middle domain of eIF4GII or both, hence this domains could not be located any more by molecular replacement. Despite of the 33 % sequence identity the middle domain of human eIF4GII could share insufficient structural similarity with eIF4G(572-853)(K591R) to be used as a search model.

A major problem was the reproducibility of the crystals. Until today we failed to regrow the crystals and therefore heavy-atom methods could not be used to explore these crystals further.

Table 5: Matthew’s coefficients of eIFG(572-853)(K591R)/eIF4A(1-395)

<table>
<thead>
<tr>
<th>Nmol/asym</th>
<th>Matthews Coeff</th>
<th>Solvent (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.4</td>
<td>77.1</td>
</tr>
<tr>
<td>2</td>
<td>2.7</td>
<td>54.2</td>
</tr>
<tr>
<td>3</td>
<td>1.8</td>
<td>31.4</td>
</tr>
<tr>
<td>4</td>
<td>1.4</td>
<td>8.5</td>
</tr>
</tbody>
</table>
Table 6: Molecular-replacement data of the N-terminal domain of eIF4A

a) Cross-rotation function.

<table>
<thead>
<tr>
<th>Peak No.</th>
<th>α (°)</th>
<th>β (°)</th>
<th>γ (°)</th>
<th>Peak height (σ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>145.99</td>
<td>43.42</td>
<td>90.14</td>
<td>12.31</td>
</tr>
<tr>
<td>2</td>
<td>259.28</td>
<td>50.22</td>
<td>67.51</td>
<td>4.67</td>
</tr>
<tr>
<td>3</td>
<td>12.56</td>
<td>31.63</td>
<td>32.87</td>
<td>4.52</td>
</tr>
<tr>
<td>4</td>
<td>115.13</td>
<td>35.88</td>
<td>225.40</td>
<td>4.40</td>
</tr>
<tr>
<td>5</td>
<td>54.72</td>
<td>85.26</td>
<td>97.66</td>
<td>3.91</td>
</tr>
<tr>
<td>6</td>
<td>289.40</td>
<td>48.61</td>
<td>25.37</td>
<td>3.89</td>
</tr>
<tr>
<td>7</td>
<td>213.19</td>
<td>47.03</td>
<td>248.01</td>
<td>3.88</td>
</tr>
<tr>
<td>8</td>
<td>129.76</td>
<td>35.79</td>
<td>224.98</td>
<td>3.84</td>
</tr>
<tr>
<td>9</td>
<td>316.57</td>
<td>77.51</td>
<td>344.56</td>
<td>3.81</td>
</tr>
<tr>
<td>10</td>
<td>354.45</td>
<td>48.38</td>
<td>219.53</td>
<td>3.81</td>
</tr>
</tbody>
</table>

b) Translation function with one monomer fixed. $X_{frac}$, $Y_{frac}$ and $Z_{frac}$ are fractional Cartesian coordinates.

<table>
<thead>
<tr>
<th>Peak No.</th>
<th>$X_{frac}$</th>
<th>$Y_{frac}$</th>
<th>$Z_{frac}$</th>
<th>$R$ factor</th>
<th>Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.953</td>
<td>0.500</td>
<td>0.834</td>
<td>0.539</td>
<td>0.553</td>
</tr>
<tr>
<td>2</td>
<td>0.939</td>
<td>0.500</td>
<td>0.330</td>
<td>0.575</td>
<td>0.495</td>
</tr>
<tr>
<td>3</td>
<td>0.447</td>
<td>0.000</td>
<td>0.581</td>
<td>0.907</td>
<td>0.128</td>
</tr>
<tr>
<td>4</td>
<td>0.945</td>
<td>0.907</td>
<td>0.082</td>
<td>0.713</td>
<td>0.234</td>
</tr>
<tr>
<td>5</td>
<td>0.945</td>
<td>0.090</td>
<td>0.082</td>
<td>0.710</td>
<td>0.236</td>
</tr>
<tr>
<td>6</td>
<td>0.956</td>
<td>0.207</td>
<td>0.833</td>
<td>0.663</td>
<td>0.252</td>
</tr>
<tr>
<td>7</td>
<td>0.887</td>
<td>0.003</td>
<td>0.129</td>
<td>0.706</td>
<td>0.217</td>
</tr>
<tr>
<td>8</td>
<td>0.122</td>
<td>0.501</td>
<td>0.858</td>
<td>0.720</td>
<td>0.213</td>
</tr>
<tr>
<td>9</td>
<td>0.945</td>
<td>0.648</td>
<td>0.081</td>
<td>0.720</td>
<td>0.219</td>
</tr>
<tr>
<td>10</td>
<td>0.757</td>
<td>0.499</td>
<td>0.293</td>
<td>0.722</td>
<td>0.212</td>
</tr>
</tbody>
</table>

Figure 6: SDS-Page (+ pole at bottom) of eIF4G(572-853)(K591R)/eIF4A(1-395). 4 crystals were washed 3 times in mother liquor and analysed to identify the crystal content. Both expected complex components were found. 1: Marker (66, 45, 24.6, 18.7, 14.3 kDa); 2: eIF4A (44.5 kDa), eIF4G (32 kDa).
Due to the failure to solve the structure with molecular replacement, (SeMet)-labelled protein was produced to perform a SAD/MAD experiment. Three different complexes were crystallised incorporating (i) only labelled eIF4A, (ii) only labelled eIF4G, and (iii) labelled eIF4A and eIF4G. Even though all three complexes crystallised under the same conditions as the native complex, none of the (SeMet)-labelled crystals could be frozen. The observed diffraction pattern could not be indexed. Several cryoprotectants were tested, namely Glycerol (5-20 %, direct and serial transfer), Glycerol with 10 mM AMP (5-20 %, direct and serial transfer), MPD with 10 mM AMP (5-35 %, direct and serial transfer), PEG 400 with 10 mM AMP (5-40 %, direct and serial transfer), Paratone N, silicone oil and paraffin oil.

\( \text{eIF4G}_{(572-853)(K591R)/eIF4A}_{(9-395)} \)

The eIF4A construct used for complex formation and crystallisation lacks the first eight N-terminal residues which are not conserved and not visible in the electron density of the N-terminal domain alone [23]. The removal of this flexible residues was thought to lead to better ordered crystals.

The crystals obtained from the complex eIF4G\(_{(572-853)(K591R)/eIF4A}_{(9-395)}\) are shown in Figure 4E. The crystallisation conditions are reported in Table 3. The initially obtained crystals could be improved by the addition of 3 % (v/v) MPD (Figure 4F).

Prior to freezing in liquid nitrogen, the crystals were soaked in mother liquor containing 35 % MPD. The crystals were measured at the Swiss Light Source at beamline X06SA at a temperature of 100 K. The crystals belong to the space group \( \text{P4\(_1\)}2_2 \) or its enantiomorph \( \text{P4\(_3\)}2_2 \) respectively with unit cell parameters of \( a=b=67.70 \text{ Å, } c=316.18 \text{ Å} \). The data collection statistics are reported in Table 4.

Due to the small dimensions of the crystals (200×200×5 μm\(^3\)) radiation damage was a serious problem during the experiment. So we either were not able to measure a complete
dataset from one crystal or we had to shift the crystal during the experiment to collect a dataset over the required 45 degrees. The data collection statistics above are from three different datasets merged together.

The molecular replacement experiment was performed as described in the Methods. None of the three search models could be located.

eIF4G(572-853)(K591R)/eIF4A(226-395)

The full length structure of eIF4A consists of two compact domains connected by an extended 11-residue linker which is highly flexible. [21] The C-terminal domain alone is sufficient to bind to eIF4G. By crystallizing a complex between eIF4G and only the C-terminal domain of eIF4A the flexibility of the complex should be reduced.

A complex between eIF4G(572-853)(K591R) and the C-terminal domain of eIF4A (AA 226-395) could be crystallised. The crystals we obtained are very small (10 x 10 x 5 µm³). Crystals are shown in Figure 4 (Crystals G, H). The crystallisation conditions are reported in Table 3. Crystals only could be obtained with 50 mM β-mercaptoethanol. The β-mercaptoethanol might prevent aggregation of the C-terminal domain of eIF4A formed by disulfide bridges of solvent exposed Cysteins. In such a high concentration the β-mercaptoethanol also could act as a small amphiphilic molecule, i.e. a crystallisation additive.

Prior to freezing in liquid nitrogen, the crystals were soaked in mother liquor containing 20% Glycerol. Data collected in house at 110 K diffracted to 5 Å. Data collected at the X06SA beamline at the Swiss Light Source diffracted to 2 Å. The diffraction pattern could not be indexed, because the crystals were highly twinned and disordered (Figure 7).
**Figure 7**: Diffraction image of eIF4G$_{572-853}$/K591R/eIF4A$_{226-395}$ collected at X06SA at the SLS. Exposure time, 1 s; crystal-to-detector distance, 150 mm; oscillation angle, 1.0°. A MAR CCD detector was used to record the image. The frame edge in the close-up view (right panel) is at 2.04 Å resolution.

**Conclusions**

Different complexes of eIF4G and eIF4A could be crystallised and their diffraction properties were analysed by X-ray crystallography. Datasets of two different crystal forms could be collected and processed. Nevertheless, structure solution was not possible until now. The N-terminal domain of eIF4A has been found by molecular replacement but the two other fragments could not be located due to reasons not understood. Crystal reproducibility proved to be extremely difficult preventing MIR/MAD methods so far.

The most promising crystals stem from the complex between the middle domain of eIF4G (AA 572-853) and the C-terminal domain of eIF4A. The flexibility of the complex should be
reduced by lacking the N-terminal domain of eIF4A which is flexible connected with the C-terminal domain. Improving the crystal quality of this complex would be the goal of future investigations. The solvent exposed cysteines of the C-terminal domain of eIF4A should be mutated against alanine or serine to prevent aggregations. Different detergent screens should be carried out to replace the β-mercaptoethanol which putatively acts as detergent.

