Tryptophan Synthase Complexes:
Evolution, Substrate Specificity, and Quaternary Structure

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFTEN (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von
Florian Busch
aus
Kösching

im Jahr 2015

Das Promotionsgesuch wurde eingereicht am:

29.5.2015

Die Arbeit wurde angeleitet von:

Prof. Dr. Reinhard Sterner

Unterschrift:

This work was done in the period from November 2010 to May 2015 in the group of Prof. Dr. Reinhard Sterner (Biochemistry II, Institute of Biophysics and Physical Biochemistry, University of Regensburg).

‘It is apparent that our knowledge of the structure and function of this enzyme [the tryptophan synthase complex] is providing a glimpse of the sophistication that is feasible in natural protein design.’

Edith Miles

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

FORMULA INDEX

LIST OF COMMON USED ABBREVIATIONS

Enzymes

GAPDH	glyceraldehyde-3-phosphate dehydrogenase
TrpA1	α-subunit of the permanent tryptophan synthase complex
TrpA2	α-subunit homolog
TrpB1	β-subunit of the permanent tryptophan synthase complex
TrpB2	β-subunit homolog
TS	tryptophan synthase complex/ complexes

Organisms

A. thaliana	*Arabidopsis thaliana*
E. coli	*Escherichia coli*
S. solfataricus	*Sulfolobus solfataricus*
S. typhimurium	*Salmonella typhimurium*
T. maritima	*Thermotoga maritima*
T. volcanium	*Thermoplasma volcanium*

Chemicals

EPPS	3-[4-(2-Hydroxyethyl)-1-piperazinyl]propanesulfonic acid
GP	rac glycerol-3-phosphate
GAP	glyceraldehyde-3-phosphate
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IGP	indole-3-glycerol-phosphate
LB	Luria broth media
OPS	O-phospho-L-serine
PLP	pyridoxal-5-phosphate

Techniques

CD	circular dichroism
DSC	differential scanning calorimetry
HPLC	high pressure liquid chromatography
ITC	isothermal titration calorimetry
SPR	surface plasmon resonance

Amino acids are abbreviated by the 1- or 3-letter code.
Nucleotides are abbreviated by the 1-letter code.

PREFACE

The bacterial tryptophan synthase complex (TS) serves as a longstanding model for the investigation of substrate channeling and allosteric communication within elaborate enzyme complexes. It is a permanent αββα tetramer with the active sites of the α- and β-subunit being connected by a hydrophobic channel. The α-subunit catalyzes the aldolytic cleavage of indole-3-glycerol-phosphate (IGP) to glyceraldehyde-3-phosphate (GAP) and indole. The latter is channeled to the active site of the β-subunit, where it reacts with L-serine to L-tryptophan and water. Bi-directional allosteric communication between the subunits results in the precise temporal coordination of the catalytic steps and ultimately prevents the release of intermediary formed indole during the TS reaction.

The evolution of permanent TS is examined in **Chapter 1**. In order to identify early adaptations in TS evolution, the sequences of the α-subunit and β-subunit of the last common ancestor (LCA) were reconstructed. The properties of the reconstructed subunits are described.

TrpB2 enzymes share a sequence identity of approximately 30 % with the β-subunits of the permanent TS. As TrpB2 enzymes have a low affinity for L-serine, they were proposed to utilize another substrate *in vivo*. The substrate specificity of TrpB2 enzymes is examined in **Chapter 2**.

TrpB2 enzymes encoded inside of the *trp* operon (TrpB2i) are the β-subunits of a transient, ligand-dependent TS. This complex has a different quaternary structure compared to the permanent TS. The properties of transient TS and the characteristics that make TrpB2 part of a transient TS are examined in **Chapter 3**.

Subsequent to the main part, a **Summary** of the presented data in English and German as well as the sections for **Materials**, **Methods**, **References**, and **Supporting Information** can be found. The work is concluded by the **Acknowledgements**.

1 EVOLUTION OF THE PERMANENT TS

1.1 Introduction

1.1.1 The heterotetrameric tryptophan synthase complex

Members of the InterPro enzyme families (Mitchell et al., 2015) IPR002028 and IPR006654 are the α- and β-subunits of a permanent heterotetrameric $(\alpha\beta)_2$ tryptophan synthase complex (TS). This enzyme complex serves as model system to understand the mechanisms underlying protein-protein interaction and allosteric regulation. The subunits, their assembly to the TS and the regulation of subunit activities will be discussed in this introduction, based on the available data for the enzymes from *Escherichia coli*, *Salmonella typhimurium*, *Thermotoga maritima* and *Pyrococcus furiosus*. If not stated otherwise, the assertions on permanent TS are assumed to apply to all known IPR002028-IPR006654 complexes.

1.1.2 Catalyzed reactions

The TS catalyzes the last two steps in tryptophan biosynthesis (Pan et al., 1997). The α-subunit catalyzes the aldolytic cleavage of indole-3-glycerol-phosphate (IGP) to indole and glyceraldehyde-3-phosphate (GAP) and the β-subunit catalyzes the condensation reaction of indole and L-serine to L-tryptophan and water (Figure 1).

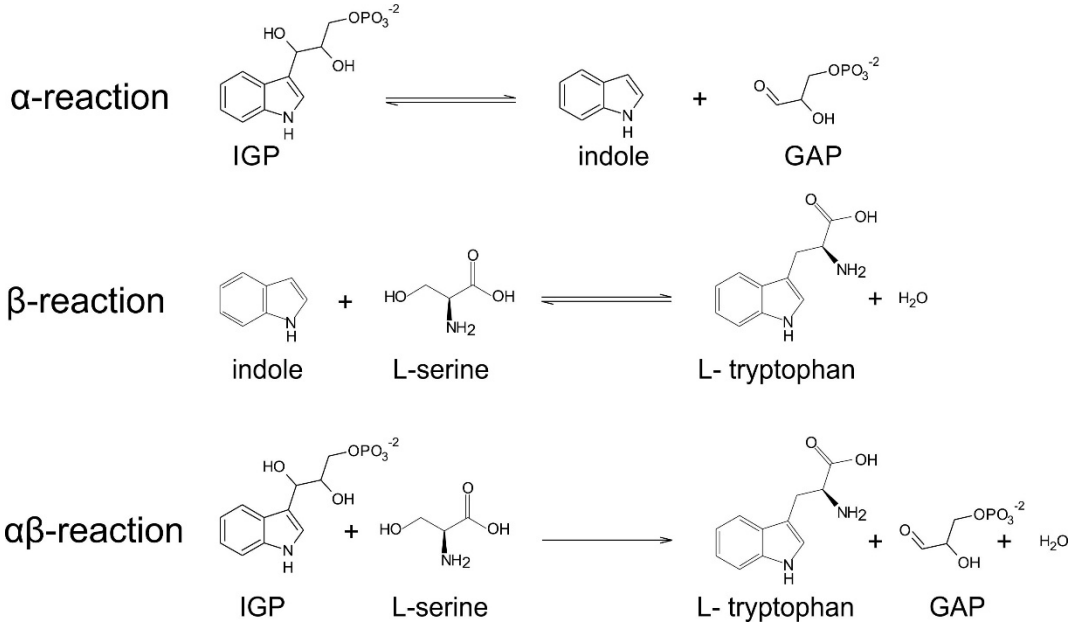


Figure 1: Reactions catalyzed by the α- and β-subunits.

1.1.3 The reaction at the α-subunit

The α-subunit is a general acid-base catalyst. The aldolytic cleavage reaction proceeds through an indolenine tautomer intermediate (Figure 2).

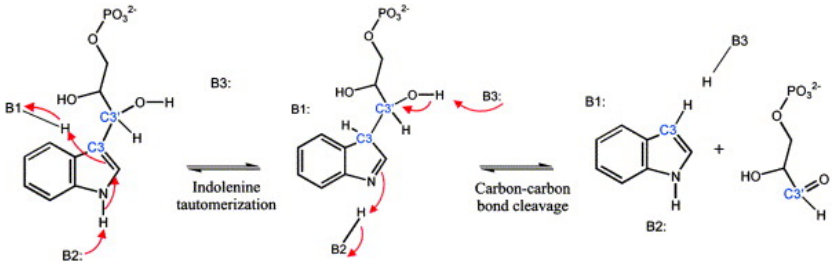


Figure 2: Enzymatic mechanism of the α-subunit.
The C3-C3' bond becomes aldolytically cleaved. In the α-subunit of *S. typhimurium*, Asp60 acts as B2 and Glu49 acts as B1H and B3. Figure from (Kulik et al., 2005).

The chemical balance lies on the side of the back reaction of indole and GAP to IGP (Weischet and Kirschner, 1976).

1.1.4 The reaction at the β-subunit

The β-subunit is a PLP dependent enzyme. The cofactor acts as electron sink in the course of the reaction (Barends et al., 2008). Main chemical steps of the reaction cycle are shown in Figure 3 (Schiaretti et al., 2004).

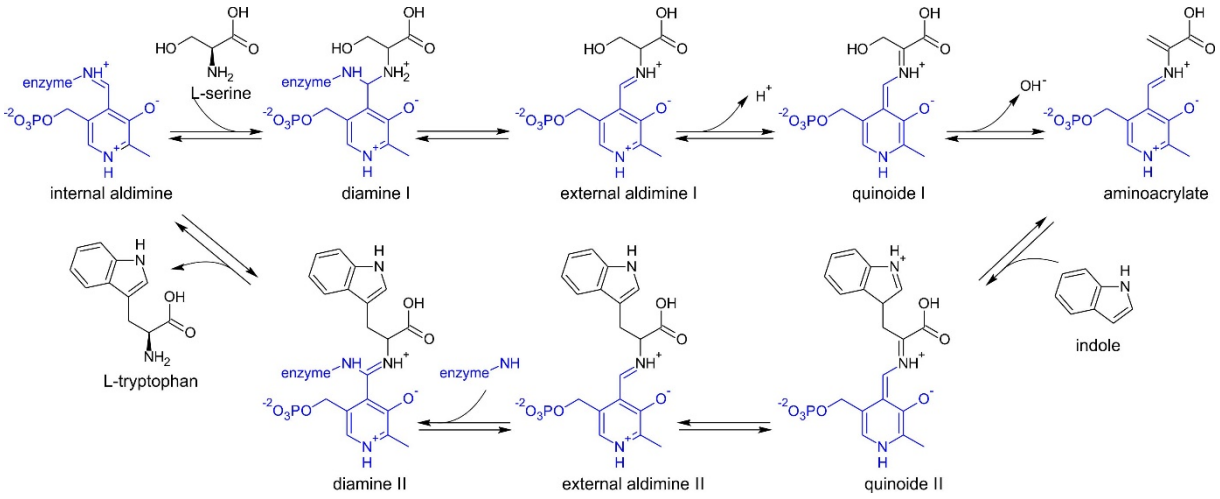


Figure 3: Enzymatic mechanism of the β-subunit.
The cofactor PLP and the lysine of the enzyme (enzyme-NH_2) are blue-colored. Substrates, intermediates, and products of the reaction are black-colored.

Initially, the cofactor PLP is covalently bound as internal aldimine by an enzyme's lysine. L-serine reacts with this internal aldimine to form diamine I. Subsequently, external aldimine I is formed. Dehydratization proceeds via quinoide I and leads to

the reactive aminoacrylate. A nucleophilic attack of the aminoacrylate by indole results in quinoide II, which is further converted to external aldimine II. L-tryptophan is subsequently released by its replacement with the enzyme's lysine via intermediate diamine II. All chemical steps are reversible. The equilibrium between external aldimine and aminoacrylate is influenced by many factors including monovalent cations, complex formation, ligands bound at the associated α-subunit, pressure, co-solvents, chaotropic reagents and temperature (Ahmed et al., 1996; Ahmed and Miles, 1994; Fan et al., 1999; Fan et al., 2000; Hur et al., 2002; Miles, 2001; Schiaretti et al., 2004).

1.1.5 Subunits and subunit assembly

Isolated monomeric α-subunits and dimeric β-subunits are stable and active enzymes. The subunits readily assemble to a permanent heterotetrameric complex with αββα stoichiometry (Creighton and Yanofsky, 1966; Goldberg et al., 1966). The catalytic efficiencies of the subunits are enhanced by complex formation (Lane et al., 1984). Enzymatic parameters of the isolated α- and β-subunits in comparison to the enzymatic parameters of the α- and β-subunits in the TS are shown for the enzymes from *E. coli* and *T. maritima* (Table 1).

Table 1: Effect of complex formation on the α- and β-reactions.

		α-reaction		β-reaction			
E. coli	k_{cat}	0.034 s^{-1} 0.0016 s^{-1}	21 x ↑	2.7 s^{-1} 0.09 s^{-1}	30 x ↑		
	K_m	0.14 mM 0.48 mM	3 x ↓(IGP)	0.43 mM 0.95 mM	2 x ↓(L-serine)	0.015 mM 0.014 mM	+/- (indole)
T. maritima	k_{cat}	2.8 s^{-1} 0.174 s^{-1}	16 x ↑	4 s^{-1} 1.9 s^{-1}	2 x ↑		
	K_m	0.19 mM 1.62 mM	9 x ↓(IGP)	3.7 mM 110 mM	30 x ↓(L-serine)	0.025 mM 0.04 mM	2 x ↓(indole)

Enzymatic parameters of the α-reaction (IGP-cleavage) and of the β-reaction (condensation of indole with L-serine) were determined for the isolated subunits and for the subunits in the *E. coli* TS at 25° C and in the *T. maritima* TS at 80° C. Parameters for the isolated subunits are blue-colored and parameters for the subunits in the TS are black-colored. The changes in k_{cat} and K_m resulting from complex formation are indicated as fold increase (↑) and fold decrease (↓). Changes that increase the catalytic efficiency (k_{cat}/K_m) are green-colored. K_m- and k_{cat}-values were deduced from (Hettwer and Sterner, 2002).

The effect of complex formation on the enzymatic parameters differs for the enzymes from *E. coli* and *T. maritima*. The catalytic efficiencies of the β-subunits are equally increased by complex formation, but by different means.

All TS share the structures of their subunits and the quaternary structure (Figure 4).

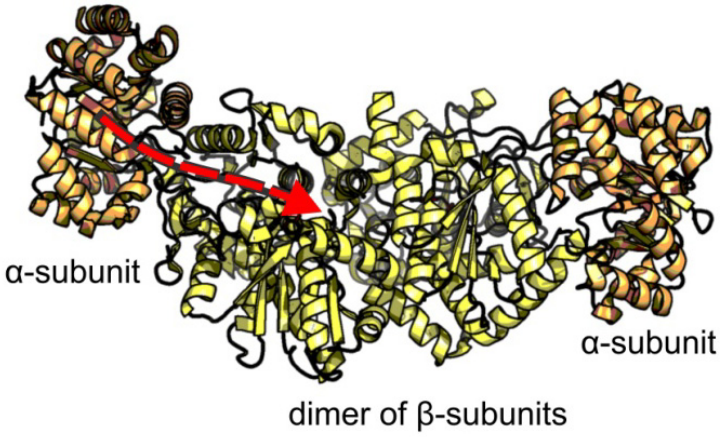


Figure 4: Subunit assembly and quaternary structure.
Cartoon representation of the TS quaternary structure from *P. furiosus* (1WDW). The α-subunits are orange-colored and the β-subunits are yellow-colored. The red arrow indicates the pathway of indole within the TS from the active site of the α-subunit to the active site of the β-subunit.

The α-subunit has a $(\beta\alpha)_8$-barrel topology and the β-subunit has a PLP dependent fold type II topology. Within the TS, subunits are assembled in a nearly linear αββα-arrangement and the active αβ-sites are connected by a hydrophobic tunnel (Hyde et al., 1988). The hydrophobic tunnel enables the transfer of indole from the active site of the α-subunit to the active site of the β-subunit (Pan et al., 1997).

1.1.6 Allostery within the TS

Indole is the product of the α-reaction and the substrate of the β-reaction. Such being the case, the TS catalyzes the reaction of IGP and L-serine to L-tryptophan, G3P and water (Figure 1). As a consequence of the coupled αβ-reaction, the α-reaction becomes largely irreversible because the reaction product indole is removed by its reaction with L-serine, which shifts the chemical equilibrium towards formation of indole. The catalytic efficiency of the α-subunit is increased and the activity of the β-subunit is decreased by allosteric communication within the TS (Table 2).

Table 2: Effect of allosteric communication on the α- and β-reactions.

		α-reaction		β-reaction	
E. coli	k_{cat}	0.7 s^{-1} 0.034 s^{-1}	21 x ↑	0.7 s^{-1} 2.7 s^{-1}	4 x ↓
	K_m	0.069 mM 0.14 mM	2 x ↓ (IGP)	0.34 mM 0.43 mM	+/- (L-serine)
T. maritima	k_{cat}	2.07 s^{-1} 2.8 s^{-1}	+/-	2.13 s^{-1} 4 s^{-1}	2 x ↓
	K_m	0.016 mM 0.19 mM	12 x ↓ (IGP)	0.13 mM 3.7 mM	28 x ↓ (L-serine)

Enzymatic parameters of the α-reaction (IGP-cleavage), of the β-reaction (condensation of indole with L-serine) and of the αβ-reaction (conversion of IGP and L-serine to tryptophan) were determined for the *E. coli* TS at 25° C and for the *T. maritima* TS at 80° C. Parameters of the α- and β-reactions are blue-colored, parameters of the αβ-reaction are black-colored. The changes in k_{cat} and K_m are indicated as fold increase (↑) and fold decrease (↓). Changes that increase the catalytic efficiency (k_{cat}/K_m) are green-colored. Changes that decrease the catalytic efficiency (k_{cat}/K_m) are red-colored. K_m- and k_{cat}-values were deduced from (Hettwer and Sterner, 2002).

The effect of allosteric communication differs considerably in the *E. coli* TS and in the *T. maritima* TS. Allosteric mechanisms, which lead to an increase in the activity of the α-reaction seem to have evolved in an ancestor of Enterobacteria. Whereas no detailed study on the allosteric communication in the *E. coli* TS is available, the reactions have been studied for the TS from the closely related Enterobacterium *S. typhimurium* (Figure 5).

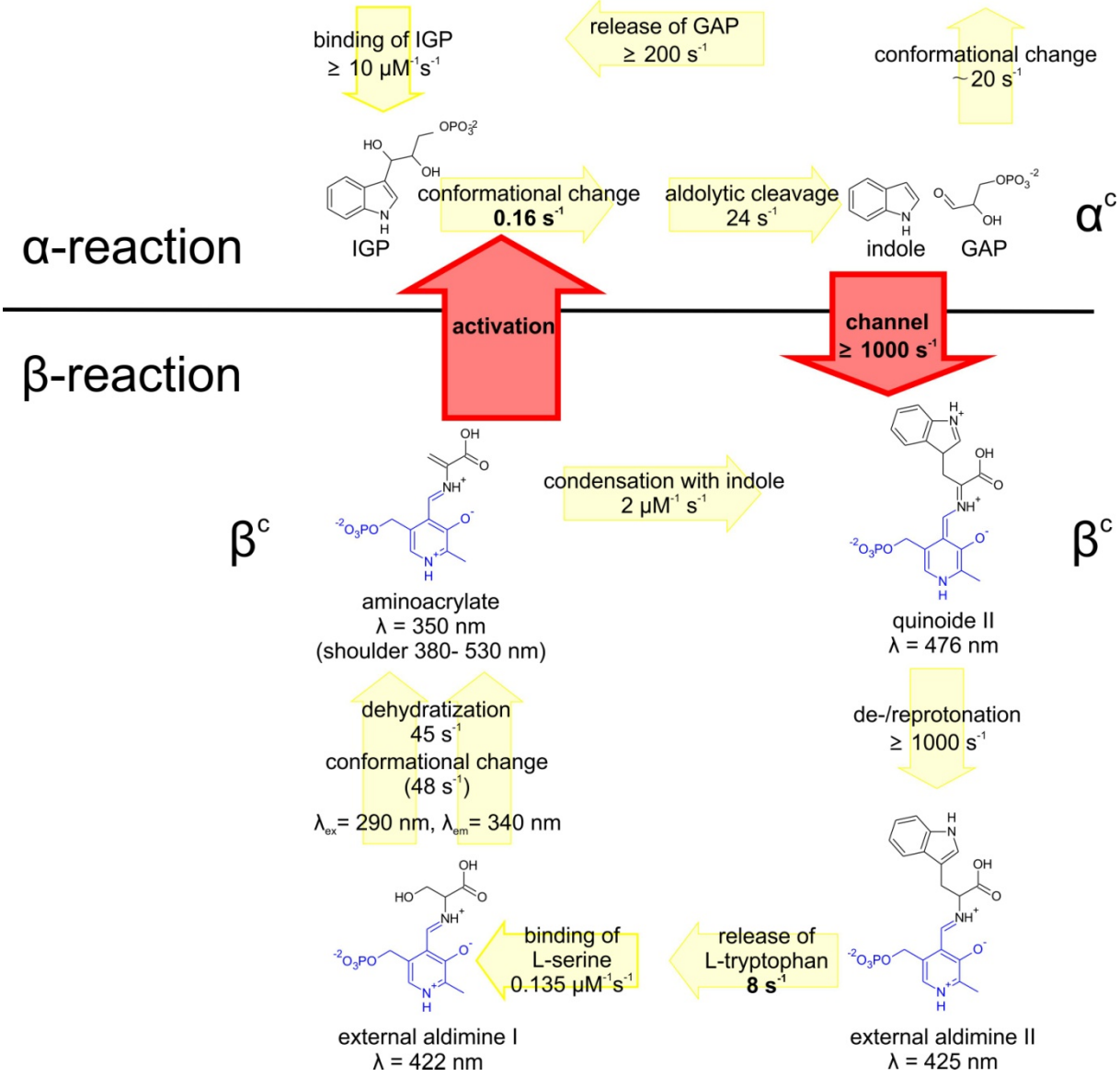


Figure 5: The coupled αβ-reaction in the *S. typhimurium* TS.
Rate constants for the enzymatic steps within the *S. typhimurium* TS in the sodium-bound state were determined at 37° C and pH 8.7 by chemical quench-flow and stopped-flow methods (Anderson et al., 1991). Intermediates of the β-reaction are spectroscopically detectable at the indicated wavelength(s) (Anderson et al., 1991; Schiaretti et al., 2004). Closed conformations of the subunits are indicated as α^c and β^c (Niks et al., 2013).

Within the *S. typhimurium* TS, the formation of the aminoacrylate is accompanied by a conformational change of the β-subunit. This speeds up IGP-cleavage by 150 fold as the α-subunit is turned into the active conformation. Within the working TS, indole is rapidly channeled to the active site of the β-subunit. The passage of indole is tightly controlled by the channel-lining residue Phe280 (Janda et al., 2014; Ruvinov et al., 1995). A pre-steady-state burst of L-tryptophan formation at a rate of 33 s^{-1} indicates that the steady-state turnover rate in the working TS is limited by the slow release of L-tryptophan at a rate of 8 s^{-1} under the investigated conditions (Anderson et al., 1991).

During catalysis, both subunits within the TS switch between an open conformation with low catalytic activity and a closed conformation with high catalytic activity (Leja et al., 1995; Miles, 2013). The temporal coordination of conformational changes regulates the formation of indole and the subsequent reaction to tryptophan during the TS reaction. The influence of ligand binding and enzymatic reactions on the subunit conformation and activity within the *S. typhimurium* TS is shown in Figure 6.

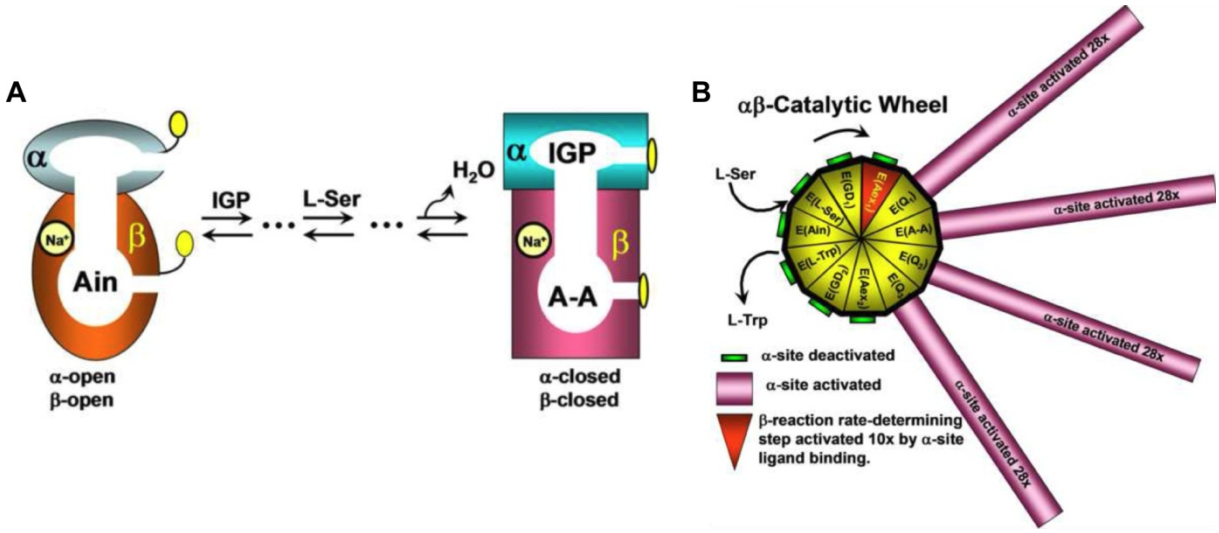


Figure 6: Regulation of subunit activities within the *S. typhimurium* TS.
(**A**) Binding of IGP and aminoacrylate formation (A-A) results in the TS with the subunits in the closed conformation. The access from the exterior to the active sites is restricted in the TS with the subunits in the closed conformation as indicated by the yellow circles. (**B**) Conformational changes trigger the bi-directional activation of the subunits during the catalytic cycle. In the case of the *S. typhimurium* TS, aminoacrylate formation at the β-subunit results in a 28-fold activation of the α-subunit and *vice versa* results IGP binding at the α-subunit in a 10-fold activation of the β-subunit. Figures are from (Dunn, 2012).

The *S. typhimurium* TS has been crystallized with both subunits in the open conformation as well as with both subunits in the closed conformation. For instance, both subunits are in the open conformation in the structure of the wild type TS and both subunits are in the closed conformation in the βK87T TS with bound GP and L-serine. A comparison of both structures is shown in Figure 7.

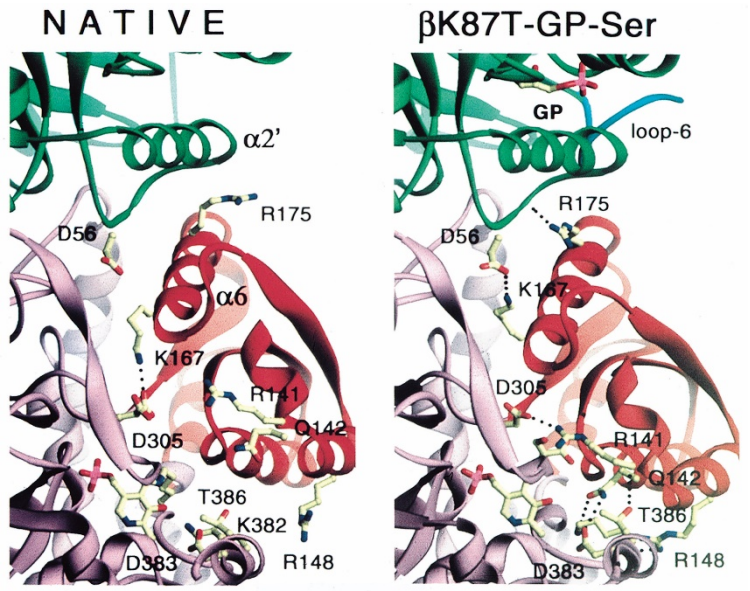


Figure 7: Conformational changes within the *S. typhimurium* TS.
The structure of the wild type *S. typhimurium* TS (1BKS) and of the βK87T *S. typhimurium* TS with bound GP and L-serine (2TSY) are depicted as ribbon diagram. The α-subunits are green-colored and the β-subunits are lavender-colored. The moveable communication domain (COMM domain) of the β-subunit is highlighted by red color. Hydrogen bonds are indicated by dashed lines and hydrogen bonding residues are shown as sticks. Figures are from (Miles, 2001).

Different hydrogen bonding networks exist in the open conformation TS and in the closed conformation TS. For instance, the βLys167-αAsp56 salt-bridge, which provides the major pathway for allosteric communication, is only formed in the closed conformation TS (Fan et al., 2000; Rowlett et al., 1998).

1.1.7 The TS as model system

The TS became a model system for the investigation of interaction and allosteric communication within a protein complex as it is amenable to a detailed experimental characterization. For instance, complex formation and allosteric communication within the complex can be studied by measuring the effect on catalysis. Furthermore, the reaction intermediates and conformational changes can be detected spectroscopically. The structure of the *S. typhimurium* TS was studied in the absence and in the presence of ligands as it can be expressed and purified in large quantities and is crystallizable. From an engineer's standpoint, the enterobacterial TS seems to be the 'perfectly' evolved extant TS.

1.1.8 Ancestral protein reconstruction

How has this fine-tuning in the extant TS evolved? Extant proteins and protein complexes are the result of an evolutionary process, in which vertically transmitted genetic trait is subtly changed from generation to generation. Comparing the characteristics of all members of an enzyme family reveals ‘shared’ conserved features, which were most likely also present in the common ancestor, and ‘exclusive’ features, which might or might not have been present in the common ancestor. In order to differentiate early and late adaptations in the evolution of an enzyme family, ancestral protein reconstruction is used. Recent developments in computational biology allow for the reliable reconstruction of protein ancestors, which have existed millions of years ago (Wilson et al., 2015; Yokoyama et al., 2014) or had even existed at the rise of the last common universal ancestor (LUCA) about 3.5 billion years ago (Perez-Jimenez et al., 2011; Reisinger et al., 2014). Those reconstructions help to understand the evolution of protein-protein interactions (Finnigan et al., 2012) and protein-ligand interactions (Kuang et al., 2006). They also help to detect shifts in thermodynamics and conformational flexibility, which had occurred in the course of protein evolution and specification (Hart et al., 2014; Wilson et al., 2015). Basis for the protein reconstruction is a multiple sequence alignment (MSA) of representative extant proteins, which is subsequently used to calculate a most likely phylogenetic tree. Nodes within this tree represent ancestors of the connected leaves and the root of the tree represents the universal ancestor. The distribution of phyla and the root are not mandatorily identical in the phylogenetic tree of a certain protein family and in the universal tree of life due to the individual history of protein families and horizontal gene transfer events.

1.2 Remarks

The TS ancestor (LCA TS) was calculated by Prof. Dr. Rainer Merkl. Stopped-flow experiments were done in collaboration with Dr. Sandra Schlee.

1.3 Significance of this work

The enterobacterial permanent TS is the result of a stringent evolutionary process yielding the efficient conversion of IGP to L-tryptophan without loss of intermediary formed indole. With tracing back its evolution one would be able to

identify ancient characteristics preceding the allosteric regulation in extant TS, to define the starting point in the development of modern TS and to gain information about the requirements of an ancient organism on this two-step catalysis. Recent proceedings in computational biology allow for the reconstruction of ancient enzymes and were successfully applied to resurrect an ancient TS. This work documents the high accuracy of *state of the art* computational enzyme reconstruction and provides first-time insights into the evolution of an allosteric regulated protein-protein interaction.

1.4 Results and discussion

1.4.1 Reconstruction of an ancient LCA TS

A phylogenetic tree was constructed using a multiple sequence alignment of 52 extant TS sequences. The sequences were derived from the bacterial phyla of Chloroflexi, Deinococci, Nitrospirae, Verrucomicrobia, Proteobacteria and Firmicutes, and the archaeal phyla of Euryarchaeota (Figure 8).

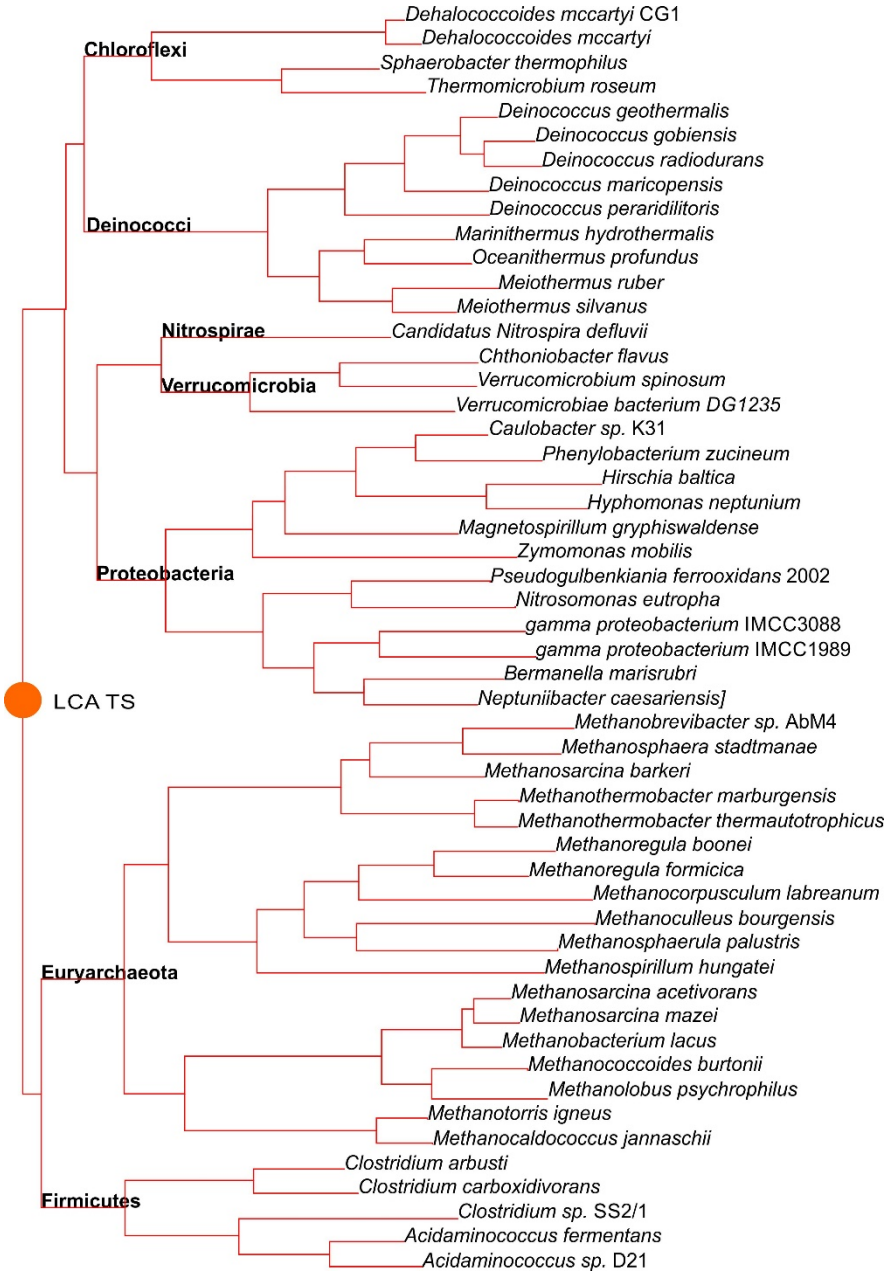


Figure 8: Phylogenetic tree for the reconstruction of an ancient TS.
The constructed tree was rooted by taking the phylogenetic relationship between extant TS into consideration. The location of LCA TS (**L**ast **C**ommon **A**ncestor of **T**ryptophan **S**ynthase Complexes) within the phylogenetic tree corresponds to the root of the constructed tree and is indicated by an orange-colored dot.

The evolution of permanent TS could be traced back to their first occurrence in a time before the diversification of Bacteria. Archaea have evolved an ‘alternative’ transient TS that is still conserved in Crenarchaeota (Ehrmann et al., 2010), but was replaced by the bacterial TS in Euryarchaeota (Merkl, 2007). According to the constructed phylogenic tree, this replacement in Euryarchaeota is the result of a horizontal gene transfer derived from an ancestor of Firmicutes. The hierarchy of bacterial phyla within the constructed TS tree is consistent with the global phylogenetic relationship of species (Ciccarelli et al., 2006), which is depicted in Figure 9.

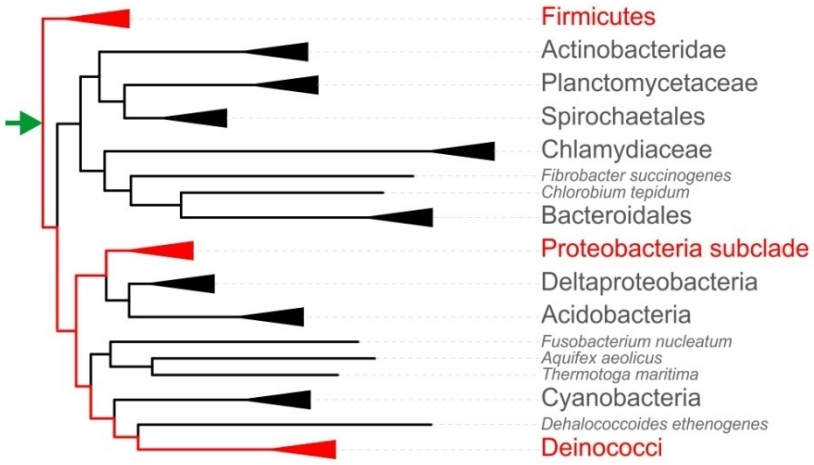


Figure 9: Phylogenetic tree of life.
A dataset of 150 sequenced bacterial species was used and 31 concatenated orthologs were considered for this phylogenetic reconstruction. The origin of Bacteria is indicated by a green-colored arrow. Phyla that had also been taken into account for the TS reconstruction are red-labeled (Letunic and Bork, 2007; Letunic and Bork, 2011).

1.4.2 Cloning of LCA subunits

The LCA TS sequence consists of the concatenated sequences of the LCA α-subunit and LCA β-subunit. The corresponding genes were optimized for protein expression in *E. coli* and synthesized by Life Technologies (9.1). Genes were cloned into pET21a(+) at the NdeI/ XhoI restriction sites to allow for the expression of proteins with a C-terminal His_6-*tag* (Figure 10).

LCA α-subunit

MNRIAEAFEELKKKGEKALIPFITAGDPDLETTLELVRALVEAGADIIELGIPFSDPLADGPTIQRASQRALASGTTLDKVFEMVREL
REKNTDVPIVFLTYYNPIFRYGIERFVKECAEAGVDGLIVPDLPPEEAADLAAAAEKYGVDLIFLVAPTSTDERIKMIAKHASGFVYC
VSVTGVTGARSEIAADLAELVSRIRKHTDLPIAVGFGISTPEQAAEVAQVADGVIVGSAIVKRIEENQDEEDIVEEVREFVRELREAV
K**LEHHHHHH**

LCA β-subunit

MIGRFGKYGGQYVPETLMPALEELEEAYERAKNDPEFQAELEYYLRDYVGRPTPLYFAENLTKDLGGAKIYLKREDLNHTGAHKI
NNALGQALLAKRMGKKRVIAETGAGQHGVATATVAAMFGLECVVYMGAEDIERQALNVFRMKLLGAKVRPVTSGSRTLKDAINE
AMRD**W**VTNVEDTFYIIGSVVGPHPYPMMVRDFQSVIGEEARQQILEKEGRLPDAIVACVGGGSNAMGIFHPFIDDESVRLIGVEA
AGKGIETGKHAATLSAGRPGVLHGAMTYLLQDEDGQIIEAHSISAGLDYPGVGPEHAYLKDTGRAEYVSVTDDEALEAFQLLSRT
EGIIPALESSHAVAYAMKLAPELSKDQIIVVNLSGRGDKDVNTVARYLLGVELD**LEHHHHHH**

Figure 10: Amino acid sequences of LCA subunits.
Amino acids derived from the expression vector and the His_6-*tag* are red-colored. The single tryptophan in the LCA β-subunit is green-colored.

The closest related extant enzymes are the α-subunit from *Clostridium arbusti* (57 % sequence identity, 100 % coverage) and the β-subunit from *Caldanaerobacter subterraneus* (78 % sequence identity, 99 % coverage) according to the NCBI database (status as of 01/12/2015). The low sequence identity between the LCA α-subunit and the closest related extant α-subunit reflects the fact that α-subunits are in general less conserved than β-subunits.

1.4.3 Expression and purification of LCA subunits

For expression of LCA subunits, *E. coli* T7 Express I^q was transformed with pET21a(+)-*LCA*-α and pET21a(+)-*LCA*-β, respectively. The cells were grown at 37° C in LB with 150 µg/ ml ampicillin to OD_{600}= 0.6. The media was supplemented with 20 µM PLP for the expression of the LCA β-subunit. Protein expression was induced by addition of 0.5 mM isopropyl-β-thiogalactopyranoside (IPTG). After growth over night at 20° C, cells were harvested by centrifugation, resuspended in 100 mM potassium phosphate pH 7.5, 300 mM KCl and 10 mM imidazole (with 20 µM PLP in the case of the LCA β-subunit) and disrupted by sonication. After a heat step (20 min, 70° C), the His_6-tagged proteins were purified by metal chelate affinity chromatography using a HisTrap FF crude column. Proteins in 100 mM potassium phosphate pH 7.5 and 300 mM KCl were eluted by a linear gradient of imidazole (10- 1000 mM) and dialyzed against 100 mM potassium phosphate pH 7.5. The LCA subunits were at least 95 % pure as judged by SDS-PAGE (Figure 11).

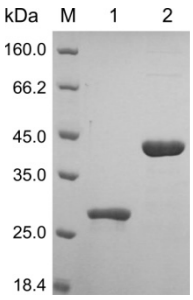


Figure 11: Purity of LCA subunits.
SDS-PAGE (12.5 % polyacrylamide) of the purified LCA subunits. Applied were (**M**) protein ladder and marker (LMW), (**1**) 15 µl LCA α-subunit (10 µM monomer concentration) and (**2**) 15 µl LCA β-subunit (10 µM monomer concentration).

Monomeric LCA α- and β-subunits have a molecular weight of 29.9 kDa and 43.6 kDa, respectively. This is consistent with their separation by SDS-PAGE. A solution of LCA β-subunit is yellow-colored, which confirms that the cofactor PLP is tightly bound.

1.4.4 Structural integrity, thermal stability, and activity of LCA subunits

The structural integrity of LCA subunits was analyzed by analytical size exclusion chromatography and far-UV circular dichroism (Figure 12).