References


Chapter II (Manuscript submitted to J. Biol. Chem.)

Crystal structure of yeast Ypr118w, a methylthioribose-1-phosphate isomerase related to regulatory eIF2B subunits

Mario Bumann\textsuperscript{1,2,§}, Siamak Djafarzadeh\textsuperscript{2,§}, Anselm Erich Oberholzer\textsuperscript{1,§}, Peter Bigler\textsuperscript{1}, Michael Altmann\textsuperscript{2}, Hans Trachsel\textsuperscript{2} & Ulrich Baumann\textsuperscript{1,*}

Affiliation:

\textsuperscript{1}Department of Chemistry and Biochemistry, University of Berne, Freiestrasse 3, CH-3012 Switzerland

\textsuperscript{2}Institute for Molecular Biology and Biochemistry, University of Berne Bühlstrasse 28, CH-3012 Switzerland

§ These three authors contributed equally to the work and are listed in alphabetical order.

* Corresponding author, e-mail: ulrich.baumann@ibc.unibe.ch, phone + 41 31 6314320, fax + 41 31 631 4887

Running Title: Structure and Function of Ypr118w

Key Words: eIF2B, methionine salvage pathway, translation initiation, crystal structure
Summary

Ypr118w is a non-essential, low-copy number gene product from *Saccharomyces cerevisiae*. It belongs to the PFAM family PF01008, which contains the α−, β-, and δ-subunits of eukaryotic translation initiation factor eIF2B, as well as proteins of unknown function from all three kingdoms. Recently, one of those latter proteins from *Bacillus subtilis* has been characterized as a 5-methylthioribose-1-phosphate isomerase, an enzyme of the methionine salvage pathway. We report here the crystal structure of Ypr118w which reveals a dimeric protein with two domains and a putative active site cleft. The C-terminal domain resembles ribose-5-phosphate isomerase (RpiA) from *Escherichia coli* with a similar location of the active site. In vivo, Ypr118w protein is required for yeast cells to grow on methylthioadenosine in the absence of methionine showing that Ypr118w is involved in the methionine salvage pathway. The crystal structure of Ypr118w reveals for the first time the fold of a PF01008 member and allows a deeper discussion of an enzyme of the methionine salvage pathway, which has in the past attracted interest due to tumor suppression and as target of antiprotozoal drugs.
Introduction

The *Saccharomyces cerevisiae* gene *YPR118W* is a non-essential gene on chromosome 16 encoding an acidic protein (pI 4.89) of 411 amino acids, Ypr118w, whose function is unknown (1) and which is of rather low abundance (922 molecules per cell (2). Ypr118w belongs to the PFAM family PF01008 and the TIGR 00512 and 00524 families, the latter also being called eIF2B-related (eIF2B_rel) (3-5). Members of PF01008 contain the α, β and δ – but not the catalytically active ε- subunit of eukaryotic translation initiation factor 2B (eIF2B) from yeast and mammals (4,5) (Figure 1). eIF2B is an important regulator of translation initiation. In eukaryotic translation initiation, a ternary complex consisting of Met-tRNA<sub>i</sub>, GTP and the heterotrimeric initiation factor eIF2 (6,7) associates with the 40S ribosomal subunit together with other initiation factors to form the 43S pre-initiation complex. This complex binds close to the 5'-end of mRNA and scans it in the 5' to 3' direction to localize the AUG initiation codon. AUG-recognition is mediated by codon-anticodon interaction and involves GTP hydrolysis stimulated by eIF5. This is a prerequisite for 80S initiation complex formation, *i.e.* the joining of the large ribosomal subunit and the release of initiation factors bound to the 40S ribosomal subunit. After its release the eIF2-GDP complex is bound by eIF2B which catalyzes the exchange of GDP for GTP. The recycling of eIF2 is modulated by posttranslational modifications of eIF2B (8,9) and by phosphorylation of the alpha subunit of eIF2 (10,11). Phosphorylated eIF2 binds 150 times more strongly to eIF2B than the non-phosphorylated form and therefore acts as a competitive inhibitor in the nucleotide exchange reaction, which consequently leads to inhibition of translation initiation (12).

Beside the well-characterized eIF2B subunits, the PF01008 family contains a subfamily of proteins (eIF2_rel) of mainly unknown function in eukaryotes, archaeae (e.g., *Pyrococcus*
furiosus) or eubacteria like Bacillus subtilis and Thermotoga maritima. Many of these proteins were initially annotated as putative translation initiation factors, despite the fact that there is no evidence for the requirement of an IF2 recycling factor in prokaryotic translation initiation. Recently, one of these proteins from Bacillus subtilis, Swiss-Prot entry MTNA_BACSU, has been functionally characterized as a 5-methylthioribose-1-phosphate isomerase (MTNA) (13). This enzyme participates in the methionine salvage pathway catalyzing the isomerisation of 5-methylthioribose-1-phosphate to 5-methylthioribulose-1-phosphate (14). The methionine salvage pathway leads to the synthesis of methionine from methylthioadenosin (for a scheme; see Figure 2A), the end-product of the spermidine and spermine anabolism in many species including mammals (15), Trypanosoma brucei (15), Saccharomyces cerevisiae (16), Klebsiella pneumoniae (17) and Bacillus subtilis (14).

The methionine salvage pathway has attracted much attention due to the finding that methylthioadenosine phosphorylase (MTAP), the first enzyme in the pathway, is deleted in a variety of human tumors (15) and acts as a tumor suppressor gene (18). This was attributed to the fact that products of the methionine salvage pathway regulate polyamine biosynthesis by inhibiting the activity of ornithine decarboxylase, an enzyme found to be upregulated in a variety of human and animal tumors (19). Furthermore, the methionine salvage pathway is a target for antimalarial drugs, as Plasmodium falciparum was shown to harbour a 5-methylthioadenosine phosphorylase which is inhibited by methylthioadenosine analogs in µM concentrations leading to growth inhibition (20).

Ypr118w shares 37 % sequence identity with the established 5-methylthioribose-1-phosphate isomerase from Bacillus subtilis and 26 - 28 % identity with the well-characterized eIF2B α-subunits from Schizosaccharomyces pombe and Saccharomyces cerevisiae. These findings raise the question of whether Ypr118w is a 5-methylthioribose-1-phosphate isomerase or a translation initiation factor. We solved the crystal structure of Ypr118w and present genetic
Evidence showing that Ypr118w is the yeast *Saccharomyces cerevisiae* 5-methylthioribose-1-phosphate isomerase.

**Experimental Procedures**

*Yeast strain*

Yeast BY4741 wild type and derivative strain Δypr118:: kanXR (Euroscarf collection) were used. The genotype of these strains is: MATa Δhis3 Δleu2 Δmet15 Δura3.

*Plasmid construct*

The YPR118w open reading frame was PCR-amplified from the start ATG to the stop codon from total yeast genomic DNA using primers to introduce a 5’ *SmaI* and a 3’ *SacI* site and subcloned into the multiple cloning site of yeast shuttle vector p301HIS3 under the control of the *GAL1/10* promoter resulting in the construct p301HIS3-YPR118.

*Cell growth on MTA*

Yeast cells were transformed with p301HIS3-YPR118 or p301HIS3 vector (negative control). Precultures of transformed cells were grown in SD or SDGal in the presence of all auxotrophic markers. Cells were washed with water and used to inoculate SD- or SDGal-cultures containing leucine, uracil and methionine (all at 20 µg/ml) or 5 mM MTA. Growth was monitored by measuring the cell density (A₆₀₀) after 40 hours of incubation at 30°C.

*Protein expression and purification*

DNA encoding full length *YPR118W* was amplified by PCR from total yeast genomic DNA using primers designed to introduce a 5' *SalI* site and a 3' *NotI* site. The PCR product was digested and inserted into pGEX-6P3 (Amersham Biosciences) expression vector. The
glutathione S-transferase (GST)-fusion protein was expressed in BL21-CodonPlus™(DE3)-RIL cells (Stratagene). Cells were grown to an A₆₀₀ of 0.9 in minimal media at 37°C and expression was induced by 1mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 3hr induction, cells were harvested and resuspended in Tris-buffered saline (100mM NaCl and 20mM Tris-HCl, pH 7.5) and disrupted using a French Press. After centrifugation, batch purification of GST-Ypr118w was performed by incubating the extract with glutathione Sepharose 4B beads (Amersham Biosciences) as described by the manufacturer, followed by three washes with TBS. The protein was eluted from the beads using reduced glutathione and the GST moiety was cleaved with Precission protease (Amersham Biosciences) at 4°C for 16hr. The protein was purified by gel filtration using a Superdex75 (Amersham) column equilibrated in 100 mM NaCl, 20 mM Tris-HCl, pH 7.5, 1mM EDTA and 1 mM DTT. The protein fraction was pooled, reincubated with glutathione Sepharose 4B beads and loaded onto a 1 ml Resource Q™ column (Amersham Biosciences), equilibrated with 20 mM Tris-HCl, pH 8.0. Bound protein was eluted with a linear NaCl gradient (0-0.5 M) in the same buffer. All purification steps were monitored by SDS-PAGE analysis (21). The protein was desalted by ultrafiltration using a Centriprep-10 (Millipore) with 5 mM Tris-HCl, pH 7.5 and concentrated to a final concentration of 10 mg/ml. Mass spectrometric analysis was carried out by the Analytical Research Services of the Chemistry Departement, University of Berne. Selenomethionine (SeMet)-labeled protein was prepared by the method of methionine biosynthesis inhibition (22) and purified following the same protocol as for wild-type protein.