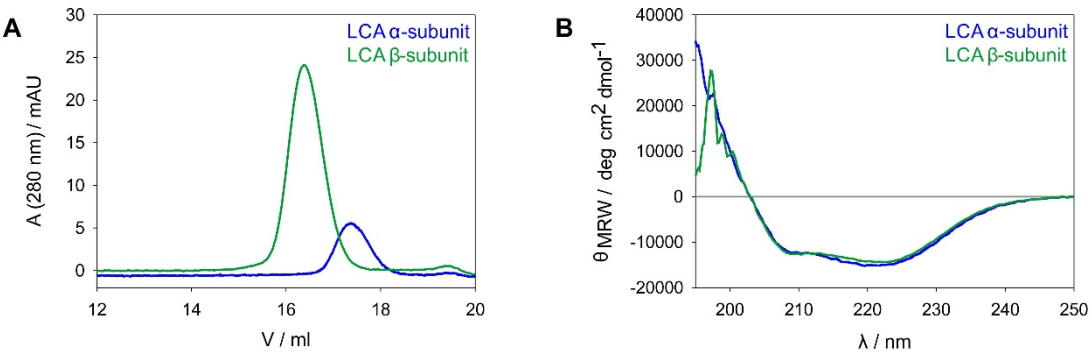


Figure 12: Structural integrity of LCA subunits.
(**A**) Analytical size exclusion chromatograms of 50 µl LCA α- and β-subunit (10 µM monomer concentration). The subunits were applied on a S200 analytical column equilibrated with 50 mM potassium phosphate pH 7.5, 300 mM KCl. Elution was performed with a flow rate of 0.5 ml/ min at 25° C. (**B**) Far-UV CD-spectra of 10 µM LCA α- and β-subunit (monomer concentration). Spectra were recorded in 50 mM potassium phosphate pH 7.5 from 195 nm to 250 nm (d= 1 mm).

The proteins are compactly folded and elute as homogenous symmetric peaks from the analytical size exclusion column. The molecular weights that correspond to the elution times were calculated with a calibration curve of standard globular proteins (9.2). According to the calculated molecular weights, the LCA subunits have the same oligomeric state as extant α- and β-subunits. As such, the LCA α-subunit is a monomer (calculated: 33.5 kDa; expected for α: 29.9 kDa) and the LCA β-subunit is

a dimer (calculated: 80.4 kDa; expected for β_2: 87.2 kDa) in solution. Both subunits have CD-spectra dominated by α-helical structures, which are virtually identical to the spectra of extant α- and β-subunits. The CD-spectrum of the β-subunit is influenced by the cofactor, which interferes with an accurate measurement of ellipticity below ~205 nm.

The thermal stability of the LCA subunits was investigated by thermal denaturation, which was followed by CD-spectroscopy and by DSC (Figure 13).

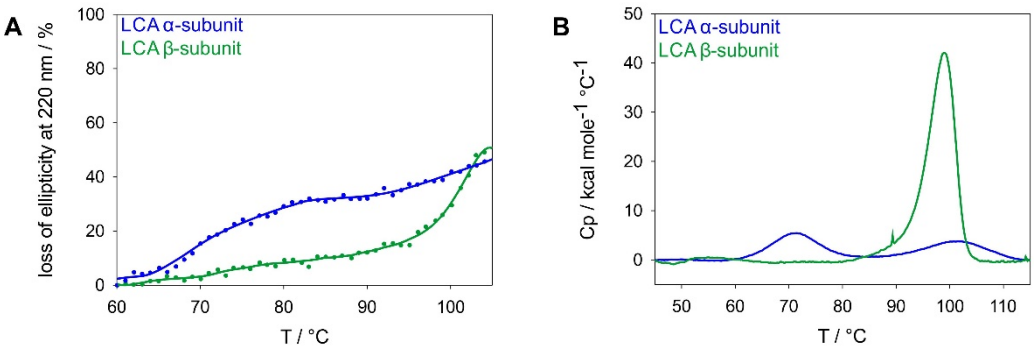


Figure 13: Thermal stability of LCA subunits.
(**A**) Thermal denaturation followed by CD-spectroscopy. The loss of ellipticity at 220 nm of 10 µM subunit (monomer concentration) in 50 mM potassium phosphate pH 7.5 was monitored from 60° C to 105° C at a scan rate of 1° C/ min (d= 1 mm). The curves connecting the data points were LOESS smoothed. (**B**) Thermal denaturation followed by DSC. Changes in heat capacity of 15 µM subunit (monomer concentration) in 50 mM potassium phosphate pH 7.5 were detected from 45° C to 115° C at a scan rate of 1° C/ min using 50 mM potassium phosphate pH 7.5 as reference. The curves were baseline corrected.

Thermal unfolding followed by CD-spectroscopy revealed that both LCA subunits retain approximately 50 % of ellipticity at 220 nm at a temperature of 105° C. The LCA α-subunit has - at least - two unfolding steps and the LCA β-subunit has one unfolding step within the operational temperature range. Thermal unfolding followed by DSC confirmed these findings and allowed to determine the apparent melting temperatures (T_m). The LCA α-subunit has T_m-values of 71° C and 102° C and the LCA β-subunit has a T_m-value of 99° C. Both subunits don't aggregate at 115° C as indicated by the symmetry of unfolding signals.

Due to their thermal stability and a presumable 'hot' environment in a time before the diversification of Bacteria, the enzymatic parameters of the LCA subunits were determined at the elevated temperature of 60° C (Figure 14).

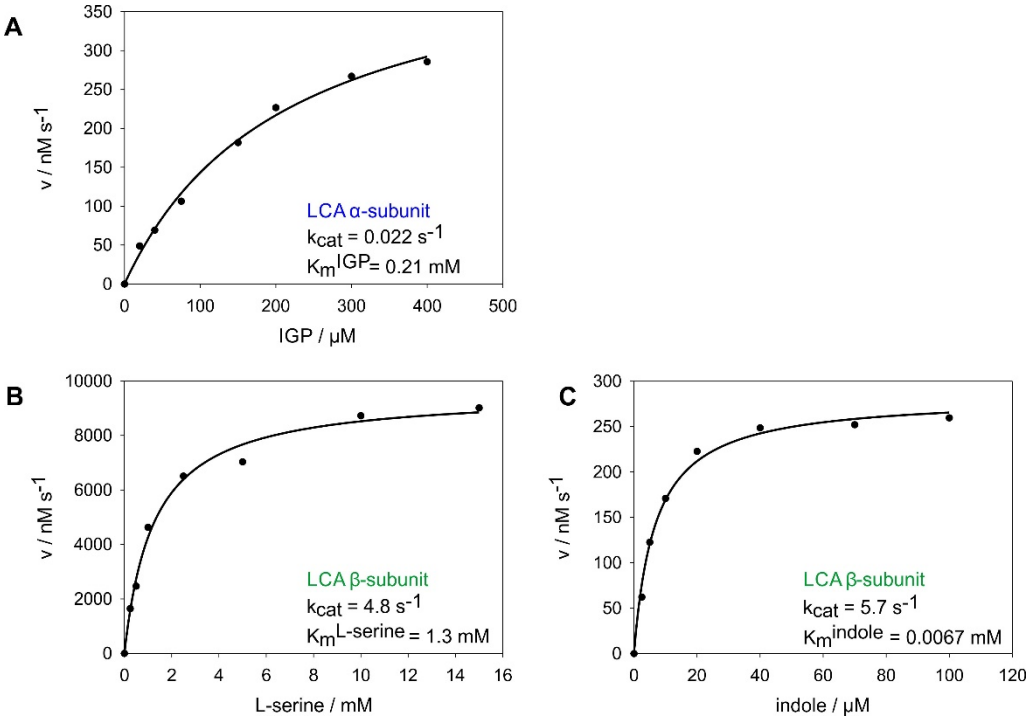


Figure 14: Enzymatic parameters of isolated LCA subunits.
(**A**) IGP-dependent LCA α-reaction. The reaction was started by the addition of IGP and followed at 60° C by the absorbance change at 340 nm caused by the production of NADH. The reaction contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 6 mM NAD^+, 20 mM arsenate, 5.5 µM GAPDH and 20 µM LCA α-subunit (monomer concentration). (**B**) L-serine-dependent LCA β-reaction. The reaction was started by the addition of 100 µM indole and followed at 60° C by the absorbance change at 290 nm. The reactions contained 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 2 µM LCA β-subunit (monomer concentration) and varying amounts of L-serine. (**C**) Indole-dependent LCA β-reaction. The reaction was started by the addition of indole and followed at 60° C by the absorbance change at 290 nm. The reactions contained 100 mM potassium phosphate pH 7.5, 180 mM KCl, and 40 µM PLP, 100 mM L-serine and 0.05 µM LCA β-subunit (monomer concentration). Data points were fitted with a single rectangular hyperbolic function (Michaelis-Menten equation).

Both subunits are catalytically active. As for extant subunits, the activity of the α-subunit is low (k_{cat}= 0.022 s^{-1}). The activity of the LCA β-subunit (k_{cat}= 4.8 s^{-1}) is significantly higher than the activity of an enterobacterial β-subunit (Table 1).

In summary, the thermal stability, oligomeric state and catalytic activity of the LCA subunits show that the reconstructed enzymes are thermostable and catalytically active.

1.4.5 Subunit assembly and activity of subunits within the LCA TS

Have the LCA enzymes already formed a complex in an early era of evolution before the diversification of Bacteria? To answer this question, a potential interaction between LCA α-subunit and LCA β-subunit was analyzed by analytic size exclusion chromatography and fluorescence titration (Figure 15).

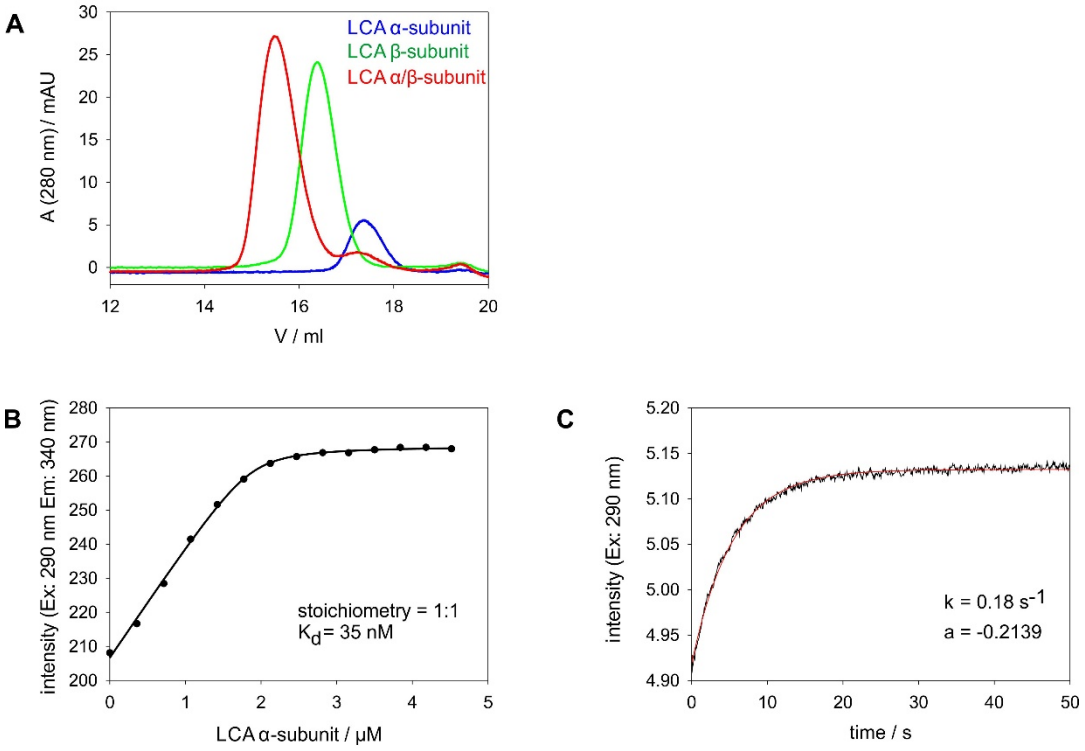


Figure 15: Assembly of subunits to a LCA complex.
(**A**) Analytical size exclusion chromatograms of 50 µl LCA α-subunit, β-subunit and αβ mixture (10 µM monomer concentration). The subunits were applied on a S200 analytical column equilibrated with 50 mM potassium phosphate pH 7.5, 300 mM KCl. Elution was performed with a flow rate of 0.5 ml/ min at 25° C. (**B**) Fluorescence titration of 2 µM LCA β-subunit (monomer concentration) with LCA α-subunit. The titration assay was performed in 10 mM potassium phosphate pH 7.5 at 25° C using a spectro-fluorimeter (d= 1 cm). Tryptophan was excited at 290 nm and the emission was detected at 340 nm. Data points were fitted with a quadratic function. (**C**) Time-dependent fluorescence change upon complex formation. 2 µM LCA β-subunit (syringe A) were mixed with 2 µM α-subunit (syringe B) in 10 mM potassium phosphate pH 7.5 in a stopped-flow instrument. Tryptophan was excited at 290 nm and the emission was detected with a 335 nm cut-off filter. Data was fitted with a single exponential function.

An interaction between the LCA α-subunit and the LCA β-subunit is detectable by analytical size exclusion chromatography. The calculated molecular weight of the eluting LCA tryptophan synthase complex (LCA TS) indicates that it has a $\alpha_2\beta_2$ stoichiometry (calculated: 176.3 kDa; expected for $\alpha_2\beta_2$: 147.0 kDa). Fluorescence titration confirmed the 1:1 stoichiometry of subunits in the LCA TS and showed that the K_d for the LCA TS is in the nanomolar range. The complex is slowly formed as indicated by stopped-flow analysis. This might be due to a rate-limiting conformational change, which precedes or accompanies complex formation. As complex formation might influence the enzymatic properties of the subunits, the enzymatic parameters of the α- and β-reactions within the LCA TS were subsequently determined (Figure 16).

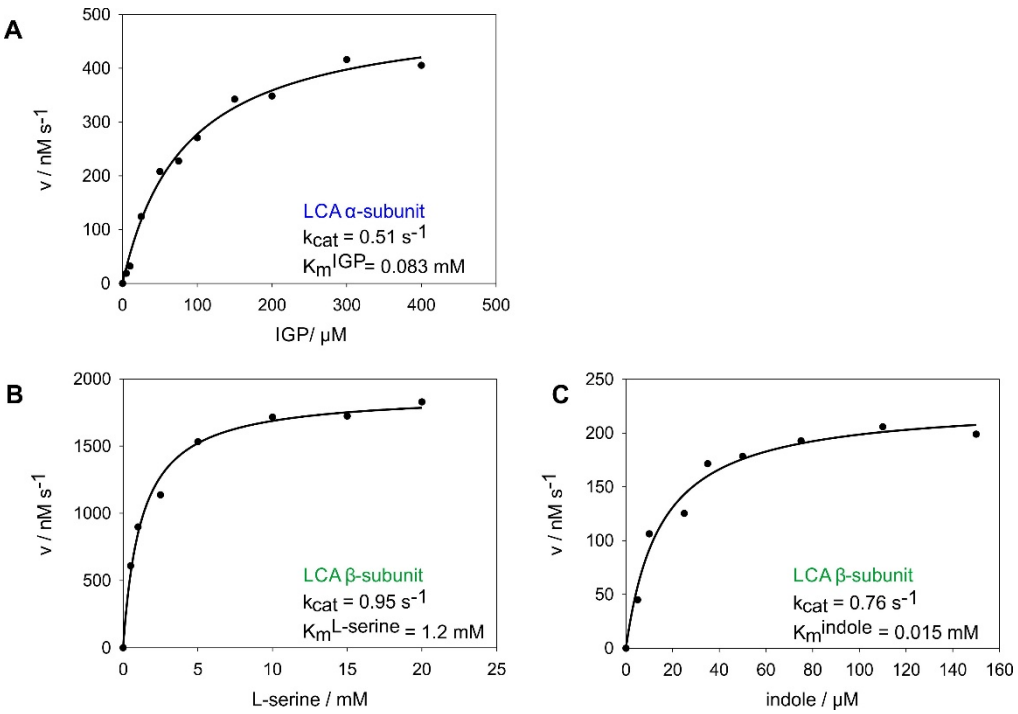


Figure 16: Enzymatic parameters of LCA subunits within the TS.
(**A**) IGP-dependent LCA α-reaction in the presence of β-subunit. The reaction was started by the addition of IGP and followed at 60° C by the absorbance change at 340 nm caused by the production of NADH. The reaction contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 6 mM NAD^+, 20 mM arsenate, 5.5 µM GAPDH, and 1 µM α- and β-subunits. (**B**) L-serine-dependent LCA β-reaction in the presence of α-subunit. The reaction was started by the addition of 100 µM indole and followed at 60° C by the absorbance change at 290 nm. The reaction contained 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 2 µM LCA α- and β-subunits and varying amounts of L-serine. (**C**) Indole-dependent LCA β-reaction in the presence of α-subunit. The reaction was started by the addition of indole and followed at 60° C by the absorbance change at 290 nm. The reactions contained 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 100 mM L-serine, and 0.3 µM LCA α- and β-subunits. Data points were fitted with a single rectangular hyperbolic function (Michaelis-Menten equation).

The catalytic parameters of both subunits are influenced by complex formation. An overview of changes induced by complex formation is given in Table 3.

Table 3: Effect of complex formation on the LCA α- and β-reactions.

		α-reaction		β-reaction			
LCA	k_{cat}	0.51 s^{-1} 0.022 s^{-1}	23 x ↑	0.86 s^{-1} 4.8 s^{-1}	6 x ↓		
	K_m	0.083 mM 0.21 mM	3 x ↓(IGP)	1.2 mM 1.3 mM	+/- (L-serine)	0.015 mM 0.0067 mM	2 x ↑(indole)

Enzymatic parameters of the α-reaction (IGP-cleavage) and of the β-reaction (condensation of indole with L-serine) were determined for the isolated subunits and for the subunits in the *LCA* TS at 60° C. Parameters for the isolated subunits are blue-colored and parameters for the subunits in the TS are black-colored. The relative change in k_{cat} and K_m resulting from complex formation is indicated as fold increase (↑) and fold decrease (↓). Changes that increase the catalytic efficiency (k_{cat}/K_m) are green-colored. Changes that decrease the catalytic efficiency (k_{cat}/K_m) are red-colored. K_m- and k_{cat}-values refer to the single active site.

The enzymatic parameters of the LCA α-subunit and extant α-subunits are similarly affected by complex formation. The k_{cat} is increased and the K_m^{IGP} is decreased upon complex formation.

Contrary to extant β-subunits (Table 1), complex formation with the LCA α-subunit leads to a reduction of the activity of the LCA β-subunit. However, the reaction of the α-subunit is slower than the reaction of the β-subunit within the LCA TS.

1.4.6 Impairment of the β-reaction within the LCA TS

The impairment of activity of the β-subunit within the LCA TS was not observed in those extant TS, which have been studied up to date. The reason underlying this impairment may be the restriction of substrate accessibility, a limitation in a catalytic step and/ or the restriction of product release brought about by complex formation. Those steps were thus analyzed in more detail by stopped-flow kinetics comparing the isolated β-subunit and the β-subunit complexed with the α-subunit (Figure 17 - Figure 20).

Binding of L-serine to the PLP-cofactor

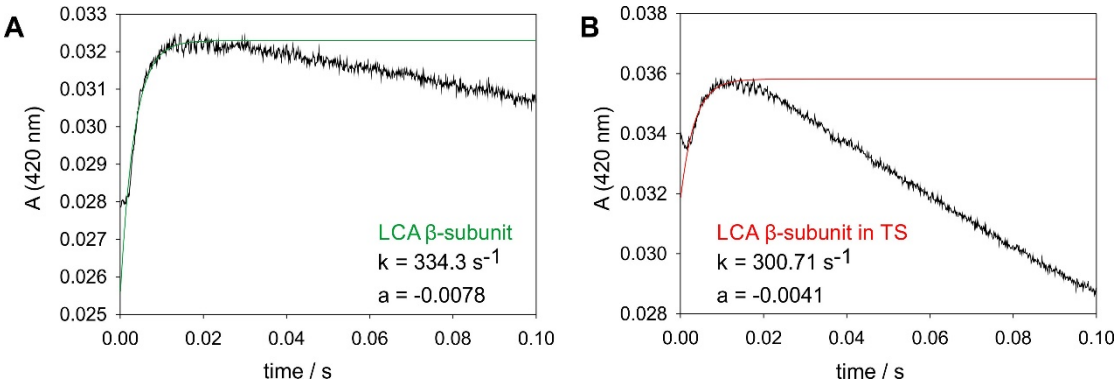


Figure 17: Binding of L-serine to the PLP-cofactor.
Influence of TS formation on the binding of L-serine to the PLP-cofactor of the LCA β-subunit at 25° C. (**A**) Time course for the reaction of 80 mM L-serine (syringe A) with 20 µM β-subunit (syringe B) in 50 mM EPPS/KOH pH 7.8. (**B**) Time course for the reaction of 80 mM L-serine (syringe A) with 20 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8. Time courses were followed by the increase in absorbance 420 nm. Data sets were fitted with single exponential functions.

The binding of L-serine to PLP leads to the formation of the external aldimine, which was monitored by the increase in absorbance at 420 nm (first phase, Figure 17). The rate of L-serine binding is not substantially altered within the complex but the amplitude is reduced two-fold. The relatively large rate constants indicate that binding

of the substrate L-serine is not limiting for the reaction of the isolated LCA β-subunit and for the LCA β-subunit in the TS. The subsequent reaction of the external aldimine to the aminoacrylate leads to a decrease in absorbance at 420 nm, which is also observable in Figure 17. It was analyzed in more detail (Figure 18).

Formation of aminoacrylate

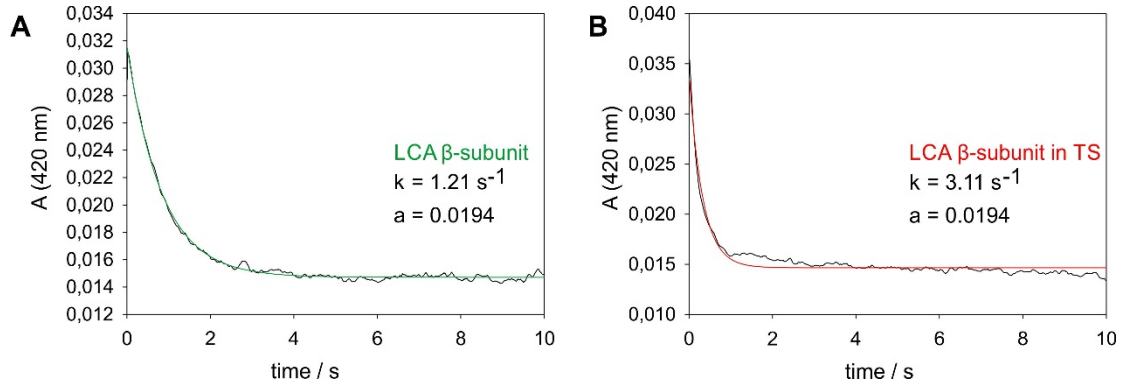


Figure 18: Formation of aminoacrylate.
Influence of TS formation on the reaction of external aldimine to aminoacrylate at 25° C. (**A**) Time course for the reaction of 80 mM L-serine (syringe A) with 20 µM β-subunit (syringe B) in 50 mM EPPS/KOH pH 7.8. (**B**) Time course for the reaction of 80 mM L-serine (syringe A) with 20 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8. Time courses were followed by the decrease in absorbance at 420 nm. Data sets were fitted with single exponential functions.

The amplitude is unaffected by complex formation. The rate of the dehydratization step, which is ascribed as the rate-limiting step in extant TS at low pH values, is increased by complex formation by a factor of three. This step might be rate-limiting for the isolated LCA β-subunit, but not for the LCA β-subunit in the TS. The observed rate is smaller than the k_{cat} of the isolated β-subunit (Figure 14) determined by steady state measurements, as it has been measured at 25° C and not at 60° C.

Reaction of indole with the aminoacrylate

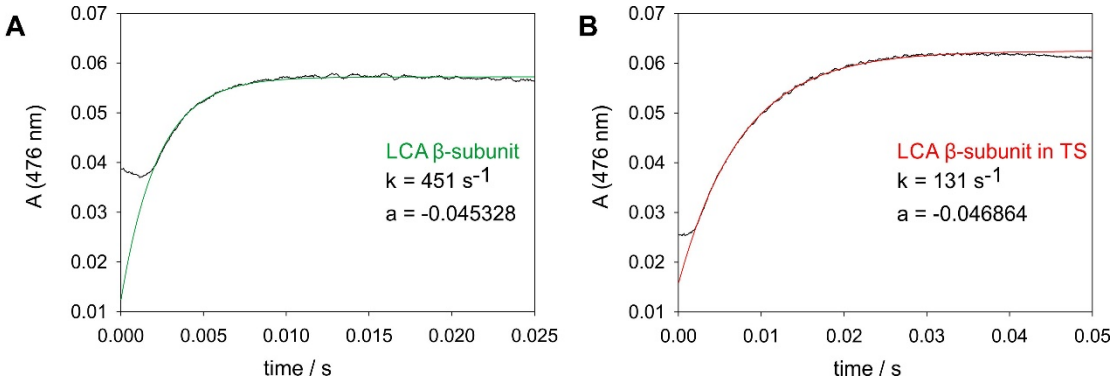


Figure 19: Reaction of indole with the aminoacrylate.
Influence of TS formation on the reaction of indole with the aminoacrylate at 25° C. (**A**) Time course for the reaction of 2 mM indole (syringe A) with 20 µM β-subunit (syringe B) in 50 mM EPPS/KOH pH 7.8 and 40 mM L-serine. (**B**) Time course for the reaction of 2 mM indole (syringe A) with 20 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8 and 40 mM L-serine. Time courses were followed by the increase in absorbance at 476 nm. Data sets were fitted with single exponential functions.

The conversion of the aminoacrylate to the quinoide II was monitored by the increase in absorbance at 476 nm (Figure 19). In the TS, the rate of the reaction of indole with the aminoacrylate is decreased, but the amplitude is unaffected by complex formation. The reaction is fast, which indicates that it is not rate-limiting for the reaction of the β-subunit in the TS. Complex formation seems to have no drastic effect on the accessibility of indole to the active site of the β-subunit.

Release of L-tryptophan

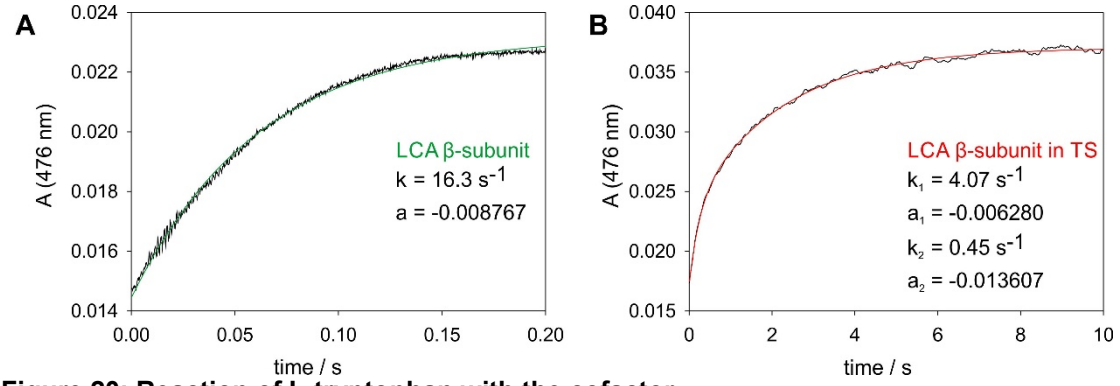


Figure 20: Reaction of L-tryptophan with the cofactor.
Influence of TS formation on the reaction of L-tryptophan with the cofactor of the LCA β-subunit at 25° C. (**A**) Time course for the reaction of 2 mM L-tryptophan (syringe A) with 20 µM β-subunit (syringe B) in 50 mM EPPS/KOH pH 7.8. (**B**) Time course for the reaction of 2 mM L-tryptophan (syringe A) with 20 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8. Time courses were followed by the increase in absorbance at 476 nm. Data sets were fitted with single exponential or double exponential functions.

The reverse reaction (binding of L-tryptophan and subsequent reaction to quinoide II) was monitored by the increase in absorbance at 476 nm (Figure 20). Within the TS, the reaction is dominated by a slow rate of 0.45 s^{-1}.

The conversion of the external aldimine to the aminoacrylate might be rate-limiting for the isolated LCA β-subunit. Within the TS, this reaction is faster. Quinoide II formation and/ or L-tryptophan release might be rate-limiting for the complexed LCA β-subunit as the reverse reactions are negatively influenced by complex formation. Pre-steady-state data in dependency of substrate concentrations will be collected at 60° C to allow for a direct comparison with the steady-state data.

1.4.7 Allosteric communication within the LCA TS

The enzymatic parameters of extant subunits are also influenced by allosteric communication within the TS. Ligand binding and catalysis at one subunit thereby help to coordinate the catalytical steps at the other subunit. To analyze whether allosteric communication was already present in the ancestor of permanent TS, the enzymatic parameters of the coupled LCA TS-reaction were determined (Figure 21).

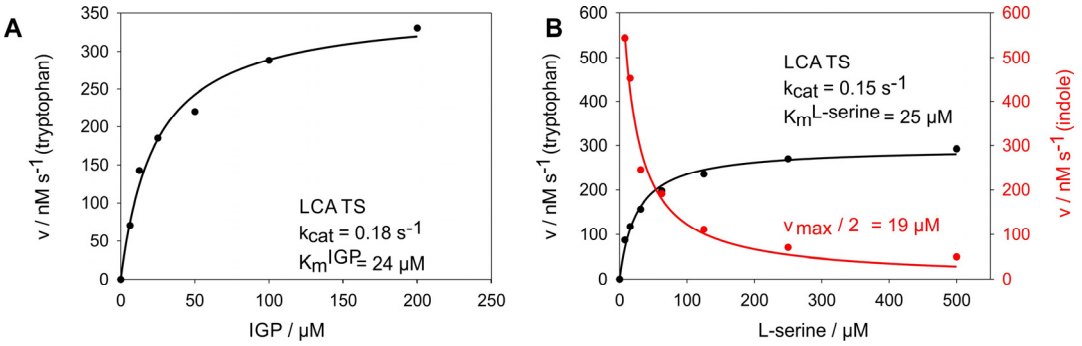


Figure 21: Enzymatic parameters of the working TS.
(**A**) IGP-dependent LCA TS-reaction. 100 µl reactions were started by the addition of IGP at 60° C in a thermocycler. The reactions contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 10 mM L-serine and 2 µM α- and β-subunits. Reactions were quenched with 50 µl 1N KOH. 100 µl of the quenched reactions were mixed with 100 µl methanol. The conversion to tryptophan was determined by reversed-phase HPLC using program 3. The program was adapted to a flow-rate of 0.5 ml/ min. Data points were fitted with a single rectangular hyperbolic function (Michaelis-Menten equation). (**B**) L-serine-dependent LCA TS-reaction. 100 µl reactions were started by the addition of 100 µM IGP at 60° C in a thermocycler. The reactions contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, varying amount of L-serine and 2 µM α- and β-subunits. Reactions were quenched with 50 µl 1N KOH. 100 µl of the quenched reactions were mixed with 100 µl methanol. Tryptophan and indole were quantified by reversed-phase HPLC using program 4. The program was adapted to a flow-rate of 0.5 ml/ min. Data points were fitted with a single rectangular hyperbolic function (Michaelis-Menten equation) and with a single hyperbolic decay function, respectively.

The reaction of IGP to indole and the reaction of IGP to tryptophan both lead to an absorbance change at 290 nm. In order to quantify the amount of formed indole and

tryptophan, the reactions were thus quenched and analyzed by HPLC. At saturating concentrations of L-serine, only tryptophan is formed (Figure 21 A). At sub-saturating concentrations of L-serine, tryptophan and indole are formed (Figure 21 B).

The changes in enzymatic parameters resulting from allosteric communication within the LCA TS are summarized in Table 4.

Table 4: Effect of allosteric communication on the LCA α- and β-reactions.

		α-reaction		β-reaction	
LCA	k_{cat}	0.18 s^{-1} 0.51 s^{-1}	3 x ↓	0.15 s^{-1} 0.86 s^{-1}	6 x ↓
	K_m	0.024 mM 0.083 mM	4 x ↓(IGP)	0.025 mM 1.2 mM	48 x ↓(L-serine)

Enzymatic parameters of the α-reaction (IGP-cleavage), the β-reaction (condensation of indole with L-serine) and the αβ-reaction (conversion of IGP and L-serine to tryptophan) within the LCA TS were determined at 60° C. Parameters of the α- and β-reactions are blue-colored, parameters of the αβ-reaction are black-colored.The relative changes in k_{cat} and K_m are indicated as fold increase (↑) and fold decrease (↓). Changes that increase the catalytic efficiency (k_{cat}/K_m) are green-colored. Changes that decrease the catalytic efficiency (k_{cat}/K_m) are red-colored.

The reaction of IGP and L-serine to tryptophan proceeds at a slower rate than the single reactions within the TS. This might indicate that the reaction at either the α-subunit or the β-subunit negatively influences one or more catalytic steps of the associated subunit. The reaction at the α-subunit leads to an increase in the affinity of the β-subunit for L-serine. This is similar as in the TS from *T. maritima* (Table 2) and might provide a mechanism that ensures the presence of L-serine at the β-subunit when indole is formed at the α-subunit.

1.4.8 Substrate channeling within the LCA TS

Within extant TS, the αβ-active sites are connected by a hydrophobic tunnel. This tunnel allows for the transfer of indole from the α-subunit to the β-subunit and prevents the loss of intermediary formed indole by free diffusion. Is the hydrophobic tunnel within the permanent TS an ancient ‘invention’ driven by a stringent selection for the quantitative conversion of IGP to L-tryptophan? Within interconnected TS, external indole has to enter the active site of the α-subunit to get to the active site of the β-subunit. As such, binding of the α-subunit ligand GP should block the access to the active site of the β-subunit and should prevent external indole from being converted to tryptophan. In this work, the accessibility of nucleophiles to the active

site of the LCA β-subunit was investigated. It was evaluated how complex formation with the α-subunit and additional blockage by the α-subunit ligand GP influences accessibility. In a first step, the small nucleophile N-methylhydroxylamine (NMHA) was used as probe for the accessibility of the β-subunit active site. Its reaction with the aminoacrylate at the β-active site leads to a change in the absorption at 456 nm (Figure 22).

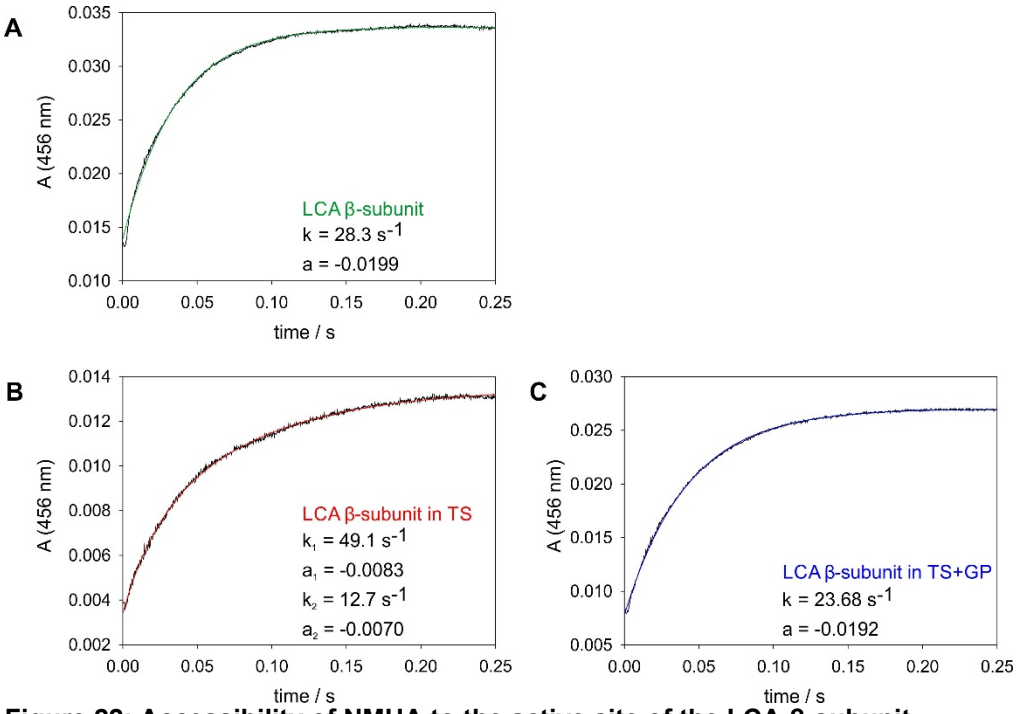


Figure 22: Accessibility of NMHA to the active site of the LCA β-subunit.
Influence of TS formation and GP binding on the reaction of NMHA with the aminoacrylate at the active site of the β-subunit at 25° C. (**A**) Time course for the reaction of 150 mM NMHA (syringe A) with 20 µM β-subunit in 50 mM EPPS/KOH pH 7.8 and 40 mM L-serine. (**B**) Time course for the reaction of 150 mM NMHA (syringe A) with 4 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8 and 40 mM L-serine. (**C**) Time course for the reaction of 150 mM NMHA (syringe A) with 4 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8, 40 mM L-serine and 100 mM GP. Time courses were followed by the increase in absorbance at 456 nm. Data sets were fitted with single exponential or double exponential functions.

Neither complex formation nor binding of GP at the active site of the α-subunit significantly influences the reaction rate of NMHA with the aminoacrylate. This is consistent with the observation that the blockage by complex formation and GP is insufficient to prevent the passage of small nucleophiles to the active site of the β-subunit (Dunn et al., 1990). However, a large nucleophile similar in size to indole should be efficiently prevented from entering the active site of the β-subunit by complex formation and GP binding if it has to pass through the active site of the α-subunit to get to the active site of the β-subunit. Benzimidazole (BZI) was used as such a large nucleophilic probe as displacement of β-active site bound aniline by BZI can be monitored by the decrease in absorbance at 466 nm (Figure 23).

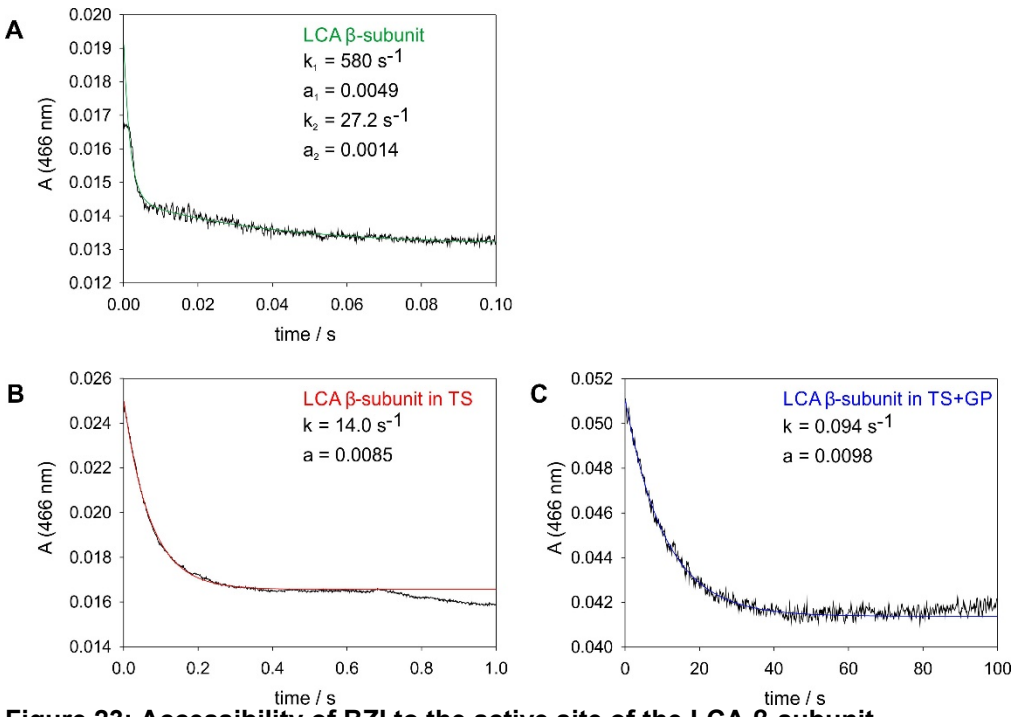


Figure 23: Accessibility of BZI to the active site of the LCA β-subunit.
Influence of TS formation and GP binding on the reaction of BZI with the aniline-quinoide at the active site of the β-subunit at 25° C. (**A**) Time course for the reaction of 20 mM BZI (syringe A) with 10 µM β-subunit (syringe B) in 50 mM EPPS/KOH pH 7.8, 100 mM anilinium chloride and 40 mM L-serine. (**B**) Time course for the reaction of 20 mM BZI (syringe A) with 10 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8, 100 mM anilinium chloride and 40 mM L-serine. (**C**) Time course for the reaction of 20 mM BZI (syringe A) with 10 µM α- and β-subunits (syringe B) in 50 mM EPPS/KOH pH 7.8, 100 mM anilinium chloride, 40 mM L-serine and 100 mM GP. Data sets were fitted with single exponential or double exponential functions.

The significant decrease in the reaction rate upon complex formation and GP binding indicates that a large nucleophile has to enter the active site of the α-subunit to get to the active site of the β-subunit. Thus the αβ-connection is the most plausible physical pathway for indole in the LCA TS like it is in extant TS (Dunn et al., 1990).

1.5 Conclusion

Most enzymes reconstructed up to date are relatively young and are thus almost identical to their extant descendants (Chandrasekharan et al., 1996; Malcolm et al., 1990; Stackhouse et al., 1990; Thomson et al., 2005). Only some enzymes have been reconstructed, which strikingly differ from their extant descendants. Among those are the 3-isopropylmalate dehydrogenase LeuB from the ancestor of Bacteria (Hobbs et al., 2012) and the imidazole glycerol phosphate synthase HisF from the last universal common ancestor (Reisinger et al., 2014). Those share only 76% respectively 78% sequence identity with their closest related extant enzymes.

In this work, the α- and β-subunits of the TS from the last common ancestor of Bacteria (LCA) were reconstructed. Those enzymes can be supposed to be active in the LCA due to their central role in primary metabolism. The presence of enzymatic activity thus serves as a reliable indicator for the reconstruction success. As the LCA α- and β-subunits are catalytically active, an accurate reconstruction was shown to be achievable even in the case of the LCA α-subunit, which shares only 57% sequence identity with the closest related extant α-subunit. A recent study on bacterial ancestral nucleoside diphosphate kinases showed that they were thermostable. This finding let the authors to the conclusion that the environmental temperature in the LCA era was ~80- 90° C (Akanuma et al., 2013). Consistently, the LCA α- and β-subunits are thermostable as well.