**Crystallization**

Crystals of Ypr118w and (SeMet)-Ypr118w were obtained within 4 days by the sitting-drop vapour-diffusion method. Drops were set up by mixing 2 µl of protein solution (10 mg/ml) with 2 µl of reservoir solution (200 mM ammonium sulfate, 25 % PEG 3350, Bis-Tris, pH 5.5) and equilibrated against 120 µl reservoir solution at 291 K. Typical crystals were needle-
shaped with an average size of $30 \times 30 \times 300 \, \mu\text{m}^3$. They belong to the orthorhombic spacegroup $P2_12_12_1$ with cell parameters of $a = 60.7 \, \AA$, $b = 105.6 \, \AA$, $c = 263.3 \, \AA$ and contain four monomers per asymmetric unit.

Data collection

Before data collection crystals were flash-cooled in a nitrogen stream at 110 K after raising the glycerol concentration of the crystallization solution to 15 % (v/v). Data were collected in a MAD experiment on beamline BW7A at the EMBL Hamburg Outstation at the DESY at 100 K, employing a MAR-165 CCD (MAR X-ray-research, Hamburg, Germany) detector. A second MAD experiment was performed at the Grenoble ESRF MAD beamline BM14 at 100 K, employing a Mosaic 225 CCD (MAR X-ray-research, Hamburg, Germany) detector. Due to the alignment of the crystals, only 0.2 degree/frame could be collected. Exposure times varied from about 30 seconds (Hamburg) to about 5 seconds (Grenoble). All datasets were integrated and scaled with XDS (23,24).

Structure solution, refinement and analysis

The structure was solved by MAD using selenomethionine-labeled protein. The Hamburg as well as the Grenoble MAD datasets were sufficient to solve the structure, both yielding clean electron density maps. However, the Grenoble dataset extended to higher resolution due to the brighter beam and larger detector and was therefore later used to obtain an experimental density map suitable for automated model building.

Sixteen out of twenty expected selenium positions were determined by SHELXD (25). Phases were computed using SOLVE (26) and RESOLVE (27). Further density modification, phase extension and automatic model building was done by ArpWarp (28) version 6.0. Refinement was effected using REFMAC (29) version 5.1.24, for model rebuilding the program O (30) version 9.0.3 was used.
Data collection and refinement statistics are given in Table I.

Sequence alignment was done using CLUSTALW (31), Figure 1 was prepared using ESPRIPT (32). Comparison of three-dimensional structures was effected using DALI (33). Structure figures were created using the program PYMOL (www.pymol.org).

Results

Ypr118w is a 5-methylthioribose-1-phosphate isomerase

The sequence similarity between Ypr118w and 5-methylthioribose-1-phosphate isomerase from *Bacillus subtilis* led us to test whether Ypr118w is involved in the methionine salvage pathway. In this pathway, methylthioadenosine (MTA), a byproduct of the polyamine biosynthetic pathway, is recycled in cells by reconversion to methionine. This requires several enzymatic steps. The first step is the conversion of MTA to 5-methylthioribose-1-phosphate (MTR) by MTA-phosphorylase followed by its isomerization to methylthioribulose-1-P (Figure 2A). Recent studies led to the identification of some genes involved in the methionine salvage pathway in *Saccharomyces cerevisiae*. Among those, *Meu1* was identified as encoding the enzyme MTA-phosphorylase (16), whereas the gene encoding the MTR-isomerase could so far not be identified.

In order to find out whether Ypr118w is the MTR isomerase we tested whether a Δypr118w-knockout yeast strain auxotrophic for methionine was able to grow in a synthetic medium lacking methionine when supplemented with MTA. As opposed to the isogenic wild-type strain BY4741 (not shown) the Δypr118-knockout strain was not able to grow on MTA, neither in the presence of glucose nor of galactose (2%) as carbon source, but it grew well in the presence of methionine (Figure 2B). We then transformed the Δypr118-knockout strain with a plasmid carrying the gene *YPR118W* under the control of the galactose-inducible
GAL1/10 promoter. Cells transformed with this construct (but not with the control vector) were able to grow on MTA in the presence of galactose (Figure 2B) but not of glucose, clearly demonstrating that expression of Ypr118w was required to convert MTA into methionine. This indicates that YPR118W encodes for an enzyme of the methionine salvage pathway.

We also wanted to proof directly the isomerization of methylthioribose-1-phosphate to methylthioribulose-1-phosphate by Ypr118w. As methylthioribose-1-phosphate is not commercially available, we performed a two-step enzymatic reaction using recombinant Meu1p and 5-thiomethyladenosine to produce methylthioribose-1-phosphate. In a second step Ypr118w was added to the reaction mixture. We could readily detect the formation of a compound that was positive in an assay for reducing sugars (34). The NMR spectra of the expected compounds 5-methylthioribose-1-phosphate and 5-methylthioribulose-1-phosphate were identified (17), though the coupled reaction leading to 5-methylthioribulose-1-phosphate was not very efficient (not shown).

**Monomer structure**

The current model contains residues 1 to 407 without the amino acids 212 to 220 which are located in a disordered surface loop in all four monomers. This segment contains a unique and large insertion in the PF01008 family. The four crystallographically independent monomers do not show any significant deviation from each other, the RMS deviation of the Cα-positions is about 0.3 Å while single exposed surface residues show deviations of up to 1.4 Å.

The Ypr118w monomer structure has dimensions of 55 x 55 x 73 Å and can be formally divided into two domains (Figure 3A). The N-terminal 138 amino acids fold into a three-stranded antiparallel β-sheet followed by 6 α-helices with helices 1, 2, 4 and 5 building the core of this domain. Helix 5 is 38 amino acids long and bridges the N- and C-terminal
domains. This helix is kinked in the middle by about 15 degrees. The pollen allergen phl p 6 (PDB entry 1NLX) and the ATP synthase subunit C (PDB entry 1C17) were identified as closest structural homologs of the N-terminal domain with a DALI Z-score of 5.0, where 77 or 91 amino acids where aligned with a RMS deviation of 3.2 and 4.5 Å, respectively. The C-terminal domain possesses an $\alpha$$\beta$$\alpha$-fold resembling ribose-5-phosphate isomerase (RpiA) from *Escherichia coli* (PDB entry 1LKZ, DALI Z-score 11.6, 158 amino acids aligned with an RMS deviation of 3.2 Å) (35), and glutaconate coenzyme A-transferase from *Acidaminococcus fermentans* (PDB entry 1POI, DALI-score 8.4) (36). This domain contains a central six-stranded, mostly parallel, $\beta$ sheet, surrounded by four $\alpha$-helices and the C-terminal part of helix 6. The core resembles the well-known Rossman-type fold with three $\beta$$\alpha$$\beta$-units.

An “arm”, consisting of residues 341-367 wraps around the molecule with a short alpha-helix and two beta strands connected by a loop forming the tip (“hand”) of this segment (residues 338 to 372). This arm is rather unique in the whole family PF01008.

*Ypr118w is a dimer*

Gel filtration experiments indicate that Ypr118w is a dimer in solution (not shown). This was further corroborated by the crystal structure: the asymmetric unit contains two dimers which are identical despite rather different crystal contacts. A total of 5,379 Å$^2$ (30%) of accessible surface gets buried upon dimer formation, a surface much larger than a usual crystal contact.

The dimer has a “butterfly” shape (Figure 3B). The interactions between the monomers are mediated by the C-terminal domains, especially via the “arm” and “hand” and neighbouring segments (residues 335 to 383). Residues 231 to 260 including helices 7 and 8 also contribute to the interface.
**Putative active site**

A sulfate ion originating from the crystallization solution is deeply bound inside the cleft between the N- and C-terminal domains. It is coordinated via hydrogen bonds to the sidechains of Arg$^{51}$, Arg$^{106}$, Gln$^{237}$ and Lys$^{291}$ (Figure 4). The sidechain of Arg$^{51}$ is anchored via hydrogen bonds to the mainchain carbonyl oxygen and to the sidechain of Asn$^{182}$. These amino acids are highly, though not absolutely conserved. They vary especially in the case of bona fide translation initiation factors which supports the hypothesis that Ypr118w has a function different from translation initiation. The sulfate ion is located solvent-inaccessible deeply inside of the protein in a rather large cavity (volume 110 Å$^3$). We interpret the observation of such a tightly bound sulfate ion in all four crystallographically independent monomers as a putative substrate binding site, where a phosphate would occupy the position of the sulfate, as has been observed frequently. The cavity surrounding the sulfate binding site is in principle large enough to harbour a phosphorylated sugar molecule. Interestingly, after superposition the putative phosphate binding site is close to the active centre of ribose-5-phosphate isomerase RpiA (Figure 5) which further supports the hypothesis that these enzymes have evolved from a common ancestor.

**Conserved sequence motifs**

Kyprides and Woese, who first described the family PF01008 (3), defined six conserved sequence motifs within this family (Figure 1). Nearly all of them are located close to the putative active site assigned to the sulfate binding cavity (Figure1, Figure 3B and Figure 4). Motif 1 includes amino acids 51 to 57. Arg$^{51}$ is located at the start of helix 2 and binds probably to the phosphate moiety of the substrate.

Motif 2, which is located mainly in the turn between helix 3 and helix 4, comprises residues 105-110. Again, Arg$^{106}$ is likely to be involved in binding to a phosphate moiety.
Motif 3 includes amino acids 181 to 189 and possesses a cysteine (Cys\textsuperscript{181}) which is absolutely conserved in the MNTA (eIF2\textsubscript{rel}) subfamily of PF01008, but is absent in eukaryotic eIF2B subunits. A cysteine is also present in the active site of some non-homologous ribose-5-phosphate isomerase enzymes, e.g., in \textit{Escherichia coli} RpiB, where it may act as catalytic base (37).