The LCA α- and β-subunits form a complex. For the first time, an ancient enzyme complex was successfully reconstructed. The subunits within the complex influence each other during catalysis. This reveals an early bi-directional adaptation of enzymes within a complex and the ancient nature of allosteric communication. The rate of tryptophan synthesis of the LCA TS is similar to the rate found in 'less evolved' extant TS like the *S. solfataricus* TS (Leopoldseder et al., 2006). *S. solfataricus* is tryptophan prototroph, which testifies that a slow rate of tryptophan synthesis is sufficient for a slowly growing organism.

The interconnection between the active sites allows for the channeling of intermediary formed indole in the LCA TS. Besides allosteric communication, metabolite channeling seems to have already already existed in an ancestor of Bacteria. This mechanism prevents that free indole is produced in proportion to tryptophan.

The LCA TS has a strong IGP lyase activity in the absence of L-serine that has not been observed in extant TS. Such being the case, the LCA TS can synthesize indole or tryptophan in dependency of the intracellular concentration of L-serine, whereas extant TS can only act as tryptophan synthases. Consistently, it was proposed that ancestral enzymes are less specialized than extant enzymes (Jensen, 1976). For instance, ancestral β-lactamases and α-glucosidases were shown to have broad substrate specificities (Risso et al., 2013; Voordeckers et al., 2012). The ancestral-type GAL1 was further shown to function as galactose phosphorylating enzyme and as regulatory protein, whereas modern-type GAL1 can only act as galactose phosphorylating enzyme (Conant and Wolfe, 2008; Meyer et al., 1991).

In the extant microorganism *E. coli*, the intracellular concentration of L-serine is approximately 100 µM (Bennett et al., 2009). Taken a similar concentration of L-serine in the LCA into account, the depletion of L-serine would have resulted in an increase of the intracellular indole concentration. The occurrence of free indole might explain, why the LCA also contained a β-subunit homolog (Merkl, 2007), which utilizes free indole for the synthesis of tryptophan. Besides the utilization of free indole by other metabolic enzymes, indole might also have served as signaling molecule. It is presumed to be an universal inter-cellular signaling molecule (Bunders et al., 2011), which for instance regulates the transition from exponential to stationary phase in extant bacteria (Lelong et al., 2007). In the course of evolution, the TS might have specialized as tryptophan synthase and indole became exclusively synthesized by a tryptophanase in bacteria (Isupov et al., 1998; Ku et al., 2006) and by a IGP lyase in plants (Kulik et al., 2005).

1.6 Ongoing research and future work

Current investigations focus on the determination of the ligand-bound LCA TS structure. This structure should help to understand the molecular details underlying ancient inter-subunit communication. Protein crystals were obtained in several conditions of commercial screening kits (6.5), but their diffraction was insufficient for structural determination (>4 Å). Those conditions are currently optimized. As mentioned before, the LCA TS has to be characterized by pre-steady-state kinetics at 60° C to allow for a direct comparison with the steady-state data. Microscopic rate constants for different steps in the catalytic mechanism, e.g. binding of L-serine, reaction of external aldimine to aminoacrylate, release of tryptophan, will be determined by varying ligand concentrations and interpreting concentration-dependent secondary plots. Pre-steady-state measurements should also help to determine the rate-limiting step of the coupled TS reaction (conversion of IGP to tryptophan). For instance, the initial rate of IGP cleavage and tryptophan formation should indicate whether the release of tryptophan becomes more impaired in the working LCA TS or whether the α-reaction is rate-limiting for the overall reaction.

2 SUBSTRATE SPECIFICITY OF β-SUBUNIT HOMOLOGS

2.1 Introduction

2.1.1 Phylogenetic analysis of the TrpB enzyme family

Analysis of sequenced genomes let to the identification of a group of enzymes that share a sequence identity of approximately 30 % with the β-subunit (TrpB1) of the permanent αββα tryptophan synthase complex (TS) (Merkl, 2007; Xie et al., 2002). Accordingly, the β-subunit homologs were named TrpB2 enzymes. They were subdivided into TrpB2i enzymes, which are encoded within the *trp* operon, and TrpB2o and TrpB2a enzymes, which are encoded outside of the *trp* operon (Figure 24).

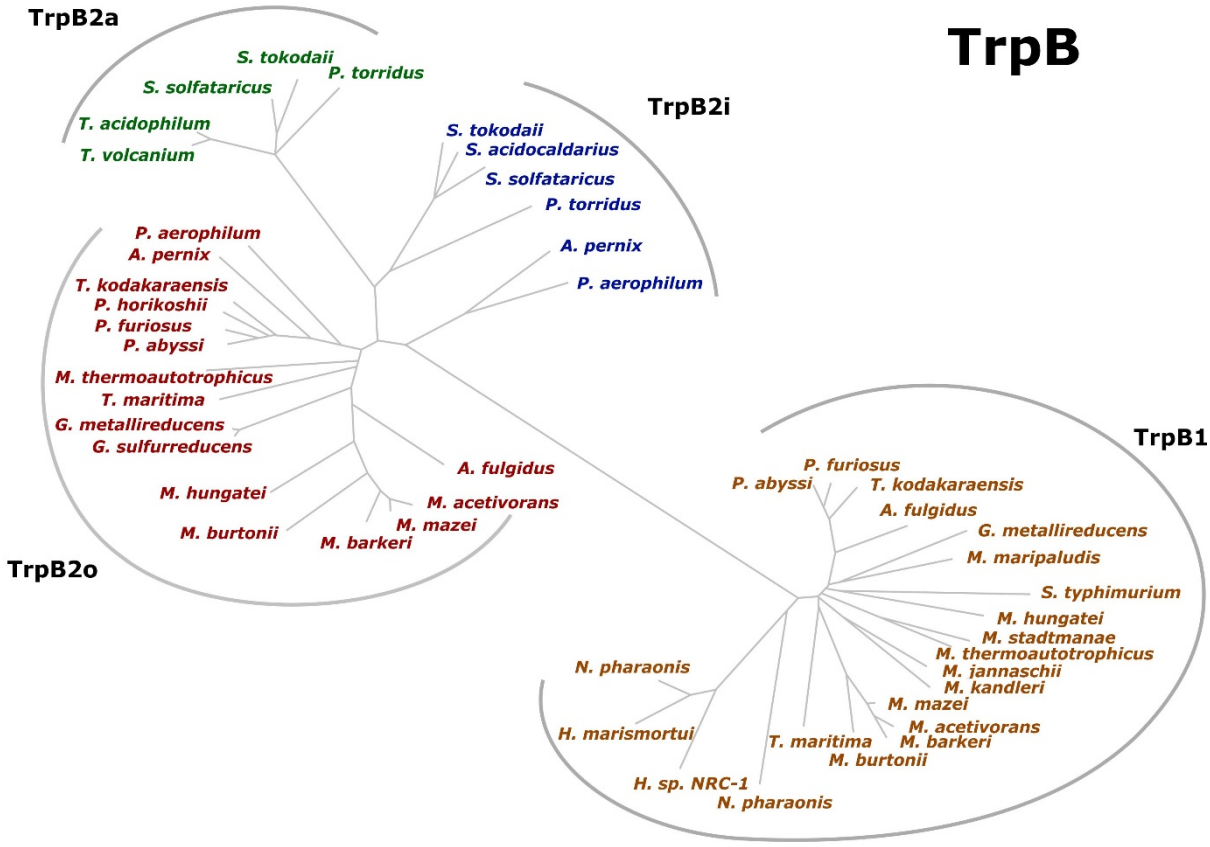


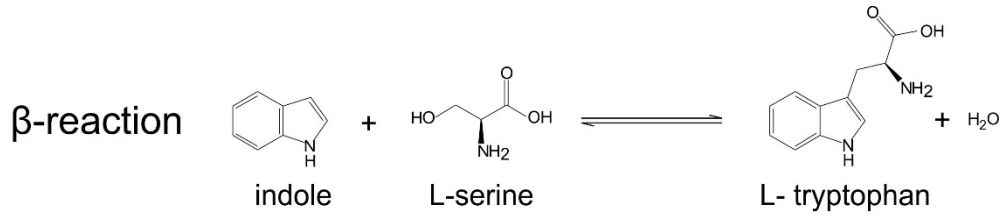
Figure 24: Phylogenetic tree of TrpB enzymes.
The phylogenetic tree is based on a multiple sequence alignment (MSA) of bacterial and archaeal TrpB genes. Figure according to (Merkl, 2007).

2.1.2 Abundance and properties of TrpB2 enzymes

*trp*B2i and *trp*B2a genes occur exclusively in Archaea that lack a *trp*B1 gene. As shown for ssTrpB2i from *Sulfolobus solfataricus*, TrpB2i is part of a transient TS with unusual αββ stoichiometry under physiological conditions (Ehrmann et al., 2010; Leopoldseder et al., 2006).

In contrast, *trp*B2o genes are found in the genomes of Archaea, Bacteria and Eukarya in addition to a *trp*B1 gene. The *trp*B2o genes are transcribed (Yin et al., 2010), but transcriptional regulation is hardly investigated and the genomic neighborhood of *trp*B2o is not conserved. For *Methanothermobacter thermoautotrophicus*, TrpB2o expression was shown to be regulated by L-tryptophan (Karr et al., 2008). However, TrpB2o enzymes are not part of a tryptophan synthase complex (TS) and no other interaction partner could be identified so far (Ehrmann, 2011; Hettwer and Sterner, 2002; Hiyama et al., 2014). As a knockout of *trp*B2o doesn't affect the ability of an organisms to grow in medium lacking L-tryptophan, TrpB2o enzymes were considered to be expandable for the biosynthesis of tryptophan *in vivo* (Ehrmann, 2011).

All TrpB2 enzymes have in common a low catalytic efficiency $k_{cat}/K_m^{\text{L-serine}}$, due to a K_m for L-serine in the high millimolar range (Figure 25).



TrpB1 (*E. coli*) k_{cat}= 0.09 s^{-1} $K_m^{\text{L-serine}}$= 0.95 mM K_m^{indole}= 0.014 mM
TrpB2o (*T. maritima*) k_{cat}= 0.44 s^{-1} $K_m^{\text{L-serine}}$= 50.2 mM K_m^{indole}= <0.77 µM

Figure 25: L-serine-dependent synthesis of tryptophan.
The enzymatic parameters of TrpB1 from *E. coli* and of TrpB2 from *T. maritima* for the L-serine-dependent synthesis of L-tryptophan are shown (Hettwer and Sterner, 2002).

It had been proposed that those TrpB2 enzymes, which are not part of a TS, might act as indole salvage enzymes *in vivo* or might have a function apart from tryptophan biosynthesis. The hypothesis of TrpB2o as indole-salvage enzyme is challenged by its high K_m for L-serine. No reaction apart from tryptophan synthesis was identified that is catalyzed by TrpB2o enzymes with reasonable catalytic efficiency (Ehrmann, 2011; Hettwer and Sterner, 2002; Xie et al., 2002; Yin et al., 2010).

2.1.3 Hypothesis on the function of TrpB2 enzymes

In order to identify the physiological function of TrpB2 enzymes, the strains *S. solfataricus* Δ*trp*B2a and *T. kodakaraensis* Δ*trp*B2o were used. A comparison with the corresponding wild-type strains revealed that the knockout strains thrive equally well in medium lacking tryptophan (Busch, 2010; Ehrmann, 2011). This confirmed that the TrpB enzyme, which is encoded within the *trp* operon (TrpB1 or TrpB2i) is sufficient for the biosynthesis of tryptophan *in vivo*. Comparing the metabolomes and proteomes of wild-type and knockout strains gave no clear-cut hint on the physiological role of 'additional' TrpB2 enzymes (Busch, 2010; Ehrmann, 2011). Thus 'additional' TrpB2 enzymes may either have no unique function or are dispensable under the investigated conditions. The lack of a clear-cut phenotype asked for a different strategy for the functional annotation of TrpB2 enzymes. Within the last years, knowledge-based approaches, which combine computational and experimental methods were highly successful in functional annotation projects (Brown and Babbitt, 2012; Pandya et al., 2014; Zhao et al., 2013). Thus such an approach was also used for the functional annotation of TrpB2 enzymes. Information on the occurrence, catalytic mechanism, structure and conserved residues was combined for an initial *in vitro* screening with putative substrates. The substrate specificity was subsequently confirmed by X-ray crystallography and mutagenesis studies.

2.2 Remarks

Dr. Chitra Rajendran collected the X-ray data and solved the structure of ssTrpB2a with bound O-phospho-L-serine. Patrick Löffler modeled the ssTrpB2a structure and created sequence logos. Parts of this chapter have been published equally worded in (Busch et al., 2014). The publication was written by Prof. Dr. Reinhard Sterner and myself.

2.3 Significance of this work

The rapid increase in the number of sequenced genomes asks for the functional annotation of the encoded enzymes. A combined computational-structural approach was used to determine the function of the TrpB2 subgroup of the tryptophan synthase beta chain/ beta chain-like TrpB1-TrpB2 family. The results showed that

TrpB2 enzymes are O-phospho-L-serine-dependent tryptophan synthases, whereas TrpB1 enzymes catalyze the L-serine-dependent synthesis of tryptophan. A single residue is responsible for the different substrate specificities of TrpB1 and TrpB2. This was confirmed by mutagenesis studies and crystallographic analysis of a TrpB2 enzyme with bound O-phospho-L-serine.

2.4 Results and discussion

2.4.1 Phylogenetic distribution of TrpB1 and TrpB2 enzymes

In order to gain as complete a picture as possible on the distribution of TrpB enzymes, the InterPro database (v.49) was searched for the occurrence of TrpB1 enzymes (IPR006654) and TrpB2 enzymes (IPR006316) in different phyla (Figure 26).

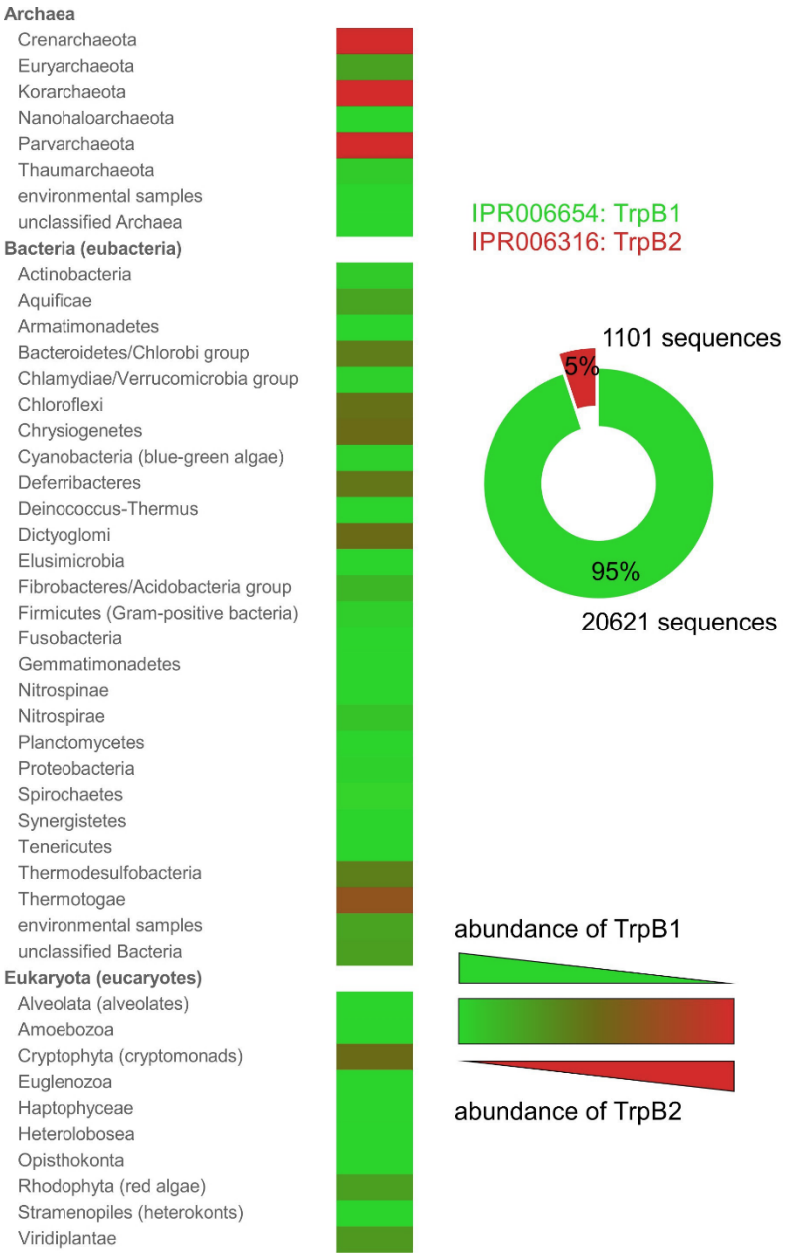
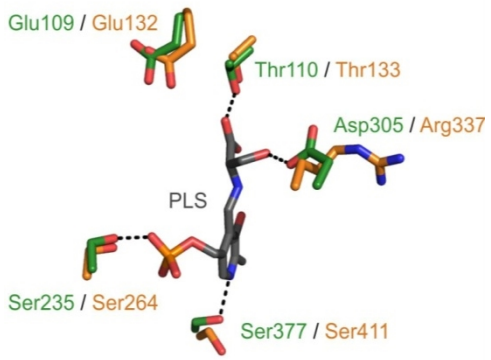


Figure 26: Occurence of TrpB enzymes.
Information on the abundance of TrpB1 and TrpB2 was deduced from the InterPro database (v.49). The relative abundance is depicted as color gradient ranging from green (only TrpB1 enzymes in a phylum) to red (only TrpB2 enzymes in a phylum).

TrpB2 enzymes occur exclusively in Parvarchaeota, Korarchaeota and Crenarchaeota, but they co-occur with TrpB1 enzymes in phyla of all three domains of life. The widespread distribution of TrpB2 enzymes might indicate that they have arisen before the diversification of life. However, many archaeal, bacterial and eukaryotic lineages seem to have lost the *trp*B2 gene most probable due to the lack of a selectable advantage on their physiological function in the presence of *trp*B1.

2.4.2 Comparison of conserved residues in TrpB1 and TrpB2 enzymes

The detailed reaction mechanism of TrpB2 enzymes is unknown, and no crystal structure has been available for this group so far. In contrast, the structure-function relationship of the TrpB1 group is well understood, mainly based on studies of stTrpB1 from *Salmonella typhimurium*. Here, catalysis involves the formation of an external aldimine between the cofactor PLP and the substrate L-serine. In the first instance, the coordination of this intermediate in stTrpB1 was analyzed and compared to a most plausible coordination in a homology model generated for ssTrpB2a from *S. solfataricus* (Figure 27).



	stTrpB1	ssTrpB2a
binding of cofactor	Ser235	Ser264
cofactor chemistry [1-3]	Ser377	Ser411
nucleophile specificity [4]	Glu109	Glu132
amine specificity [5-6]	Thr110 Asp305	Thr133 Arg337

Figure 27: Superposition of active sites of stTrpB1 and ssTrpB2a.
The crystal structure of stTrpB1 from *S. typhimurium* and the modeled structure of ssTrpB2a from *S. solfataricus* were superimposed.The residues in close proximity to the L-serine bound cofactor (PLS) are shown as sticks. Residues of stTrpB1 are orange-colored, residues of ssTrpB2a are green-colored and the PLP-L-serine aldimine is gray-colored. The role of the depicted residues are indicated as being analyzed for stTrpB1. [1-3](Hyde et al., 1988; Jhee et al., 1998; Rhee et al., 1997), [4](Brzovic et al., 1992), [5-6](Ferrari et al., 2003; Ferrari et al., 2001).

The cofactor PLP is bound in the same position and the nitrogen of the pyridinium ring is coordinated by equivalent serine residues in stTrpB1 and ssTrpB2a, which indicates that both enzyme groups use the same cofactor chemistry. The specificity for using indole as nucleophile is determined by a glutamate residue, which is strictly

conserved in all TrpB enzymes. Moreover, both TrpB subfamilies utilize a threonine residue to coordinate the carboxyl moiety of bound L-serine. However, whereas the hydroxyl-group of L-serine is coordinated by an aspartate in TrpB1, all TrpB2 enzymes have an arginine at that position (Figure 28).

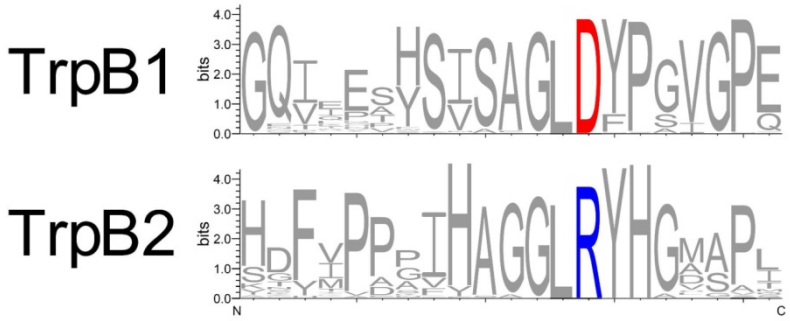


Figure 28: Sequence logos for the TrpB1 and TrpB2 enzyme group.
Sequence logos were created by WebLogo (Crooks et al., 2004; Schneider and Stephens, 1990) using a multiple sequence alignment (MSA) of each 100 randomly selected TrpB1/TrpB2 sequences. The substrate specifying residue within the depicted section is highlighted by colour.

The consequences of this difference for substrate specificity were analyzed by testing different TrpB enzymes for their ability to catalyze the conversion of indole with α-amino acids other than L-serine. For this purpose, the stTrpB1 enzyme from *Salmonella typhimurium* and various TrpB2 enzymes (ssTrpB2i and ssTrpB2a from *Sulfolobus solfataricus*, tmTrpB2o from *Thermotoga maritima* and atTrpB2o from *Arabidopsis thaliana*) were cloned, expressed and purified.

2.4.3 Cloning of TrpB1 and TrpB2 enzymes

st*trp*B1 was amplified from pBR322-st*trp*AB using the oligonucleotides AGC CAT ATG ACA ACA CTT CTC AAC CCC TAC/ CTG GTG CAA GCT TGA TTT CCC CTC GCG CTT TCA GGA TATC and inserted into pET24a(+) at the NdeI/ HindIII restriction sites.

ss*trp*B2a was amplified from pET28a(+)-ss*trp*B2a (Leopoldseder et al., 2006) using the oligonucleotides TAA TAC GAC TCA CTA TAG GG/ CCG CAA GCT TCT CCT TAA ATA ACA C and inserted into pET24a(+) at the NdeI/ HindIII restriction sites.

ss*trp*B2i cloning and insertion into pET28a(+) has been described before (Leopoldseder et al., 2006).

tm*trp*B2o was amplified from pET21a(+)-tm*trp*B2o (Hettwer and Sterner, 2002) using the oligonucleotides ACC GCA TAT GAG AAT TGT TGT GAA/ CCC AGG AAT TCA GGC TTT CAC ACG TAC GCT GT and inserted into pET28a(+) at the NdeI/ HindIII restriction sites.

at*trp*B2o was amplified from cDNA of *Arabidopsis thaliana* Col-0. The yield of amplification product was increased by two subsequent PCR reactions. In the first round of PCR, the oligonucleotides GCA GCT TTG AGA TCT ACT CA/ TTA TGG GGC CAT TCG AGC TT were used. The amplification product was the template in a second round of PCR using the oligonucleotides CTA GCT TAA GAC ATA TGG CAG CTT TGA GA/ TTA TGG GGC CAT GGA TCC TTA AAC AAC A. The final amplification product was inserted into the pET28a(+) expression vector at the NdeI/ BamHI restriction sites.

2.4.4 Expression and purification of TrpB1 and TrpB2 enzymes

For expression of stTrpB1, *E. coli* T7 Express was transformed with pET24a(+)-st*trp*B. The cells were grown at 37° C in LB with 50 µg/ ml kanamycin to OD_{600}= 0.5. Protein expression was induced by addition of 0.5 mM IPTG. After growth over night at 20° C, cells were harvested by centrifugation, resuspended in 25 ml/ l culture 50 mM Tris/HCl pH 7.5, 150 mM NaCl and 10 mM imidazole, and disrupted by sonication. The His_6-tagged protein was purified by metal chelate affinity chromatography using a HisTrap FF crude column. Proteins in 50 mM Tris/HCl pH 7.5 and 150 mM NaCl were eluted by a linear gradient of imidazole (10- 500 mM) and dialyzed against 50 mM Tris/HCl pH 7.5.

The proteins ssTrpB2i, ssTrpB2a, ssTrpB2a and tmTrpB2o were expressed with a His_6-*tag* and purified by a heat step and metal chelate affinity chromatography according to (Hettwer and Sterner, 2002) and (Leopoldseder et al., 2006).

For expression of atTrpB2o, *E. coli* (DE3) was transformed with pET28a(+)-at*trp*B2o. The cells were grown at 37° C in LB with 50 µg/ ml kanamycin to OD_{600}= 0.5. Protein expression was induced by addition of 0.5 mM IPTG. After growth over night at 37° C, cells were harvested by centrifugation, resuspended in 50 mM potassium phosphate pH 7.5, 300 mM KCl and 10 mM imidazole, and disrupted by sonication. The His_6-tagged protein was purified by metal chelate affinity chromatography using a HisTrap FF crude column. Proteins in 50 mM potassium phosphate pH 7.5 and 300 mM KCl were eluted by a linear gradient of imidazole (10-

500 mM) and dialyzed against 50 mM potassium phosphate pH 7.5. The proteins were at least 90 % pure as judged by SDS-PAGE (Figure 29).

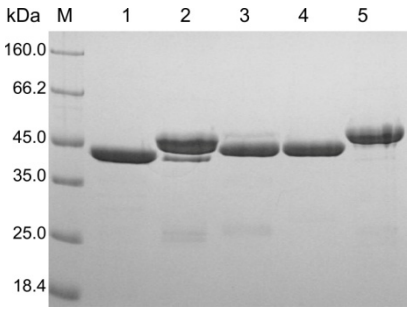
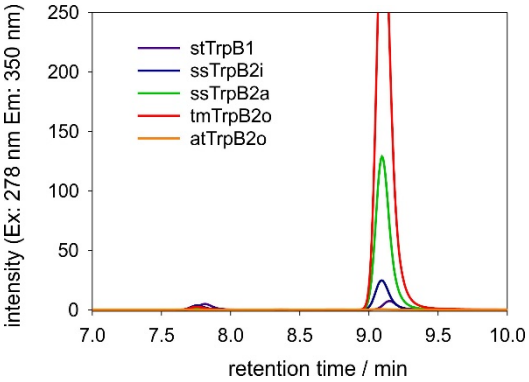


Figure 29: Purity of TrpB1- and TrpB2-proteins.
SDS-PAGE (12.5 % polyacrylamide) of purified TrpB-proteins. Applied were (**M**) protein ladder and marker (LMW), (**1**) 15 µl stTrpB1 (10 µM monomer concentration), (**2**) 15 µl ssTrpB2i (10 µM monomer concentration), (**3**) 15 µl ssTrpB2a (10 µM monomer concentration), (**4**) 15 µl tmTrpB2o (10 µM monomer concentration), (**5**) 15 µl atTrpB2o (10 µM monomer concentration).

2.4.5 Screen for the synthesis of β-substituted tryptophans

The ability of TrpB enzymes to catalyze the condensation reaction of L-threonine, D-threonine, DL-phenylserine or O-phospho-L-threonine with indole to a β-substituted tryptophan derivative was analyzed by HPLC. Solely L-threonine is converted with indole to β-methyltryptophan, a known building block for the maremycine synthesis (Zou et al., 2013). The stereoselectivity of this reaction was analyzed by HPLC separation followed by fluorescence detection (Figure 30).



	stTrpB1	ssTrpB2i	ssTrpB2a	tmTrpB2o	atTrpB2o
relative yield / %	3	8	36	100	0
ee (2S,3S/2S,3R)	17	73	98	99	0

Figure 30: Synthesis of β-methyltryptophan.
Detection of β-methyltryptophan by HPLC. 10 µM TrpB enzyme were incubated with 500 µM indole and 250 mM L-threonine in 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP for 30 min at 40°C (stTrpB1, atTrpB2o) or 80°C (tmTrpB1, tmTrpB2o, ssTrpB2a, ssTrpB2i). The reactions were quenched with methanol and analyzed by HPLC using program 2. The retention time for 2S,3R-methyltryptophan was 7.8 min and the retention time for 2S,3S-methyltryptophan was 9.1 min consistent with previous investigations (Zou et al., 2013).

All TrpB enzymes except atTrpB2o are able to synthesize β-methyltryptophan using L-threonine, albeit with low efficiency. TrpB2 enzymes seem to be stereoselective for the synthesis of 2S,3S-methyltryptophan. The low efficiency for the synthesis of β-methyltryptophan indicates that this reaction is not of physiological relevance. The utilization of L-threonine as substrate may be an evolutionary relict, as TrpB enzymes and threonine synthases share a common ancestor (Parsot, 1987).

2.4.6 Screen for the synthesis of tryptophan

Next, the ability of TrpB enzymes to catalyze the condensation reaction of L-serine, D-serine, L-cysteine, D-cysteine, DL-diaminopropionate, O-acetyl-L-serine, O-phospho-L-serine or O-phospho-D-serine with indole to tryptophan was analyzed by HPLC (Table 5).

Table 5: Conversion of different amino acids to tryptophan.

	stTrpB1	ssTrpB2i	ssTrpB2a	tmTrpB2o	atTrpB2o
L-serine	+++	++	-	-	-
D-serine	-	-	-	-	-
L-cysteine	-	+	-	-	-
D-cysteine	-	-	-	-	-
DL-diaminopropionate	-	-	-	-	-
O-acetyl-L-serine	++	++	-	+	-
O-phospho-L-serine	-	+++	+	+++	+++
O-phospho-D-serine	-	-	-	-	-

Conversion of 500 µM indole in 100 mM potassium phosphate pH 7.5, 180 mM KCl and 40 µM PLP with 2 mM of the indicated α-amino acid by 5 µM TrpB enzyme was analyzed after 30 min at 25° C (stTrpB1) or 60° C (ssTrpB2i, ssTrpB2a, tmTrpB2o, atTrpB2o) by HPLC using program 3. Symbols: +++: conversion of > 60 % indole, ++: conversion of 20-60 % indole, +: conversion of 10-20 % indole, -: conversion of < 10 % indole.

Whereas stTrpB1 preferentially utilizes L-serine as substrate, all tested TrpB2 enzymes have a preference for O-phospho-L-serine (OPS) as substrate. HPLC chromatograms for the OPS-dependent tryptophan synthesis are shown in Figure 31.

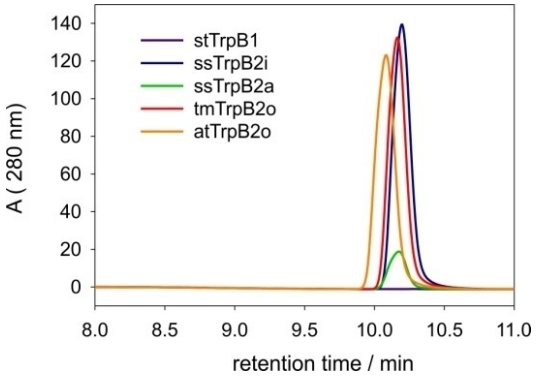


Figure 31: OPS-dependent synthesis of tryptophan.
HPLC chromatograms for the OPS-dependent synthesis of tryptophan. 100 µl reaction mixtures contained 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, 500 µM indole, 2 mM OPS and 5 µM enzyme. After incubation for 30 min at 25° C (stTrpB1) or 60° C (ssTrpB2i, ssTrpB2a, tmTrpB2o, atTrpB2o), reactions were quenched by the addition of 400 µl methanol. Tryptophan was detected by HPLC using program 2.

Next, the steady-state kinetic parameters of the TrpB2 enzymes for the OPS-dependent synthesis of tryptophan were determined (Figure 32).

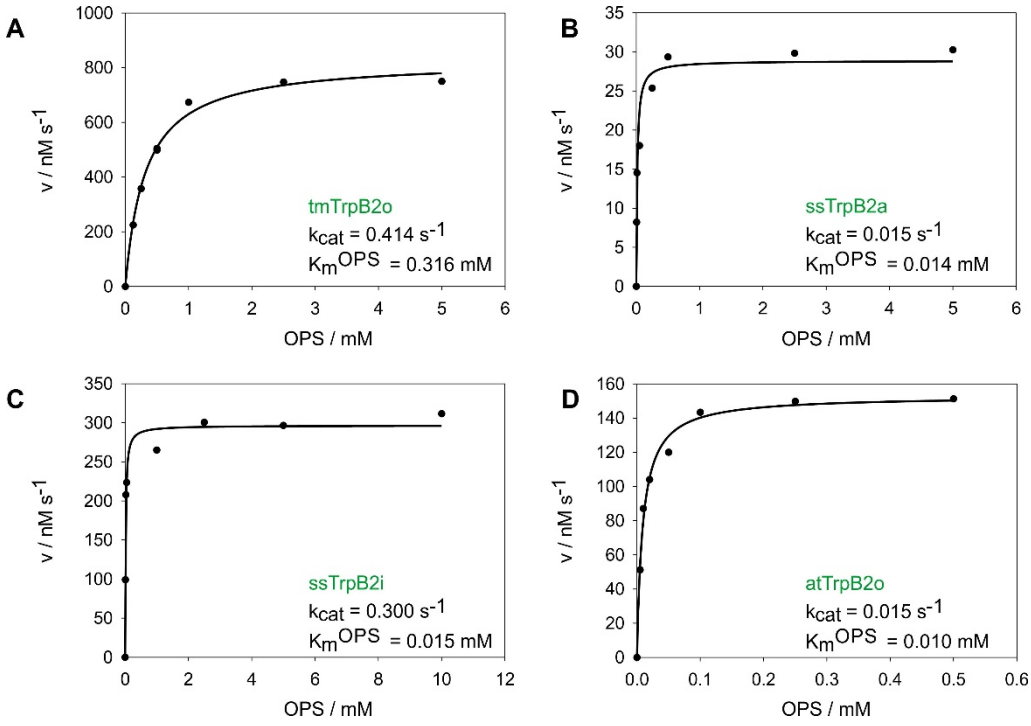


Figure 32: Enzymatic parameters for the OPS-dependent synthesis of tryptophan.
OPS-dependent TrpB2 reaction. The reactions were started by the addition of 100 µM indole and followed by the absorbance change at 290 nm. The reactions contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, varying amounts of OPS and TrpB2 protein. The reactions were measured (**A**) at 80° C with 2 µM tmTrpB2o, (**B**) at 60° C with 2 µM ssTrpB2a, (**C**) at 60° C with 1 µM ssTrpB2i and (**D**) at 30° C with 10 µM atTrpB2o. Data points were fitted with a single rectangular hyperbolic function (Michaelis-Menten equation).

The determined parameters were compared with the published parameters for the L-serine-dependent synthesis of tryptophan (Table 6).

Table 6: Comparison of the OPS- and L-serine-dependent TS reaction.

	L-serine-dependent reaction		
	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
tmTrpB2o	0.44	50.2	$8.7 \cdot 10^{0}$
ssTrpB2a	0.032	151	$2.1 \cdot 10^{-1}$
ssTrpB2i	0.20	35	$5.7 \cdot 10^{0}$
atTrpB2o	0.016	35	$4.5 \cdot 10^{-1}$
	OPS-dependent reaction		
	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($M^{-1}s^{-1}$)
tmTrpB2o	0.414	0.316	$1.3 \cdot 10^{3}$
ssTrpB2a	0.015	0.014	$1.1 \cdot 10^{3}$
ssTrpB2i	0.300	0.015	$2.0 \cdot 10^{4}$
atTrpB2o	0.015	0.010	$1.5 \cdot 10^{3}$

Parameters for the L-serine-dependent reaction at 80° C (tmTrpB2o), 60° C (ssTrpB2a, ssTrpB2i) and 30° C (atTrpB2o) are according to (Hettwer and Sterner, 2002), (Leopoldseder et al., 2006) and (Yin et al., 2010). Conditions for the OPS-dependent tryptophan synthase reaction: 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP and 100 µM indole.

The K_m^{OPS} values of all TrpB2 lie within range of 10- 1000 µM, which is found for ~60 % of enzymes in the KEGG database (Bar-Even et al., 2011). Since these values are much lower than their $K_m^{L\text{-serine}}$ values and as the corresponding turnover numbers k_{cat} are similar, the catalytic efficiencies k_{cat}/K_m^{OPS} are higher by about 3- 4 orders of magnitude than the catalytic efficiencies $k_{cat}/K_m^{L\text{-serine}}$. This indicates that TrpB2 enzymes catalyze the reaction of indole with OPS *in vivo*. This is a new role for OPS, which has been known up to now only as an intermediate of L-serine, L-cysteine and L-cystathionine biosynthesis (Helgadottir et al., 2007; Mino and Ishikawa, 2003). The coordination of OPS within the active site of a TrpB2 enzyme was subsequently analyzed by x-ray crystallography.

2.4.7 Crystallization of ssTrpB2a

ssTrpB2a was expressed with a His_6-*tag* and purified by a heat step and metal chelate affinity chromatography according to (Hettwer and Sterner, 2002) and (Leopoldseder et al., 2006). Subsequently, it was further purified by preparative size exclusion chromatography and dialyzed against 10 mM HEPES pH 7.5 and 25 mM NaCl. Crystallization trials with ssTrpB2a were performed with the vapour diffusion method in 24 well plates at 18° C based on previously established conditions (O. Mayans, unpublished data). Drops contained 1 µl of ssTrpB2a (12.5 mg/ ml) in

10 mM HEPES pH 7.5 and 25 mM NaCl and 1 µl of reservoir solution. Equilibration was done against 500 µl reservoir solution. First crystals appeared after 1 week at 18°C with 25 % PEG-4000, 0.1 M Tris/HCl pH 7.0 and 100 mM NaCl as reservoir solution (Figure 33).

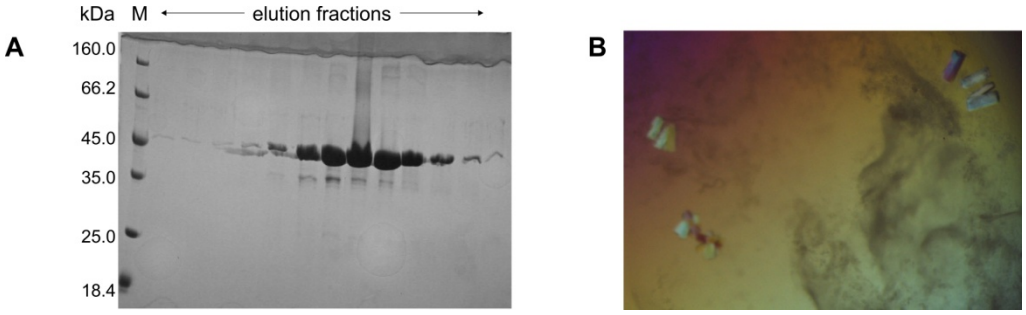


Figure 33: Crystallization of ssTrpB2a.
(**A**) SDS-PAGE (12.5 % polyacrylamide) of ssTrpB2a. Applied were the elution fractions from preparative size exclusion chromatography. (**B**) Crystals of ssTrpB2a. The crystals appeared after 1 week at 18° C with 25 % PEG-4000, 0.1 M Tris/HCl pH 7.0 and 100 mM NaCl as reservoir solution.

Initially crystals appeared as clusters and were multicrystals. Seeding was used to produce bigger single crystals with 20-25 % PEG-4000, 0.1 M Tris/HCl pH 7.0/7.5/8.0/8.5 and 100 mM NaCl as reservoir solution.

2.4.8 Structure of ssTrpB2a with bound OPS

After 1 month at 18°C, single crystals were collected, soaked with 200 mM O-phospho-L-serine and flash frozen in liquid nitrogen. The data collection and refinement statistics are shown in Table 7.

Table 7: Data collection and refinement statistics.

Protein	**ssTrpB2a**
Wavelength (Å)	1.0 Å
Unit cell (Å)	a=55.56 b=61.903 c=109.57 α=90 β=98.23 γ= 90
(Space group)	SG P 1 21 1
Total reflections	169446 (9645)
Unique reflections	52794 (4406)
Redundancy	3.2 (2.2)
Resolution (Å)	1.94
Completeness (%)	96.28 (81.64)
I/σ(I)	8.22 (1.52)
R_{merge}	0.084 (0.64)
Refinement	
Resolution (Å)	46.44 - 1.94 (2.00 - 1.94)
R_{work} / R_{free} (%)	0.195 (0.298) / 0.252 (0.319)
Av. B (Å^2)	34.50
Number of atoms	
Non-hydrogen	6711
Water	216
ligands	41
R.m.s.deviations	
Bond (Å)	0.009
Angles (°)	1.14
Ramachandran plot	
Residues in	
favourable (%)	96
allowed (%)	2.4
outliers (%)	1.6
PDB code	**4QYS**

The obtained structure of the ssTrpB2a dimer with external aldimine between PLP and OPS at one subunit and internal aldimine between PLP and Lys111 at the other subunit is depicted in Figure 34.

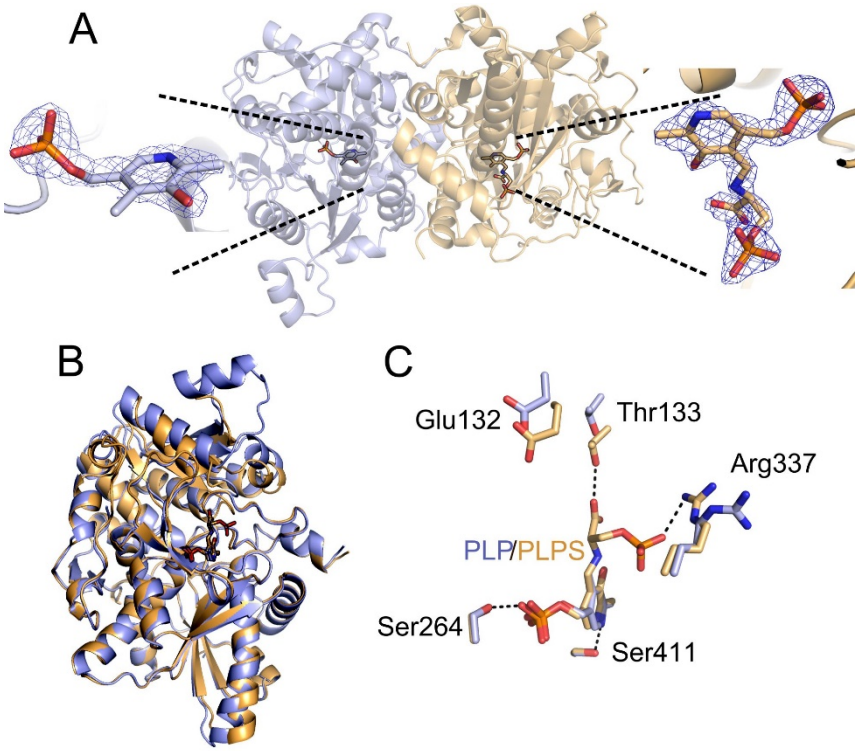


Figure 34: Structure of ssTrpB2a with bound OPS.
(**A**) Crystal structure of the ssTrpB2a homodimer (PDB ID: 4QYS). The subunit with bound PLP is blue- colored, the subunit with the external aldimine between PLP and OPS (PLPS) is yellow-colored. Both subunits are shown as ribbon diagrams, and the ligands are shown as sticks. Electron densities for PLP (left panel) and PLPS (right panel) are indicated by a refined 2Fo-Fc map, contoured at 0.586 e/ Å^3 (2.01 rmsd). (**B**) The PLP-bound subunit with open conformation and the PLPS-bound subunit with closed conformation are superimposed. (**C**) View of the active sites of ssTrpB2a with bound PLP and PLPS. Side chains are shown as balls and sticks, and H-bonds are indicated by black dashes.