Motif 4 includes residues 231 to 240 and is located between \(\beta\)-strand 5 and the beginning of helix 7. This motif contains Gln\textsuperscript{237} which coordinates to the sulfate ion and is replaced by serine or glutamic acid in most eIF2B subunits. Pro\textsuperscript{234} is in the \textit{cis}-conformation and absolutely invariant as is Gly\textsuperscript{238}, as even an alanine residue in this position would lead to steric hinderance. Motif 4 is also involved in the dimer interface. Two inter-subunit salt bridges related by the dimer twofold axis are formed by the Glu\textsuperscript{231} – Arg\textsuperscript{233} and Arg\textsuperscript{233} – Glu\textsuperscript{231} pairs. These two amino acids are also mostly conserved in the eIF2B subunits, although they are present as monomers in the pentameric eIF2-B holo-complex. The absence of the “arm” (residues 341-367), which is crucially participating in dimer formation, may explain the monomeric nature of eIF2B subunits.

Motif 5 covers amino acids 275 to 293. Lys\textsuperscript{291} is one of the putative phosphate ligands and is conservatively substituted by arginine in the mouse and human eIF2-\(\delta\) subunit but is a leucine in the \(\alpha\)-subunit from yeast. The invariant glycine residues exhibit backbone dihedral angles which are energetically unfavourable for other amino acids. Thr\textsuperscript{294} forms hydrogen bonds with the sidechain of Asn\textsuperscript{290} and the mainchain carbonyl oxygen of Lys\textsuperscript{291}. This motif also consists of Asp\textsuperscript{280}, a residue not conserved in eIF2B subunits where it is replaced by glutamate or histidine. Asp\textsuperscript{280} is completely buried and lines the putative active site pocket where it forms hydrogen bonds to a buried water molecule but does not contact the sulfate ion (Figure 4). \textit{Escherichia coli} RpiA carries at this position a conserved aspartic acid (Asp\textsuperscript{84}) which binds to the 3'-OH of the arabinose moiety of an inhibitor bound to the active site (38). Replacement of Asp\textsuperscript{84} reduces activity by a factor of 250 with only little change of the \(K_M\)
value. This aspartic acid is located in a five-amino acid consensus sequence G A D R/E I/V which is conserved in RpiA and eIF2B_rel and seems to be the only residue in the active centre which is invariant in RpiA and the eIF2B_rel family (Figure 5).

Motif 6 comprises amino acids 377 to 396. Phe<sup>379</sup> lines the active site cavity (Figure 4) and is close to the SO<sub>4</sub><sup>2-</sup> ion and to Arg<sup>106</sup> from motif 2. The absolutely conserved Asp<sup>380</sup> is completely buried and forms two hydrogen bonds with the mainchain amide groups from Ile<sup>293</sup> and Leu<sup>295</sup>. The C-terminal end of the motif is located on the surface of the molecule. The reason for the absolute conservation of Thr<sup>392</sup> is not clear.

**Discussion**

Our genetic experiments (Figure 2) together with the structural data indicate that the gene *YPR118W* encodes the yeast *Saccharomyces cerevisiae* 5-methylthioribose-1-phosphate isomerase, a protein required for the methionine salvage pathway: deletion of the gene *YPR118W* results in the inability of yeast cells to grow on methylthioadenosine (MTA) in the absence of methionine. Transformation of these cells with a plasmid carrying *YPR118W* restores growth on MTA in the absence of methionine. Additionally, we could detect *in vitro* the formation of 5-methylthioribulose-1-phosphate, the isomerization product catalysed by Ypr118w.

With regard to the sequence homology (37% identity) between Ypr118w and the recently characterized *Bacillus subtilis* 5-thiomethylribose-1-phosphate isomerase (13) we conclude that Ypr118w is indeed the yeast ortholog. There is no sequence similarity to any other enzyme of the methionine salvage pathway. Interestingly, the threedimensional structure
resembles ribose-1-phosphate isomerase A, which shows a hitherto unnoted lineage between different aldo-keto isomerases.

Also at the level of primary sequence Ypr118w resembles more the MTNA-subfamily of PF01008 than the eIF2B-αβδ subfamily. Especially pronounced are the differences at position 181 which is for all known and putative MTNAses a potentially catalytically active cysteine, while this residue is missing in all established eIF2B subunits. Similarly, Asp\textsuperscript{280}, a catalytically quite important residue is absolutely conserved in RpiA and MTNases but not in the eIF2B subunits.

The three-dimensional structure of Ypr118w reveals in its C-terminal domain similarity to ribose-5-phosphate isomerase RpiA, including the location of the putative active site and the conservation of a catalytically important aspartic acid. This is in agreement with the function of an aldose-ketose isomerase.

It is currently not clear whether all members of the eIF2B_rel family, \textit{i.e.} those PF01008 members which are not translation initiation factor subunits, are indeed methylthioribose-1-phosphate isomerases. The Swiss-Prot entry E2B_AQUAE, a homolog from \textit{Aquifex aeolicus}, has an N-terminal extension which codes for a thioesterase domain (PFAM PF03061). As mentioned above, the fold of Ypr118w is also related to glutaconate coenzyme a-transferase, an enzyme which catalyses the transfer of CoA from acetyl-CoA to the 1-carboxylate of glutaconate. This could hint at another function since a thioesterase-activity is not obviously required in the methionine salvage pathway. However, from the sequence alignment E2B_AQUAE fits seamlessly into the methylthioribose-1-phosphate isomerase branch of PF01008. Striking is the similarity between Ypr118w and the TREMBL entry Q9BV20. The latter is a human protein which is annotated as GTP-binding and involved in translation initiation. We assume that it is actually the Ypr118w ortholog. This protein is of potential
interest in cancer research: as mentioned above, ornithine decarboxylase activity is downregulated by the production of downstream metabolites of the methionine salvage pathway (16), and MTA, the first enzyme in the pathway, is a tumor suppressor. It will be interesting to examine the effect of the downstream enzymes on tumorigenesis.

Another open question is also whether the two homologous subfamilies (eIF2B-αβδ and eIF2B_rel) of PF01008 share some biological function. Binding of GTP and GDP to Saccharomyces cerevisiae eIF2B was established with binding constants of about 1 µM (39), however, the catalytic activity resides in the non-homologous ε-subunit. Therefore it seems to be likely that GTP or GDP binds to the ε-subunit, and not to the αβδ subunits. There are two other phosphate moieties involved in the biological function of eIF2B. First, eIF2B binds strongly to eIF2-α(P-51), the phosphorylated form of eIF2α. It was shown that in Saccharomyces cerevisiae the GCD2-GCD7-GCN3 subcomplex (i.e. the regulatory αβδ-subcomplex) mediates regulation of eIF2B activity in response to eIF2-α phosphorylation (40,41) while the γε-complex promoted release of GDP from eIF2 at a higher rate than the ε-subunit alone, but stable binding to eIF2 was not affected by eIF2-α phosphorylation (41). Hence the sulfate binding site determined by us in the Ypr118w structure could represent the eIF2-α-(Ser-51)-phosphate binding site in the eIF2B regulatory subunits. Second, it was reported that glucose-6-phosphate is required to maintain the activity of eIF2B by a hitherto unknown mechanism (42). Regarding the homology to RpiA and MTNA it is tempting to speculate that α, β and δ subunits of eIF2B bind glucose-6-phosphate. Apparently the ancestral protein possessed a sugar-phosphate binding/processing activity from which the two subfamilies of PF01008 and the RpiAs have emerged.
Acknowledgements

We thank Gustavo Chicaiza, Elisabeth Kislig and Sandra Nansoz for excellent technical assistance. This work was supported by the Swiss National Science Foundation (grant No 31-55423.98 to H.T. and No 31-67253.01 to U.B.), the Berner Hochschulstiftung and the University of Berne. A Roche Research Foundation fellowship to A.E.O. is gratefully acknowledged. We thank Martin Walsh from BM14 in Grenoble and Santosh Panjikar from the EMBL outstation at DESY in Hamburg for their invaluable help with data collection.

References

Figure 1: Sequence alignment. Coloring is according to similarity. Sequences were divided into two groups, one containing the putative MTNA (eIF2_rel) members of PF01008 (sequences 1-10) and the second comprising established eIF2B subunits (sequences 11 to 17). A red box with a white character means strict identity, a red character indicates similarity within a group. Red characters with a blue frame symbolize similarity between the two groups while an orange box indicates differences between the two conserved groups. The sulfate ion ligands have a cyan box. Sequence numbering is according to Ypr118w (YP18_YEAST). Q8T5K0 is annotated as translation initiation factor from Anopheles gambiae, Q9BV20 is a similar human protein sharing 41% identity with Ypr118w. E2B_AQUAE and E2B1_PYRHO are annotated as translation initiation factors from Aquifex aeolicus and Pyrococcus horikoshii, respectively. About 120 to 140 amino acid long N-terminal extensions occurring in the eIF2B-δ subunits and in E2B-AQUAE have been omitted from the alignment.
**Figure 2**

A. A scheme of the methionine salvage pathway is presented. Only the first steps leading to the production of 5-methylthioribulose-1-phosphate and the final product, methionine are presented.