The binding of OPS to the active site of ssTrpB2a leads to a conformational change from an open to a closed state. Such a conformational change upon substrate binding was also observed for stTrpB1 (Dunn et al., 2008). As predicted by homology modeling, the cofactor within the active site in the crystal structure is coordinated with hydrogen bonding interactions to Thr133, Ser264 and Ser411. Binding of OPS leads to a reorientation of Thr133 and Arg337, which facilitates the coordination of the carbonyl and the phosphate groups of PLPS. The role of the arginine in coordinating the phosphate group of bound O-phospho-L-serine was further investigated by its mutagenesis to aspartate in ssTrpB2a and atTrpB2o.

2.4.9 Cloning of TrpB2 variants

ss*trp*B2a-R337D was generated by QuikChange site-directed mutagenesis from pET28a(+)-ss*trp*B2a using the oligonucleotides TAT GCA GGT GGG CTA GAT TAT CAT GGA GTA GCC/ GGC TAC TCC ATG ATA ATC TAG CCC ACC TGC ATA.

at*trp*B2o-R350D was generated by QuikChange site-directed mutagenesis from pET28a(+)-at*trp*B2a using the oligonucleotides TGC CAT CCC ATG GTA ATC TAA TCC ACC GGC ATG/ CAT GCC GGT GGA TTA GAT TAC CAT GGG ATG GCA.

2.4.10 Expression and purification of TrpB2 variants

The proteins ssTrpB2a-R337D and atTrpB2o-R350D were expressed and purified like the wild-type proteins. Finally, they were dialyzed against 50 mM potassium phosphate pH 7.5. The proteins were at least 95 % pure as judged by SDS-PAGE (Figure 35).

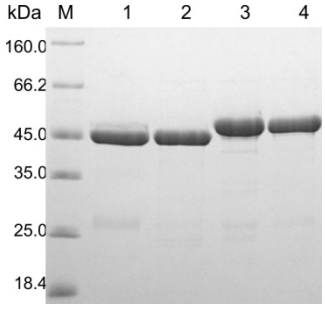


Figure 35: Purity of TrpB2 variants.
SDS-PAGE (12.5 % polyacrylamide) of the purified TrpB proteins. Applied were (**M**) protein ladder and marker (LMW), (**1**) 15 µl ssTrpB2a (10 µM monomer concentration), (**2**) 15 µl ssTrpB2a-R337D (10 µM monomer concentration), (**3**) 15 µl atTrpB2o (10 µM monomer concentration), (**4**) 15 µl atTrpB2o-R350D (10 µM monomer concentration).

2.4.11 Binding properties of TrpB2 variants

Changes in substrate binding due to the replacement of Arg by Asp were analyzed by fluorescence titration. For the TrpB2 wild-type proteins, a significant change in fluorescence was only observed in the presence of OPS. In contrast, a significant change in fluorescence in the presence of L-serine was only observed for the Arg to Asp variants ssTrpB2a-R337D and atTrpB2o-R350D (Figure 36).

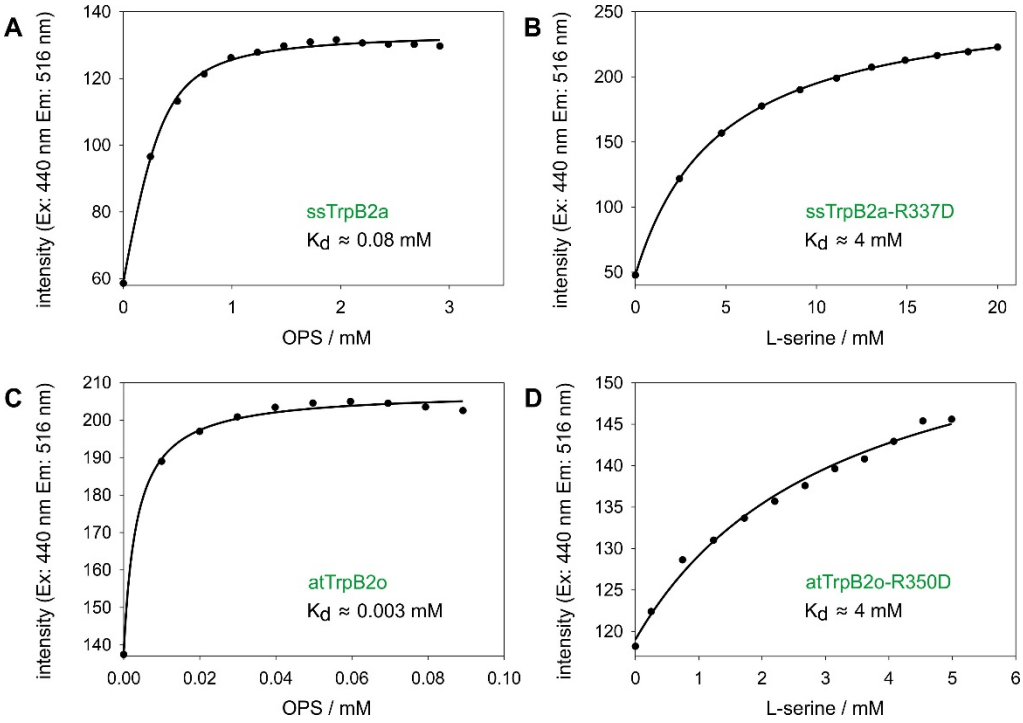


Figure 36: Binding properties of TrpB2 and TrpB2 variants.
Binding of L-serine and OPS was followed by fluorescence detection at 520 nm after excitation at 440 nm. The proteins (**A**) ssTrpB2a, (**B**) ssTrpB2a-R337D, (**C**) atTrpB2o and (**D**) atTrpB2o-R350D with a subunit concentration of 1 µM were titrated with amino acid at 25° C in 100 mM potassium phosphate pH 7.5.

The replacement of Arg by Asp using site-directed mutagenesis results in the inversion of substrate binding from OPS to L-serine as indicated by fluorescence titration experiments. This confirms the significant role of Arg in TrpB2 enzymes in specifying for OPS as substrate.

2.4.12 Divergent evolution in the cystein synthase family

The tryptophan synthase family contains two subgroups (TrpB1 and TrpB2), which differ in their substrate specificities (L-serine and OPS) but catalyse the synthesis of the same product. A similar case was found for the cysteine synthase family CysM/CysK. It shares the same fold with the tryptophan synthase enzyme family and also catalyzes a β-replacement reaction via the same α-aminoacryl intermediate (Burkhard et al., 1999). Recent investigations revealed that some CysK2 and CysM cysteine synthases use OPS instead of O-acetyl-L-serine as substrate (Agren et al., 2008; Mino and Ishikawa, 2003; Nakamura et al., 2012; Oda et al., 2005). Like in TrpB2 enzymes, the coordination of the phosphate leaving group seems to be accomplished by an arginine residue in CysK2 enzymes (Oda et al.,

2005). In the organism *Mycobacterium tuberculosis*, even three cystein synthases were identified, which differ in their substrate specificity (Figure 37).

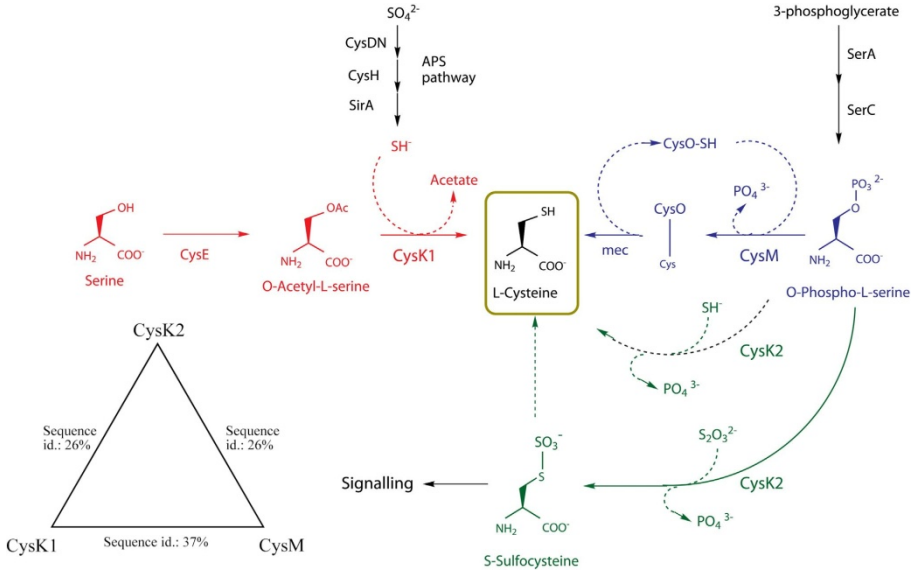


Figure 37: Divergence of substrate specificity in the cysteine synthase family.
The cysteine synthases CysM, CysK1 and CysK2 share a sequence identity of ~30 %. Their substrate specificities and catalysed reactions are indicated. Figure from (Schnell et al., 2014).

A selective advantage for maintaining several homologous cysteine synthases may be a better fine-regulation of the metabolism and the connection of one particular metabolic pathway with multiple other pathways.

2.5 Conclusion

The ability of TrpB2 to synthesize tryptophan *in vivo* was recently analyzed using the hyperthermophile *Thermococcus kodakaraensis,* which can be genetically manipulated (Sato et al., 2005; Sato et al., 2003). *T. kodakaraensis* has a tkTrpB1 enzyme, which is part of the permanent αββα tryptophan synthase complex and a tkTrpB2o enzyme that does not interact with the α-subunit. A *T. kodakaraensis* Δ*trp*B1 strain was created, which lacks TrpB1 activity and has an insufficient TrpA activity due to the missing activation by the binding partner. However, Δ*trp*B1 thrived equally well as the wild-type strain in minimal medium supplemented with indole in oder to compensate for the lacking TrpA1 activation (Hiyama et al., 2014). The natural source of indole for TrpB2o enzymes is unknown up to date.

The transcription of some *trp*B2 genes was shown to be similarly regulated as *trp*B1 (Hiyama et al., 2014; Karr et al., 2008). These findings suggest that TrpB2 enzymes act as tryptophan synthases *in vivo* and primordially evolved for the

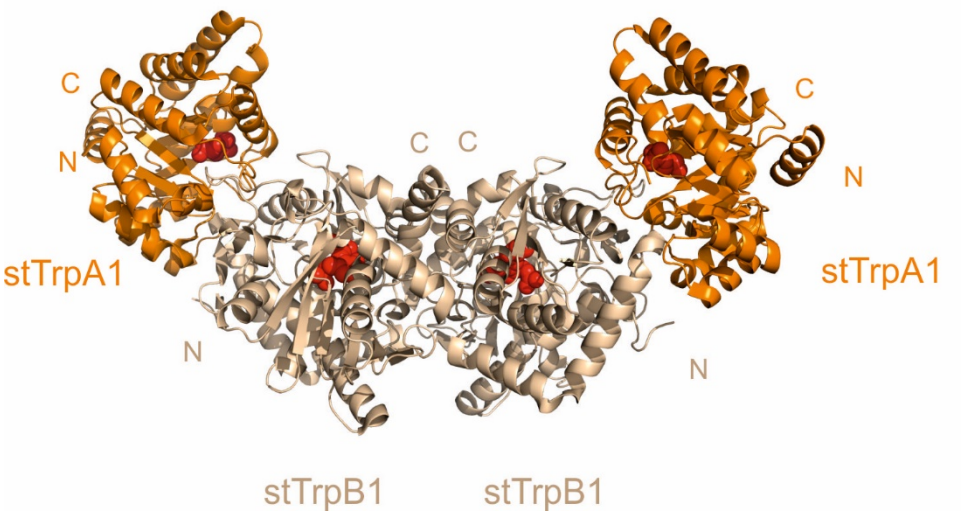
catalysis of the last step in tryptophan biosynthesis. Interestingly, several organisms harbouring TrpB2i or TrpB1 enzymes still possess TrpB2o or TrpB2a proteins. The presence of two TrpB2 enzymes seems to be a minor advantage as indicated by the uneven distribution of TrpB2a in closely related *Sulfolobales* (TrpB2a is present in *S. solfataricus* and *S. tokodaii* but absent in *S. acidocaldarius*). In contrast, differences of TrpB1 and TrpB2 in substrate specificity and affinity may provide a selective advantage by helping to regulate the intracellular concentration of free indole. The specific reason is unclear, however it is known that indole is involved in various biological processes like biofilm formation (Hu et al., 2010), cell cycle control (Chant and Summers, 2007; Chattoraj, 2007) and regulation of gene expression (Mueller et al., 2009; Wang et al., 2001).

3 THE QUATERNARY STRUCTURE OF TRANSIENT TS

3.1 Introduction

3.1.1 Structure of the permanent TS

The enterobacterial TS is a permanent αββα heterotetrameric complex that is known for its sophisticated allosteric communication and regulation. It consists of two α-subunits (TrpA1) and a dimer of β-subunits (TrpB1). The structure of the TS from *S. typhimurium* was solved more than two decades ago (Hyde et al., 1988). It displays a nearly linear arrangement of the subunits and has a length of approximately 150 Å (Figure 38).



stTrpA1 (α-subunit of the permanent TS):

MERYENLFAQLNDRREGAFVPFVTLGDPGIEQSLKIIDTLIDAGADALELGVPFSDPLADGPTIQNANLRAFAAGVTPAQCFEMLALIREKHPTIPIGLLMYANL
VFNNGIDAFYARCEQVGVDSVLVADVPVEESAPFRQAALRHNIAPIFICPPNADDDLLRQVASYGRGYTYLLSRSGVTGAENRGALPLHHLIEKLKEYHAA
PALQGFGISSPEQVSAAVRAGAAGAISGSAIVKIIEKNLASPKQMLAELRSFVSAMKAASRA

stTrpB1 (β-subunit of the permanent TS):

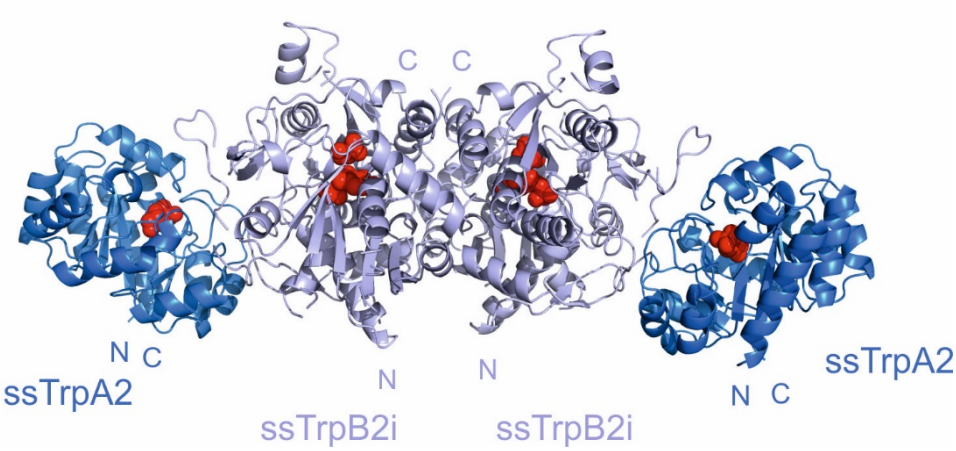
MTTLLNPYFGEFGGMYVPQILMPALNQLEEAFVSAQKDPEFQAQFADLLKNYAGRPTALTKCQNITAGTRTTLYLKREDLLHGGAHKTNQVLGQALLAKRMGKS
EIIAETGAGQHGVASALASALLGLKCRIYMGAKDVERQSPNVFRMRLMGAEVIPVHSGSATLKDACNEALRDWSGSYETAHYMLGTAAGPHPYPTIVREFQRMIGEE
TKAQILDKEGRLPDAVIACVGGGSNAIGMFADFINDTSVGLIGVEPGGHGIETGEHGAPLKHGRVGIYFGMKAPMMQTADGQIEESYSISAGLDFPSVGPQHAYLN
SIGRADYVSITDDEALEAFKTLCRHEGIIPALESSHALAHALKMMREQPEKEQLLVVNLSGRGDKDIFTVHDILKARGEI

Figure 38: Quaternary structure of the permanent TS.
The structure of the *S. typhimurium* TS (2J9X) is depicted as ribbon diagram. The α-subunits are orange-colored and the β-dimer is brown-colored. The bound ligands GP and L-serine are depicted as red spheres. The amino acid sequences of the α- and β-subunits are indicated. Residues located within 4 Å of the hetero-oligomer interface are highlighted by size and color.

Several parts of both subunits contribute to the protein-protein interface. The relative orientation of the subunits results in the interconnection of the αβ-active sites. This interconnection allows for the channeling of indole during the conversion of IGP to tryptophan (Huang et al., 2001; Raboni et al., 2005). The catalytic efficiencies of both subunits are affected by complex formation and allosteric regulation within the TS.

3.1.2 Structure of the transient TS

Recently, the TrpA1 and TrpB1 homologs ssTrpA2 and ssTrpB2i were identified as α- and β-subunits of a TS in the Crenarchaeon *S. solfataricus*. The interaction between the subunits is ligand-dependent, and the observed αββ stoichiometry testifies to a pronounced negative cooperativity for binding of the second α-subunit (Ehrmann et al., 2010). Moreover, the quaternary structure of the *S. solfataricus* TS strikingly differs from that of the *S. typhimurium* TS (Figure 39).



ssTrpA2 (α-subunit of the transient TS):

MEMGKMLVVYMTLGYPNVQSFKDFIIGAVENGADILELGIPPK**YAKY**DG**P**VI**R**KSYDKVKGLDIWPLIEDIRKDVGVPIIALTYLEDWVDQLENFLNMIKDVKLDGILFPD**L**
LIDY**ID**DLDKIDGIIKNKGLKNVIFTS**PS**V**PD**L**L**IHKVSKISDLFLYYGVR**PT**TG**VPIPVS**VK**QL**IN**R**VRNLVENKLIVGFGLSSESDLRDALSAGADGIAIGTVFI
EEIERNGVKSAINLVKKFRAILDEYK

ssTrpB2 (β-subunit of the transient TS):

MVKEDEILPKYWYNI**IP**D**L**P**KPLPP**P**RDPQGA**YF**S**R**I**D**LL**R**SILPKEV**LR**QQ**FTIE**RYIKIPEEVRDRYLSIGRPTPLFRAKRLEEYLKTPARIYFKYEGAT
PTGSHKINTAIPQAYFAKEEGIEHVVTETGAGQWGTAVALAASMYNMKSTIFMVKVSYEQKPMRRSIMQLYGANVYASPTNLTEYGRKILETNPQHPGSLGIAMSEAIEYALKN
EFRYLVGSVLDVVLLHQSVIGQETITQLDLLGEDADILIGCVGGGSNFGGFTYPFIGNKKGKRYIAVSSAEIPKFSKGEYKYDFPDSAGLLPLVKMITLGKDYVPPPIYAGGLRYH
GVAPTLSLLTKEGIVEWREYNEREIFEAAKIFIENQGIVPAPESAHAIRAVVDEAIEARKNNERKVIVFNLSGHGLLDLSNYESMMKRLNGNG

Figure 39: Quaternary structure of the transient TS.
The structure of the *S. solfataricus* TS (O. Mayans, unpublished data) is depicted as ribbon diagram. The α-subunits are blue-colored and the β-dimer is light blue-colored. The bound ligands GP and L-serine are depicted as red spheres. The amino acid sequences of the α- and β-subunits are indicated. Residues located within 4 Å of the heterooligomer interface are highlighted by size and color.

As a consequence of the high protein concentration used for crystallization, the obtained structure displays a non-physiological αββα stoichiometry. Several parts of the α-subunit, but only one N-terminal part of the β-subunit contribute to the protein-protein interface. The residues within the interface are not sufficiently well resolved to gain insights into the interactions between single residues at the protein-protein interaction surface.

3.1.3 The non-interacting enzyme ssTrpB2a

S. solfataricus also contains an additional ssTrpB2a, which is encoded by a gene located outside of the *trp* operon. It shares a sequence identity of ~54 % with ssTrpB2i. In contrast to ssTrpB2i, ssTrpB2a does not bind to ssTrpA2 under any applied experimental conditions (Ehrmann et al., 2010; Leopoldseder et al., 2006). It is not obvious whether a lack of selective pressure on ssTrpB2a had prevented its evolution to become part of a transient TS or whether an interaction with ssTrpA2 was lost in the course of evolution.