B. Growth of a ypr118w-knockout strain on MTA. Cells transformed with construct p301HIS3-YPR118 (or control vector) were cultivated for 40 hours in minimal medium containing MTA or methionine in the presence of galactose or glucose as carbon source (for further details, see Materials and Methods)
Figure 3A:
Stereo cartoon representation of the monomer. The rainbow coloring runs from blue (N-terminus) to red (C-terminus). The sulfate ion is shown as sticks, the cavity in grey.
Figure 3B:
Cartoon representation of the dimer. One monomer is cyan, the other grey. The positions of the six conserved sequence motifs are indicated: motif 1 (red), motif 2 (magenta), motif 3 (yellow), motif 4 (blue), motif 5 (orange), motif 6 (green).
Figure 4: Close-up of the putative active site including the sulfate ion and the amino acids discussed in the text. Waters are shown as red spheres.

Figure 5: Overlay of RpiA from *Escherichia coli* (cyan, PDB code 1OB8) and Ypr118w (orange). The structures were superimposed using the program LSQMAN (43). Represented by sticks is the RpiA competitive inhibitor arabinose-5-phosphate and the conserved aspartic acid residues 84/280, as well as the sulfate ion found in the Ypr118w structure.
### Table I: Data collection and Refinement

<table>
<thead>
<tr>
<th>Hamburg BW7A</th>
<th>Peak</th>
<th>Inflection</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data collection (XDS)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9781</td>
<td>0.9785</td>
<td>0.9747</td>
</tr>
<tr>
<td>Resolution range(^a)</td>
<td>40-2.78 (2.95-2.78)</td>
<td>40-2.78 (2.95-2.78)</td>
<td>40-2.78 (2.95-2.78)</td>
</tr>
<tr>
<td>No. observations</td>
<td>467781(^b)</td>
<td>330284(^b)</td>
<td>313278(^b)</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>81766(^b)</td>
<td>81570(^b)</td>
<td>82182(^b)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.6 (98.9)</td>
<td>100 (100)</td>
<td>97.4 (89.8)</td>
</tr>
<tr>
<td>(R_{\text{sym}}) (%)</td>
<td>8.5 (24.6)</td>
<td>8.0 (22.9)</td>
<td>8.4 (23.8)</td>
</tr>
<tr>
<td>(I/\sigma(I))</td>
<td>17.6 (6.8)</td>
<td>15.0 (5.8)</td>
<td>13.9 (5.5)</td>
</tr>
<tr>
<td>FOM(^d) MAD (overall)</td>
<td></td>
<td>0.41</td>
<td></td>
</tr>
<tr>
<td>FOM(^e) solvent flattened</td>
<td></td>
<td>0.72</td>
<td></td>
</tr>
</tbody>
</table>

| Grenoble BM14                | Low remote            | Peak                    | Inflection            | High remote           |
|------------------------------|-----------------------|-------------------------|-----------------------|
| **Data collection (XDS)**    |                       |                        |                       |
| Wavelength (Å)               | 0.9879                | 0.9788                  | 0.9790                | 0.8856                |
| Resolution range\(^c\)       | 40-1.75 (outer shell) | 40-2.40 (2.56-2.40)    | 40-2.40 (2.56-2.40)   | 40-2.40 (2.56-2.40)   |
| No. observations             | 674837                | 509232\(^b\)           | 250563\(^b\)          | 512565\(^b\)          |
| No. unique reflections       | 169666                | 127394\(^b\)           | 124154\(^b\)          | 127394\(^b\)          |
| Completeness (%)             | 99.6 (98.9)           | 100 (100)               | 97.4 (89.8)           | 100 (100)             |
| \(R_{\text{sym}}\) (%)      | 7.3 (29.7)            | 7.6 (19.0)              | 5.3 (13.4)            | 9.3 (25.1)            |
| \(I/\sigma(I)\)             | 11.6 (4.1)            | 13.6 (6.5)              | 10.6 (5.3)            | 11.6 (5.1)            |
| FOM\(^d\) MAD (40 – 2.4 Å)  | 0.51                  |                         |                       |
| FOM\(^e\) solvent flattened | 0.76                  |                         |                       |

| **Refinement (REFMAC)**      |                       |                        |                       |
| Resolution range (Å)         | 30-1.75               |                         |                       |
| No. reflections working set  | 167734                |                         |                       |
| No. reflections test set     | 2208                  |                         |                       |
| No. non hydrogen atoms       | 13362                 |                         |                       |
| Solvent water molecules      | 934                   |                         |                       |
| Sulfate ions                 | 4                     |                         |                       |
| \(R/\text{free}\) (%)       | 17.1/19.8             |                         |                       |
| RMSD bond length (Å)         | 0.008                 |                         |                       |
| RMSD bond angles (deg.)      | 1.4                   |                         |                       |

\(^a\) The values in parentheses of resolution range, completeness, \(R_{\text{sym}}\) and \(I/\sigma(I)\) correspond to the outermost resolution shell

\(^b\) Friedel pairs were treated as different reflections.

\(^c\) \(R_{\text{sym}} = \frac{\sum_{hkl} |I(hkl,j)| - (\langle |I(hkl)\rangle)}{\sum_{hkl} (\langle |I(hkl)\rangle)}\) where \(I(hkl;j)\) is the \(j\)th measurement of the intensity of the unique reflection \((hkl)\) and \(\langle |I(hkl)\rangle\) is the mean over all symmetry-related measurements.

\(^d\) Figure-of-merit as computed by SOLVE.

\(^e\) Figure-of-merit as computed by RESOLVE.
Chapter III

Crystal structure of the phosphoenolpyruvate-binding enzyme I-domain from the *Thermoanaerobacter tengcongensis* PEP: sugar phosphotransferase system (PTS)*

Anselm Erich Oberholzer‡§, Mario Bumann‡§, Philipp Schneider‡, Christoph Bächler‡, Christian SieboldVert, Bernhard Erni‡* and Ulrich Baumann‡*

‡Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012 Bern, Switzerland

§Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Rosevelt Drive, Oxford OX3 7BN, UK

*Authors to whom correspondence should be addressed:

Departement für Chemie und Biochemie, Universität Bern, Freiestrasse 3, CH-3012 Bern, Switzerland; Tel. ++4131/631 43 46 (43 20); Fax ++41 (0)31/631 48 87; E-mail: bernhard.erni@ibc.unibe.ch, ulrich.baumann@ibc.unibe.ch

§ These two authors contributed equally to the work.

Running title: Crystal structure of the *T. tengcongensis* EIC domain of the PTS
Summary

The phosphoenolpyruvate: sugar phosphotransferase system (PTS) is a carbohydrate import mechanism unique to bacteria which tightly couples transport to substrate phosphorylation. The high-energy phosphate originates from PEP and is transferred sequentially to the sugar via various phosphoprotein intermediates. The first phosphate acceptor in this chain is enzyme I (EI), which in turn transfers the phosphate to HPr. Enzyme I is a multi-domain 64 kDa protein which is strongly conserved in bacteria. Therefore, it represents an attractive target for the development of novel antibiotics. In order to supplement structural information we present the crystal structure of the C-terminal domain of enzyme I of the thermophilic bacterium *Thermoanaerobacter tengcongensis* at 1.82 Å resolution. The dimeric protein folds in to an \((\beta\alpha)_8\) barrel with some large extrusions in the loops connecting \(\beta2\) and \(\alpha2\), \(\beta3\) and \(\alpha3\) and \(\beta6\) and \(\alpha6\). The dimer interface is very large burying 3749 Å² of accessible surface per monomer. A comparison to the homologous pyruvatephosphate dikinase reveals the conservation of the conformation of all the active site residues. The structure provides a rational explanation for the kinetic effects displayed in some recent inhibitor studies.
Introduction

The bacterial phosphoenolpyruvate: sugar phosphotransferase system (PTS) plays a dual role in transport and phosphorylation of sugars and hexitols on one hand and regulation of the cellular metabolism in response to the availability of these carbohydrates on the other. The molecular basis of its multiple functions is a protein phosphorylation cascade comprising four phosphoprotein units which sequentially transfer phosphoryl groups from phosphoenolpyruvate to the transported carbohydrates (for a review see [1]). Enzyme I (EI) is the first protein in this chain. It transfers phosphate from PEP to HPr, the general high energy phosphate carrier protein of the PTS (reviewed in [2]). HPr itself then serves as phosphoryl donor to the PTS transporters (EII, enzymes II) of different and sometimes overlapping substrate specificity. These transporters consist of three protein subunits or domains. The EIIA and EIIB domains serve as phosphate transfer relay to the carbohydrate substrate which is being translocated by the membrane spanning EIIC domain [3]. An increasing rate of sugar transport leads to a decrease of the phosphorylation of EII subunits which serves as the “input” for the PTS-mediated signal transduction. Phosphoryl transfer occurs through phosphohistidine and phosphocysteine intermediates. With PEP as substrate, EI links the protein phosphorylation relay of the PTS with glycolysis. PTS occur in eubacteria ([4-8] but not in eucaryotes and archaebacteria. EI of the PTS is one of the best conserved bacterial proteins with only minimal similarity to animal proteins [9].

EI is a 64 kDa two-domain protein [10-12]. The amino-terminal domain (EIN) contains a histidine which by double displacement transfers phosphate from PEP to HPr. The three dimensional crystal structure of EIN from *E. coli* has been elucidated, and its mode of interaction with HPr characterized by NMR spectroscopy [13, 14]. EIN is composed of the HPr-binding α-helical subdomain and an α/β subdomain which is structurally similar to the
phospho-histidine swivel domain of pyruvate phosphate dikinase (PPDK [15]), the latter bearing the phospho-histidine, which is His\textsuperscript{189} in \textit{E.coli}. The carboxy-terminal domain (EIC) contains the PEP binding site [16]. EIC of \textit{E. coli} plays a crucial role in reversible dimerization of EI [17, 18]. It is proteolytically unstable. The unfolding transition temperature of EI varies between 41°C in wild-type EI in the absence of substrates and 60°C in the H189A mutant in the presence of PEP and Mg\textsuperscript{2+} [19, 20]. Attempts to express the \textit{E. coli} EIC resulted in only partial success so far. The EIC (residues 240-572) fragment could be purified [21], but a twenty residues shorter variant could be purified only as fusion protein with the maltose binding protein MalE from which EIC was then released with a sequence specific protease (genenase) [22]. EIC of \textit{Mycoplasma capricolum} was purified in an active form [18].

The mode of action of \textit{E. coli} EI and how its activity is controlled is not yet fully understood. EIC is responsible for dimer formation which is strongly temperature-dependent [23]. The dissociation constant of \textit{E. coli} EI has been reported to shift from 20 µM at 6°C to 0.9 µM at 30°C. The presence of divalent cations (Mg\textsuperscript{2+} or Mn\textsuperscript{2+}) and PEP also affects the equilibrium and the association and dissociation rate constants. The association rate constant is 2-3 orders of magnitude slower than in other dimeric proteins, suggesting that dimerization is accompanied by major conformational rearrangements of the interacting EIC domains [24-27]. The activity of \textit{E. coli} EI at rate-limiting concentrations was enhanced by the addition of an inactive mutant [28, 29]. Changes of activity and fluorescence of pyrene labeled EI were correlated which was interpreted to reflect a correlation between function and dimerization [26, 27]. EI of \textit{M. capricolum} has a 10\textsuperscript{3} to 10\textsuperscript{5} times larger association constant, suggesting that reversible dimer-monomer transitions are less likely to be part of the catalytic mechanism [18]. The cellular concentration of EI is around 5 µM [30].
Whereas the structure of the N-terminal domain of *E. coli* EI has been determined [13, 14], the structures of EIC or full-length protein are not yet known. *E. coli* EIC has 28% sequence identity with the PEP-binding domain of pyruvate phosphate dikinase (PPDK), an enzyme from the C4 pathway of plants which catalyses the transfer of phosphate from ATP to pyruvate. The X-ray structure of the PPDK-substrate complex has been solved, and the C-terminal domain been shown to adopt an (βα)8 barrel fold [15, 31-33]. Here we present the X-ray structure of the EIC domain from *Thermoanaerobacter tengcongensis*, a gram-negative, saccharolytic bacterium which optimally grows at 75°C [34, 35]. This EIC displays 54% amino acid sequence identity with the *E. coli* EIC domain.

**Materials and Methods**

*Protein expression and purification*

*T. tengcongensis* full-length EI and the N-terminal (His)n-tagged construct of the C-terminal domain of Enzyme I (EIC(251-573)) were made by cloning the PCR product derived from a *Thermoanaerobacter tengcongensis* genomic DNA library into the *NdeI* and *BamHI* sites of the pET28a vector (Novagen) using the primers 5’-ACGTACATATGGAAGGATTAAGCAGTTAAAAG-3’ and 5’-ACGTAGGATCCTTAGCCAATATCTTTTATCACG-3’. EIC was expressed in Rosetta™(DE3) cells (Novagen). Cells were grown at 37°C to an A550 of 0.8 and the expression was induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG). After 12 hr induction, cells were harvested and resuspended in 20 mM Tris-HCl, pH 8.0, 400 mM NaCl, 10 mM β-mercaptoethanol, 0.2 mM PMSF and disrupted using a French press. EIC was purified on 20 ml Ni-NTA SUPERFLOW (Qiagen) according the manufacturer’s instructions. Typical yields were 30 mg/l pure EIC. After purification on Ni-NTA the eluate was
concentrated by Centriprep-30 (Millipore) to a volume of 2 ml. The protein was further purified by gel filtration using a Superdex 75 (Amersham Bioscience) column equilibrated in 10 mM HEPES, pH 7.5, 400 mM NaCl, 2 mM DTT. The purification steps were monitored by SDS-PAGE analysis [36]. The peak fractions were pooled and dialysed against 5 mM HEPES, pH 7.5, 2 mM DTT, 5 mM CaCl₂ for 2 hr at 4°C and concentrated to a final concentration of 10 mg/ml. Selenomethionine (SeMet)-labelled protein was prepared by the method of methionine biosynthesis inhibition [37]. Cells were grown at 37°C to an A550 of 0.6 and the expression was induced by 0.2 mM IPTG. After 12 hr of induction at 4°C, cells were harvested and the protein purified as the wild type.

**Crystallisation**

Trigonal crystals of EIC were obtained within 4 days by the sitting-drop vapour-diffusion method. Drops were set up by mixing 3 µl of protein solution (14 mg/ml) with 3 µl of reservoir solution (0.1 M sodium acetate trihydrate, pH 4.6, 12 % PEG 4000) and equilibrated against 120 µl reservoir solution at 291 K.

Crystals of selenomethionine-EIC were obtained within 4 days by mixing 2 µl of protein solution (10 mg/ml) with 2 µl of reservoir solution (0.1 M sodium acetate trihydrate, pH 4.6, 12 % PEG 4000, 7 % MPD). They belong to the orthorhombic space group P2₁2₁2₁ with cell parameters of a = 82.81 Å, b = 91.85 Å, c = 185.96 Å and contain four monomers per asymmetric unit. Typical crystals have an average size of 100 × 100 × 110 µm³. Only the selenomethionine-labelled crystals were suitable for structure determination.

**Data collection**

Before data collection crystals were flash-cooled in a nitrogen stream at 110 K after raising the MPD concentration of the crystallization solution to 35 % (v/v). Data were collected in a MAD experiment at the Grenoble ESRF MAD beamline BM14 at 100 K, employing a Mosaic
225 CCD (MAR X-ray-research, Hamburg, Germany) detector. Due to the alignment of the crystals in the cryoloop only 0.2 degree/frame could be collected. All datasets were integrated and scaled with XDS [38, 39]. Peak, inflection and high-remote data were collected using one crystal. For refinement purposes a high-resolution low-remote data set was collected on a different crystal which was larger in size. The two crystals turned out to be non-isomorphous. Data statistics is given in Table 1.

Structure solution, refinement and analysis
Sixty out of sixtyfour expected selenium positions were determined by SHELXD [40] using $F_{o}$-values computed by XPREP (G. Sheldrick, unpublished program) employing peak, inflection and high-remote wavelength. Phases were computed using SOLVE [41] and improved by RESOLVE [42]. Automatic model building was done by RESOLVE. 142 out of 324 residues were build by the program. Further attempts to use the low-remote high-resolution data for automated model-building failed, even after transporting the electron density into the unit cell of the high-resolution crystal using MOLREP [43]. Therefore, the remaining residues were placed manually into the electron density using the program O [44] version 9.0.3. This initial model was the placed into the unit cell of the low-remote data set followed by density modification, phase extension and automatic model building using ArpWarp [45] version 6.0. Refinement was effected using REFMAC [46] version 5.1.24. Refinement statistics are given in Table 1.

Structure-based sequence alignment was done by Indonesia [47]. Figure 3 was prepared using ESPript [48]. Comparison of structures was effected using DALI [49] and LSQMAN [50]. The structures of EIC of *E. coli* and *M. capricolum* were modelled using the SWISS MODEL server [51-53] using EIC of *T. tengcongensis* as template. Three-dimensional structures of E-
Cl-PEP and Z-CL-PEP were created using The Dundee PRODRG2 Server [54] and were superimposed on P-pyr, an inhibitor of pyruvate phosphate dikinase [33] using O.
Structure figures were created using the program PYMOL (www.pymol.org).

Results

Cloning, purification and crystallization of EIC from T. tengcongensis

The known dynamic behaviour of E. coli EI may compromise the formation of well-ordered three-dimensional crystals. In the past we subcloned the genes for EI of S. aureus, B. stearothermophilus, B. burgdorferi, E. coli, E. faecalis, H. influenzae and T. tengcongensis and overexpressed the proteins in E. coli. All EIs complemented PTS-activity in an E. coli ΔptsI mutant, and after purification displayed in vitro phosphotransferase activity, indicating that the proteins were active. EI<sup>AT</sup> of E. coli was expressed in inclusion bodies only. The recombinant EI-like domain of the multiphosphoryl protein FruB of P. aeruginosa could not be expressed in E. coli. At first EI of S. aureus appeared the most suitable of all. To prevent the protein from aggregation and to improve homogeneity, all surface-exposed cysteines were exchanged against alanine and serine. One out of 12 mutant proteins afforded crystals which diffracted to a resolution of 3.5 Å but crystal mosaicity and disorder of the crystal lattice prevented interpretation (results not shown, C. Siebold PhD thesis, University of Berne, 2002).

Thermophilic orthologues have frequently proven their use in structural biology in the past. Consequently, we focussed our efforts on Thermoanaerobacter tengcongensis, one of very few thermophilic organisms possessing a PTS. Ni<sup>2+</sup>-NTA purification of T. tengcongensis EI
expressed in *E.coli* afforded only fragments and almost no full-length protein. Four major fragments were sequenced and all were found to start with a methionine which hinted at internal translation initiation after rare codons. The most abundant fragment (>50%) turned out to be the complete C-terminal domain, EIC. The codons 252 to 573 of EI were PCR-amplified and the recombinant EIC domain expressed in an *E. coli* Rosetta strain which supplements tRNAs for rare codons. EIC could be purified to homogeneity by Ni$^{2+}$-NTA and gel-filtration chromatography. EIC is a stable homodimer with an unfolding transition temperature of 90°C (V. Navdaeva and B. Erni, unpublished results). While wildtype EIC yielded crystals of mediocre quality, selenomethionine-labelled EIC containing 16 Met (0.45 Met/kD) was prepared and crystallized in an orthorhombic spacegroup. These crystals were used for MAD experiments and structure refinement.

*Crystal structure determination*

Crystals of EIC belonged to the orthorhombic spacegroup P2$_1$2$_1$2$_1$ (a = 82.8 Å, b = 91.8 Å, c = 185.9 Å) with two physiological dimers in the asymmetric unit. Only reflections with l=4n of the (00l) reflections were strong and a native Patterson map at 4.5 Å resolution showed strong peaks (20% of the origin peak) on the sections w = 0.25 and w = 0.5. The crystals diffracted to a maximum resolution of 1.82 Å (Table 1). The structure of EIC was solved by multiple anomalous dispersion (MAD) at 2.8 Å resolution. Using the peak wavelength alone resulted in better correlation coefficients in spacegroup P2$_1$2$_1$2 than in the true space group, but no interpretable electron density was obtained. Using the right space group and $F_s$-values yielded eventually a solution with a correlation coefficient of 70% and an interpretable electron density.

Refinement of the model against 1.82 Å data resulted in reasonable $R$-factors and satisfying model geometry (Table 1). 1272 residues (92.7%) are in the most favoured regions and 90
residues (6.6%) in the additional allowed regions of the Ramachandran plot. The 4 residues (0.3 %) in disallowed regions are well defined in the electron density. Residues 251-572 of the EIC monomer are well ordered in all monomers, only the starting methionine 250 in one of the four monomers is disordered.

**Overall structure**

EIC features a (βα)₈ barrel fold (Figure 1A). The four crystallographically independent monomers of the asymmetric unit are very similar. The RMS deviation of the Cα-positions over the entire length of a molecule is 0.35 Å. Only two solvent exposed loops (302-308 and 337-351) have a more than average RMS deviations of 0.7 Å and 1.5 Å, respectively (Figure 1A). These loops are engaged in crystal contacts in 2 of the four monomers. The (βα)₈ barrel has an extra N-terminal α-helix and three extensions on the C-terminal face of the barrel. The first extension (296-309) which is inserted between β2 and α2 has an irregular structure. The second extension (333-365) between β3 and α3 comprises two short α-helices, and the third (452-479) between β6 and α6 comprises one extra α-helix. The three extensions protrude approximately 30 Å above the core of the barrel. The active site cysteine (see below) is located in the βα7 turn and accessible over the short βα8 and βα1 turns at the rim of the barrel. The shape of EIC can be compared with an “easy chair”, the three extensions forming the backrest and the C-terminal end of the barrel the cushion (Figure 1B).

EI of *E. coli* is a homodimer (Figure 1C) and EIC was identified as the dimerization domain [18]. The asymmetric unit of *T. tengcongensis* EIC contains two identical dimers. There is an extensive contact area comprising 45 residues which cover 3’749 Å² per monomer between the two monomers. 15 % of the EIC subunit surface becomes buried upon dimer formation. This large interface is formed mostly by back to back contacts between the β3/α3 and β6/α6
extensions (the backrest, Figure 1C). Additional contacts are provided by the β4/α4 and β5/α5 turns and the first half of helix 6. The interface is composed mainly of hydrophobic residues. The contact area of more than 3'700 Å² is exceptionally large for a dimerization domain which usually are of the order of 1200-2400 Å² [55].

Relationship to pyruvate phosphate dikinase

A comparison of the amino acid sequence of the maize and *Bacillus symbiosus* pyruvate phosphate dikinase (PPDK) with other protein sequences revealed homology to the mechanistically related enzymes PEP synthase and EI of the PTS [56]. PPDK catalyzes the reversible conversion of ATP, inorganic phosphate (Pᵢ) and pyruvate to AMP, pyrophosphate and phosphoenolpyruvate (PEP) [15]. The sequence similarity extends from the α/β fold (swivel domain) which contains the active site histidine to the C-terminal PEP/pyruvate binding domain [15].

EIC and the C-terminal PEP/pyruvate binding domains of PPDK (PDB entry 1DIK) display 25% sequence identity and are the closest structural homologues in the protein data bank with a DALI Z-score of 31.0. Both possess an (β/α)₈ fold and 279 amino acids out of the 360 of PPDK and 320 of EIC can be aligned with a RMS deviation of less than 2.2 Å.

Using INDONESIA, 268 residues of EIC can be superimposed on PPDK with bound phosphonopyruvate (PDB entry 1KC7) with an RMS deviation of 1.6 Å for the paired Cα-atoms (Figure 2). Figure 3 shows a structure-based sequence alignment between EIC and the pyruvate/PEP binding domain of PPDK.

EIC shares with the PEP/pyruvate binding domain of pyruvate phosphate dikinase the three extensions and the large intersubunit contact area. The PPDK extensions are however longer. The β/α2 and extensions features an α helical hairpin. β/α3 and β/α6 also have more α-helical structure. 32 out of the 45 residues making up the dimer interfaces of EIC and PPDK
are structurally conserved (Figure 3). The PEP binding site at the C-terminal end of the $\beta/\alpha$ barrel are almost completely superimposable (see below).

**Substrate binding site**

The PEP-binding site of E. coli EI has been identified by affinity labelling with the mechanism-based inhibitor Z-chloro-PEP [29] which labelled Cys-502 covalently during catalytic turnover. In PPDK, the structural details of the binding site have been characterized with crystals of PPDK in complex with the substrate analogue 3-phosphonopyruvate (P-pyr). In the phosphonopyruvate the C=O group isostERICally replaces C=CH2, while the -CH2-PO3 replaces the -O-PO3 of PEP. A rich network of interactions ensures the precise positioning of the P-pyr ligand in the PPDK binding site [33]. A superimposition of the PPDK/P-Pyr structure with EIC shows that all the residues that form the charge network with the substrate analogue and Mg$^{2+}$ are also conserved in EIC (Figure 4). P-pyr can be accommodated at the C-terminal end of the $\beta/\alpha$-barrel of the T. tengcongensis EIC structure without sterical clashes and constraints. The side chain conformations of the active site residues are remarkably well conserved, and this independently of whether a substrate is bound as in PPDK or not as in EIC. This may be taken as evidence for a rather rigid and inflexible structure of the PEP binding sites of PPDK and T. tengcongensis EIC.

The reaction mechanism at the PEP binding site of E. coli EI has previously been characterized with C-3 modified PEP-analogues [29, 57]. The K$_m$ of E. coli EI for Z-CIPEP (Z: COOH and Cl trans to C=C double bond) and E-CIPEP were found to be similar (0.04 mM and 0.12 mM) to the K$_m$ for PEP (0.14 mM), while the k$_{cat}$ values were 2 and 3 orders of magnitude lower for the Z- and E-isomer, respectively. The Cl-PEP isomers were modelled in the EIC binding site with the carboxylate, C2 and phosphate in the same position as those of P-Pyr in the PPDK model. Under these conditions the orientation of the bulky and
electronegative chloride atoms at C3 are different for both isomers. Due to sterical clashes, the chlorine atom of Z-ClPEP (Figure 5B) may change the position of the Z-ClPEP nearer to the surface whereas the E-ClPEP (Figure 5A) still fits in the binding site similar to PEP. The C-3 of Z-ClPEP is turned away from the general acid/base Cys-502 which may lead to a much slower protonation rate and dissociation of the enolate from the binding pocket before protonation can take place. This would provide an explanation for the non-stereospecific protonation of the Z-ClPEP-enolate by water after its release into the aqueous medium.

**Thermostability and salt bridges**

Electrostatic interactions are thought to act as an important factor conferring thermostability to proteins [58, 59]. This opinion is supported by the increased number of salt bridges found in many structures of thermostable proteins. On the basis of the known structure of EIC from *T. tengcongensis* the structures of EIC from *E. coli* and *M. capricolum* were modelled. The thermostable EIC from *T. tengcongensis* exhibits 17 salt bridges. Eight charged residues are forming an ion-pair network (Figure 6). In EIC of *E. coli* and *M. capricolum* are 12 and 9 salt bridges present. In both structures, six charged residues are forming an ion-pair network. The higher denaturation temperature of EIC from *T. tengcongensis* contrary to EIC from the mesophilic organisms *E. coli* and *M. capricolum* could be a result of the increased number of salt bridges and the larger ion-pair network on the surface.

**Discussion**

Enzyme I of *E. coli*, the key enzyme of the PTS, has been the object of numerous physiological, biochemical and biophysical studies. The structure of the HPr-binding N-terminal domain was solved by X-ray [13] and multidimensional NMR [14] eight years ago.
The structure of the full-length protein, however, has not been solved, possibly because the dynamic behaviour of EI and the marginal stability of the recombinant *E. coli* EIC disfavoured crystallisation, and because the structure of EIC could be predicted based on the sequence similarity with the PEP binding domain of PPDK. Hyperthermophilic bacteria such as *Thermotoga maritima*, *Thermus thermophilus* and archaeabacteria do not have a PTS. *T. tengcongensis*, a saccharolytic bacterium which optimally grows at 70°C turned out to have a PTS very similar to *E. coli*. EIC from *T. tengcongensis* is thermally stable to 90°C and more stable than the EIN domain, unlike in *E. coli* where EIN is significantly more stable than EIC. EIC of *E. coli* and *T. tengcongensis* have almost identical amino acid composition and pI. The contents of glycine, and basic amino acids are identical and only the Lys/Arg ratios are inversed. The sums of carboxylate and amide containing residues are identical but the carboxylate to amide ratio is shifted in favour of more Asn and less Gln in the thermostolerant EIC. The only Asn-Gly sequence, a linkage thought to be weak and hydrolysis-sensitive, is conserved in both proteins. EIC of *T. tengcongensis* contains 5 extra prolines in non-conserved sequences. There are thus no obvious alterations in the amino acid sequence and composition which would provide an explanation for the remarkably different thermostability. The increased number of surface exposed salt bridges in EIC of *T. tengcongensis* in comparison with EIC from *E. coli* and *M. capricolum* might one of the major reasons for the different thermostability.

The purified PTS proteins of *T. anaerobacter* displays maximum *in vitro* phosphotransferase activity at 70°C (Navdaeva and Erni, unpublished). Although this value may reflect the optimum of another PTS protein than EI, it is 20°C below the thermal unfolding temperature. This is more than the difference for *E. coli* EI (Tm 41°C to 60°C depending on the experimental conditions) and it raises the question whether thermostolerant EI can display the same complex dynamical behavior at 70°C that EI of *E. coli* displays at 37°C. EI plays the role of a gate-keeper by controlling the flux of phosphorylgroups into the PTS and it also
provides the interface between PTS and chemotaxis [60, 61]. Moreover EI can utilize not only PEP but also acetyl phosphate kinase as source of high energy phosphate [62]. The genome of T. tengcongensis contains the genes for all these proteins and therefore EI is likely to function similarly in T. tengcongensis and E. coli. But there are differences downstream along the protein phosphorylation cascade. T. tengcongensis contains no crr gene for the EIIA\textsubscript{Glc} subunit which in E. coli mediates inducer exclusion and activation of adenylcyclase, and also no gene for an adenylcyclase. On the other hand it does contain two crp genes for catabolite gene activators (cAMP receptor proteins) which are controlled by cAMP. How Crp is activated in the absence of an adenylcyclase is not clear but it points to differences downstream in the PTS system which might be compensated by differences upstream. Control and catalysis of EI in T. tengcongensis may be less complex, more robust and ultimately better adjusted to the environment. The contact areas in the EIC dimer of T. tengcongensis and in the modelled dimeric structures of E. coli and M. capricolum are comparable with respect to both, area and distribution of apolar, polar and charged residues (Figure 7) and thus also do not offer an explanation for the different stabilities.

**Footnotes**

* This research was supported by Grant 3100-063420 from the SWISS National Science Foundation and the Ciba-Geigy Jubiläumsstiftung. A Roche Research Foundation fellowship to A.E.O. is gratefully acknowledged. We thank Martin Walsh from BM14 at the ESRF in Grenoble for his invaluable help with data collection.
References


Figures and figure legends

Figure 1:

A: Cartoon representation of the monomer structure of EIC: α-helices are coloured in blue, β-sheets are coloured in red and loops are coloured in grey. Residues 302 to 308 and 337 to 351 are coloured in green.

B: Surface representation of the monomer of EIC. Residues 296 to 309 are coloured in red, residues 333-365 are coloured in blue and residues 452-479 are coloured in red.

C: Cartoon representation of the dimmer of EIC. Residues 296 to 309 are coloured in red, residues 333-365 are coloured in blue and residues 452-479 are coloured in red. The surface representation is superposed in semi-transparent.
Figure 2: Stereoview of a superimposition of EIC (red) and the PEP/pyruvate binding site of PPDK (blue).
Figure 3: Structure-based sequence alignment of EIC and PPDK. Conserved residues involved in PEP/pyruvate and Mg$^{2+}$ binding are coloured in blue. The coloured bar below the sequence shows the dimer interface. A red bar symbolises residues only involved in the EIC dimer interface, a blue bar residues of the PPDK dimer interface and a green bar residues involved in both dimer interfaces.
Figure 4: Stereo close up view of a superimposition of the putative active site of EIC with the PEP/pyruvate binding site of PPDK containing P-pyr and Mg$^{2+}$. Residues of EIC are colored in red, residues of PPDK are colored in blue. The numbers in parentheses correspond with the PPDK residues. P-pyr is shown as a stick model. Atomic colours are as follows: oxygen red, carbon white, phosphorus yellow and Mg$^{2+}$ magenta.
Figure 5:

A: Stereoview of the active site of EIC with modelled ECIPEP. ECIPEP and the side chains of Arg296, Arg332, Glu431, Asn454 and Cys502 are shown as a stick model. Atomic colors are as follows: oxygen, red; carbon, white; phosphorus and sulphur, yellow.

B: Stereoview of the active site of EIC with modelled ZCIPEP. ZCIPEP and the side chains of Arg296, Arg332, Glu431, Asn454 and Cys502 are shown as a stick model. Atomic colors are as follows: oxygen is shown in red; carbon in white, phosphorus and sulphur yellow.
Figure 6: Representation of the 8-member ion-pair network on the surface of EIC from *T. tengcongensis*. Basic residues are coloured in blue, acidic residues are coloured in red.

Figure 7: Surface comparison of the dimer interface of EIC of *T. tengcongensis* (left), EIC of *E. coli* (middle) and *M. capricolum* (right). Hydrophobic residues (Ala, Val, Cys, Leu, Ile, Phe, Tyr, Trp, Pro, Met) are coloured in yellow, basic residues (Arg, Lys, His) are coloured in blue, acidic amino acids (Glu and Asp) are coloured in red, polar residues (Asn, Gln, Ser, Thr, Gly) are shown in cyan.
Table 1: Data collection and refinement statistics

<table>
<thead>
<tr>
<th>Low Remote</th>
<th>Peak</th>
<th>Inflection</th>
<th>Remote</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crystal parameters</strong></td>
<td>P2₁,2₁,2₁; a = 82.10 Å, b = 91.43 Å, c = 181.86 Å</td>
<td>P2₁,2₁,2₁; a = 82.82 Å, b = 91.85 Å, c = 185.96 Å</td>
<td>four molecules/ a.s.u.</td>
</tr>
<tr>
<td><strong>Data collection (XDS)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beamline</td>
<td>BM14</td>
<td>BM14</td>
<td>BM14</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>0.9878</td>
<td>0.9788</td>
<td>0.9791</td>
</tr>
<tr>
<td>Resolution range</td>
<td>40-1.82 (1.94-1.82)</td>
<td>40-2.8 (2.97-2.80)</td>
<td>40-2.8 (2.97-2.80)</td>
</tr>
<tr>
<td>No. observations</td>
<td>820735</td>
<td>275471</td>
<td>277294</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>123055</td>
<td>66731</td>
<td>67336</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.9 (99.7)</td>
<td>98.8 (92.9)</td>
<td>99.7 (98.4)</td>
</tr>
<tr>
<td>R&lt;sub&gt;s&lt;/sub&gt; (%)</td>
<td>6.7 (24.4)</td>
<td>6.1 (18.4)</td>
<td>6.3 (20.3)</td>
</tr>
<tr>
<td>I/σ&lt;sub&gt;I&lt;/sub&gt;</td>
<td>16.5 (4.6)</td>
<td>13.8 (5.1)</td>
<td>13.5 (4.8)</td>
</tr>
<tr>
<td>FOM&lt;sup&gt;e&lt;/sup&gt; MAD (overall)</td>
<td>0.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FOM&lt;sup&gt;f&lt;/sup&gt; solvent flattened</td>
<td>0.77</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Refinement (REFMAC)**

| Resolution range (Å) | 30-1.82 | | |
| No. reflections working set | 116738 | | |
| No. reflections test set | 2245 | | |
| No. non hydrogen atoms | 10958 | | |
| Solvent water molecules | 837 | | |
| RMSD bond length (Å) | 0.015 | | |
| RMSD bond angles (deg.) | 1.870 | | |

*Asymmetric unit

<sup>a</sup> Friedel pairs were treated as different reflections.

<sup>b</sup> The values in parentheses of resolution range, completeness, R<sub>s</sub>, and I/σ(I) correspond to the outermost resolution shell

<sup>c</sup> R<sub>s</sub> = Σ<sub>hkl</sub>Σ<sub>j</sub>|I(hkl;j)| - (I(hkl);/Σ<sub>hkl</sub>Σ<sub>j</sub>)(I(hkl)) where I(hkl;j) is the jth measurement of the intensity of the unique reflection (hkl) and (I(hkl)) is the mean over all symmetry-related measurements.

<sup>d</sup> Figure-of-merit as computed by SOLVE

<sup>e</sup> Figure-of-merit as computed by RESOLVE
Danksagung

Bei Prof. Ulrich Baumann möchte ich mich für die hervorragende Betreuung während meiner Doktorarbeit danken.

Mein besonderer Dank gilt auch Prof. Bernhard Erni für die Möglichkeit, am Enzym I-Projekt mitzuarbeiten.

Mario Bumann, Siamak Djafarzadeh und Philipp Schneider danke ich für die hervorragende und fruchtbare Zusammenarbeit.

Herzlich möchte ich mich bei Mario Bumann, Santina Russo, Reto Meier, Dr. Jörg Benz, Dr. Stefan Hörer, Dr. Christian Siebold, Dr. Achim Stocker und Michael Locher bedanken, die dafür sorgten, dass das Labor nicht nur eine Arbeitstätte war.

Dr. Johann Schaller, Urs Kämpfer, Adrian Schindler und Michael Locher danke ich für den hervorragenden Proteinanalytik-Service.

Nicht zuletzt möchte ich meinen Eltern danken, die mir mit ihrer Unterstützung mein Studium erst ermöglicht haben.
Curriculum Vitae

Anselm Erich Oberholzer

Geburtsdatum: 23. Februar 1975
Heimatorte: Goldingen/SG und St. Gallen-Tablat/SG

Ausbildung

1982-1988 Primarschule St. Georgen
1988-1990 Sekundarschule KKSS St. Gallen
1990-1995 Kantonsschule St. Gallen
Maturität Typus C
1995 Rekrutenschule
1995-2000 Studium der Biologie biochemischer Richtung an der Universität Bern
Leiter der Arbeit: Prof. Dr. Ulrich Baumann
02/2000 Abschluss als diplomierter Naturwissenschaftler
Leiter der Arbeit: Prof. Dr. Ulrich Baumann