3.2 Remarks

Parts of this chapter had been published equal worded in (Busch et al., 2014). This chapter also contains data from Dietmar Dewitz (Dewitz, 2013), Patricia Seidel (Seidel, 2012) and Michael Schupfner (Schupfner, 2011; Schupfner, 2014). Details are indicated in the following text. This data was included in this work to draw as complete a picture as possible on transient TS.

3.3 Significance of this work

In the first part of this work, the interaction between ssTrpA2 and ssTrpB2i was investigated in more detail. This investigation revealed that the transient TS is formed with equal propensity in the presence of different ssTrpB2i ligands and that the transfer of indole within the transient TS is similarly efficient as the transfer within the permanent TS.

Up to date, solely the inside of the *trp* operon encoded ssTrpB2i was identified as part of a transient TS. In the second part of this work, several non-*trp* operon encoded TrpB2 enzymes were identified as candidates for forming a transient TS

based on the observable co-evolution in the family of α-subunits. Indeed, the co-evolved enzymes tvTrpA2 and tvTrpB2a from *T. volcanium* were shown to form a weak, ligand-dependent TS. Guided by the structure of the transient ssTrpA2-ssTrpB2i complex from *S. solfataricus*, the N-terminal part of TrpB2 enzymes was identified to determine the interaction with TrpA2. This was validated by transferring the N-terminal sequence stretches stemming from ssTrpB2i or tvTrpB2a to ssTrpB2a, which cannot interact with either ssTrpA2 or tvTrpA2. Dependent on the origin of the N-terminal stretch, the resulting chimeric proteins gained the ability to interact with either ssTrpA2 or tvTrpA2. This indicates that the ability and specificity for a protein-protein interaction can be readily transferred between those proteins.

3.4 Results and discussion

3.4.1 Cloning of subunits of a transient and a permanent TS

ss*trp*A2 cloning and insertion into pET28a(+) has been described in (Leopoldseder et al., 2006).

ss*trp*B2i was amplified by overlap extension PCR to remove the internal XhoI restriction site and the stop codon at the 3' end of the gene from pET28a(+)-ss*trp*B2i (Leopoldseder et al., 2006) using the oligonucleotides GTG CTG CTG ATG TAC ATA TGG TAA AAG AAG AC/ ATC GAT TCT GGA GAA ATA GGC ACC and GGT GCC TAT TTC TCC AGA ATC GAT/ CCG CTC GAG CCC ATT TCC ATT. The final amplification product was cloned into pET21a(+) at the NdeI/ XhoI restriction sites.

st*trp*A1 was amplified from synpBR322-st*trp*AB by PCR using the oligonucleotides AGC CAT ATG GAA CGC TAC GAA AAT TTA/ GTG GTG CAA GCT TAT GCG CGG CTG GCG GCT TTC. The amplification product was inserted into pET28a(+) at the NdeI/ HindIII restriction sites.

st*trp*B1 cloning and insertion into pET24a(+) has been described in Chapter 2.

3.4.2 Expression and purification of subunits of a transient and a permanent TS

The protein ssTrpA2 was purified by a heat step and metal chelate affinity chromatography as described (Hettwer and Sterner, 2002; Leopoldseder et al., 2006).

For expression of ssTrpB2i, *Escherichia coli* CodonPlus (DE3) RIPL cells were transformed with pET21a(+)-ss*trp*B2i. The cells were grown at 37° C in LB-medium with 20 µM PLP, 30 µg/ ml chloramphenicol, and 150 µg/ ml ampicillin overnight in the absence of IPTG. Cells were harvested by centrifugation, resuspended in 25 ml/ l culture 10 mM potassium phosphate pH 7.5, 300 mM KCl, 20 µM PLP and 10 mM imidazole, and disrupted by sonication. After a subsequent heat step (30 min at 70° C), the His_6-tagged proteins were purified by metal chelate affinity chromatography. Proteins in 10 mM potassium phosphate pH 7.5 and 300 mM KCl were eluted by a linear gradient of imidazole (10-1000 mM). Fractions containing sufficiently pure protein were dialyzed against 10 mM potassium phosphate pH 7.5. The protein was further purified by ion exchange chromatography as described

(Leopoldseder et al., 2006) and dialyzed against 100 mM potassium phosphate pH 7.5.

For expression of stTrpA1, *E. coli* BL21 Rosetta was transformed with pET28a(+)-st*trp*A1. The cells were grown at 37° C in LB with 50 µg/ ml kanamycin to a density of OD_{600}= 0.5. Protein expression was induced by the addition of 0.5 mM IPTG. After growth over night at 20° C, cells were harvested by centrifugation, resuspended in 25 ml/ l culture 50 mM Tris/HCl pH 7.5, 150 mM NaCl and 10 mM imidazole, and disrupted by sonication. The His_6-tagged proteins were purified by metal chelate affinity chromatography. Proteins in 50 mM Tris/HCl pH 7.5 and 150 mM NaCl were eluted by a linear gradient of imidazole (10- 500 mM). The proteins were dialyzed against 50 mM Tris/HCl pH 7.5. The subunits were approximately 95 % pure as judged by SDS-PAGE (Figure 40).

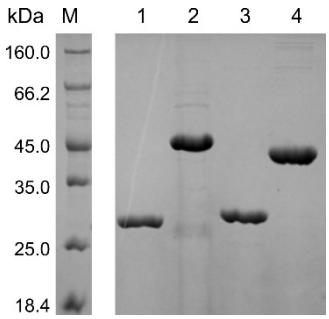


Figure 40: Purity of *S. solfataricus* and *S. typhimurium* TS subunits.
SDS-PAGE (12.5 % polyacrylamide) of the purified TS subunits. Applied were (**M**) protein ladder and marker (LMW), (**1**) 15 µl ssTrpA2 (10 µM monomer concentration), (**2**) 15 µl ssTrpB2i (10 µM monomer concentration), (**3**) 15 µl stTrpA1 (10 µM monomer concentration), (**4**) 15 µl stTrpB1 (10 µM monomer concentration).

3.4.3 Influence of different ligands on the transient TS

It has been shown that the transient ssTrpA2-ssTrpB2i TS is formed in the presence of the ssTrpB2i ligand L-serine and various ssTrpA2 ligands that mimic the substrate, the transition state or the product of the α-subunit (Ehrmann et al., 2010). In order to determine the requirements of a ssTrpB2i ligand for the interaction with ssTrpA2, SPR measurements were undertaken in the presence of GP and the ssTrpB2i ligands L-serine, O-phospho-L-serine, glycine, or O-sulfo-L-serine (Figure 41).

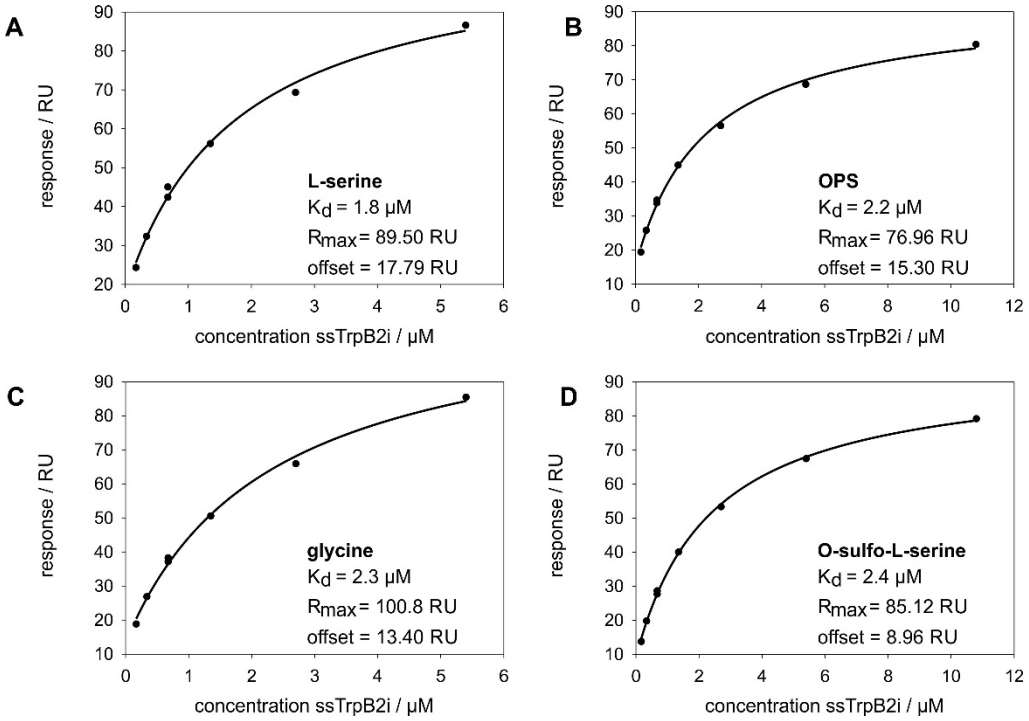


Figure 41: Ligand-dependency of the TrpA2-TrpB2i interaction.
SPR measurements for the interaction between ssTrpA2 and ssTrpB2i in the presence of different ligands. ssTrpA2 was covalently immobilized on flow cell 2 of a CM5 sensor chip using EDC/NHS chemistry. The ligand in 10 mM sodium acetate buffer pH 4.83, was injected to obtain a final signal of ~220 response units (RU). Interactions were measured at 25° C at a flow rate of 30 µl/ min using various concentrations of ssTrpB2i as analyte in HBS-EP+, 60 mM GP, and 0.05 % (w/v) sodium azide in the presence of (**A**) 1 M L-serine, (**B**) 10 mM OPS, (**C**) 1 M L-serine or (**D**) 10 mM O-sulfo-L-serine. The binding surface was regenerated after each injection with HBS-EP+. K_d values were determined by using a steady-state binding model.

The affinity between ssTrpA2 and ssTrpB2i is identical, independent of the used ssTrpB2i ligand. These results indicate that the ssTrpA2-ssTrpB2i complex is formed with equal propensity, no matter of whether the β-substituent of the ssTrpB2i ligand is –H, $-CH_2OH$, $-CH_2OSO_3^-$ or $-CH_2OPO_3^{2-}$. The K_d for the transient complex formed in the presence of GP and O-phospho-L-serine was confirmed by ITC (Figure 42).

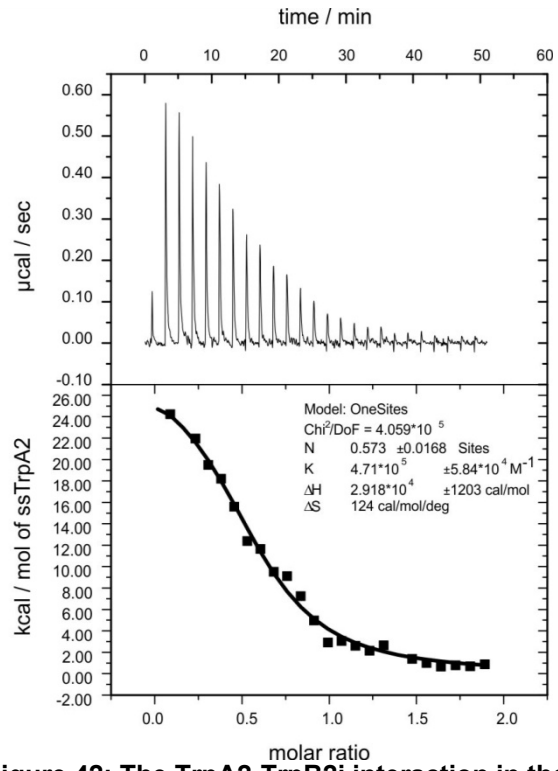


Figure 42: The TrpA2-TrpB2i interaction in the presence of GP and OPS.
ITC measurement for the interaction between ssTrpA2 and ssTrpB2i in the presence of GP and OPS. Prior to the measurement, the proteins were dialyzed against HBS-EP+, 60 mM GP, 10 mM OPS, and 0.05 % (w/v) sodium azide. The titration was conducted at 25° C by injecting 24 aliquots of 1.6 µl containing 188 µM ssTrpA2 into 200 µl containing ssTrpB2i at a subunit concentration of 21 µM. The thermodynamic parameters were calculated by a one site binding model implemented in the ITC Origin software.

The association constant K is the inverse of the dissociation constant K_d and is proportional to the free enthalpy ΔG (ΔG= -R·T·lnK). The stoichiometry N of 0.57 indicates the formation of a αββ TS. The K_d value for this TS is 2.1 µM and the ΔG value is -32 kJ/ mol. The thermodynamic parameters equals those determined for the interaction in the presence of L-serine and GP (N= 0.59, K_d= 1.8 µM, ΔG= -32 kJ/ mol; (Ehrmann, 2011)).

The data indicate that changes at the β-substituent during the TS reaction have no influence on complex stability. Most plausible, the complex remains stable throughout the whole catalytic cycle until tryptophan is released from the β-subunit.

3.4.4 Channeling within the transient TS

The activity of the α-subunit is enhanced by the liganded β-subunit within the *S. solfataricus* TS. This ensures that IGP is only cleaved to a significant amount when formed indole can be subsequently converted to tryptophan. The X-ray structure of the *S. solfataricus* TS indicates that the active sites of the α-subunit and

the β-subunit are interconnected by a hydrophobic tunnel. This interconnection might allow the channeling of intermediary formed indole (Figure 43).

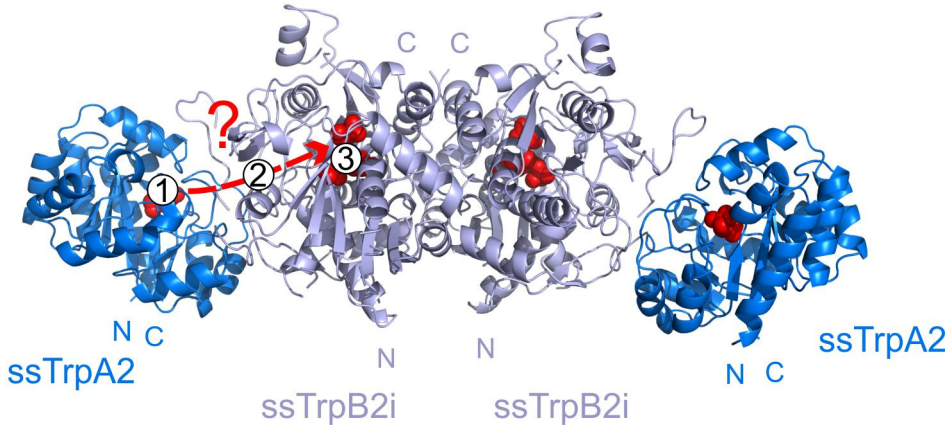


Figure 43: Putative pathway of indole within the TrpA2-TrpB2i type TS.
The structure of the *S. solfataricus* TS (O. Mayans, unpublished data) is depicted as ribbon diagram. The α-subunits are blue-colored and the β-dimer is light blue-colored. The bound ligands GP and L-serine are depicted as red spheres. The putative way of indole is indicated by a red arrow. The numbers refer to the entry point of indole (**1**), the putative path of indole within the enzyme complex (**2**) and the point of indole-to-tryptophan conversion (**3**).

An interconnection is characterized by a defined entry point at the α-subunit active site (1), a tight tunnel between the active sites (2) and a defined exit point at the β-subunit active site (3). Such being the case, the interconnection was probed regarding these three aspects: First, the accessibility of free indole to the β-active site was determined in dependency of the α-subunit ligand GP (blocking approach). Second, the amount of intermediary formed indole that escapes the complex was quantified (capturing approach). Third, the amount of indole was determined, which is converted to tryptophan by the associated β-subunits (competition approach).

Blocking approach

In tightly connected TS, external indole has to pass through the active site of the α-subunit to reach the active site of the β-subunit and to be subsequently converted to tryptophan. In the *S. typhimurium* TS, GP is able to block the active site of the α-subunit and thus prevents the conversion of external indole to tryptophan (Ahmed et al., 1991; Dunn et al., 1990). The effect of GP on tryptophan synthesis was validated for the *S. typhimurium* TS and was also determined for the enzymes from *S. solfataricus*. The results are depicted in Figure 44.

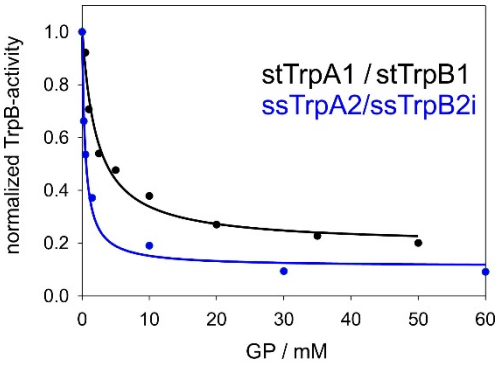


Figure 44: Effect of GP on the synthesis of tryptophan.
The TrpB-reaction was measured with 2 µM TrpB-protein and 4 µM TrpA-protein in 100 mM KP pH 7.5, 180 mM KCl, 40 µM PLP, 100 µM indole, 200 mM L-serine, and various concentrations of GP at 25° C (stTrpA1/stTrpB1) or at 60° C (ssTrpA2/ssTrpB2i).

GP binding to the *S. typhimurium* TS strongly compromises the synthesis of tryptophan from free indole. For the enzymes from *S. solfataricus*, GP binding induces TS formation and also compromises the synthesis of tryptophan by free indole. Most plausible, the connection between the active sites is the preferred route to the β-site for indole in both TS. However, it cannot be ruled out that tryptophan synthesis is limited by a different mechanism in the *S. solfataricus* TS.

Capturing approach

In tightly connected TS, indole formed at the active site of the α-subunit is not a free intermediate and does not diffuse and thus get trapped in an overlaid organic solvent. Such being the case, no indole is trapped by toluene during the conversion of IGP to tryptophan by the *Neurospora crassa* TS (Yanofsky and Rachmeler, 1958). A similar approach was used to probe for the presence of free diffusible indole during the conversion of IGP to tryptophan by the *S. typhimurium* TS and the *S. solfataricus* TS. Hexane was used as organic solvent as it has a low water solubility (0.0003 % w/w) and is able to solubilize indole. The applicability of this approach was validated by monitoring the amount of indole that is trapped during the enzymatic cleavage of IGP. Subsequently, the amount of indole being trapped during the *S. typhimurium* TS reaction and during the *S. solfataricus* TS reaction was quantified (Table 8).

Table 8: Detection of indole released during the TS reaction.

A	**indole in aqueous phase/ µM**	**indole in organic phase/ µM**	**Σ indole/ µM**
5 min	1.05 ± 0.13	2.99 ± 0.25	4.04
10 min	3.44 ± 0.57	4.14 ± 0.39	7.58
30 min	11.75 ± 0.13	7.67 ± 0.54	19.42
B	**tryptophan in aqueous phase/ µM**	**indole in organic phase/ µM**	**ratio tryptophan/ indole**
5 min	113.83 ± 1.40	0.26 ± 0.10	438
10 min	220.91 ± 1.26	0.47 ± 0.04	470
30 min	348.44 ± 3.91	0.79 ± 0.03	441
80 min	325.89 ± 30.63	0.71 ± 0.09	459
C	**tryptophan in aqueous phase/ µM**	**indole in organic phase/ µM**	**ratio tryptophan/ indole**
5 min	18.23 ± 4.65	0.07 ± 0.04	260
10 min	34.77 ± 6.55	0.18 ± 0.01	193
30 min	134.79 ± 3.64	0.62 ± 0.01	217
80 min	399.95 ± 12.76	1.08 ± 0.12	370

Reactions were performed in 150 µl volumes with 100 mM EPPS/KOH, pH 7.5, 180 mM KCl, 40 µM PLP, and 500 µM IGP with (**A**) 0.1 µM stTrpB1 and 0.1 µM stTrpA1, (**B**) 0.1 µM stTrpB1, 0.1 µM stTrpA1 and 200 mM L-serine and (**C**) with 20 µM ssTrpB2i, 20 µM ssTrpA2 and 200 mM L-serine. The reactions in 24 well plates were overlaid with 1.5 ml hexane and incubated at 40° C in a water bath for the indicated times. Indole in the organic phase was quantified spectroscopically after derivatization with p-dimethylaminocinnamaldehyde (DMACA). Indole and tryptophan in the aqueous phase were quantified by HPLC using program 1. The concentration of the two tryptophan synthases differ by a factor of 200 in order to compensate for the different enzymatic activities at the given temperature.

The *S. typhimurium* TS has a weak aldolase activity in the absence of L-serine. Free indole is equally distributed in the aqueous and the organic phase as diffusion does not result in a thermodynamic equilibrium distribution within the observed time scale. Such being the case, only half the amount of released indole can be detected in the organic phase. Taking this into account, approximately 0.5 % of indole occurs as free intermediate during the *S. typhimurium* TS reaction. Consistently, Anderson et al. observed that less than 1 % of intermediately formed indole is released during the conversion of IGP to L-tryptophan by the *S. typhimurium* TS (Anderson et al., 1991). Similarly, just about 1 % of formed indole is released in the S*. solfataricus* TS reaction. This low quantity of free indole indicates a rapid and efficient conversion of indole to tryptophan in the *S. typhimurium* TS and in the *S. solfataricus* TS.

Competition approach

In tightly connected TS, indole formed at the active site of the α-subunit remains bound within the enzyme complex till it is converted to tryptophan by the associated β-subunit. Consequently, another indole-utilizing enzyme should not compete with

the β-subunit for this indole. In this approach, the different substrate specificities of TrpB1 and TrpB2 were used to determine the amount of indole that remains in the TS and is thus exclusively converted by the associated β-subunit. The experimental setup for the permanent TS is illustrated in Figure 45.

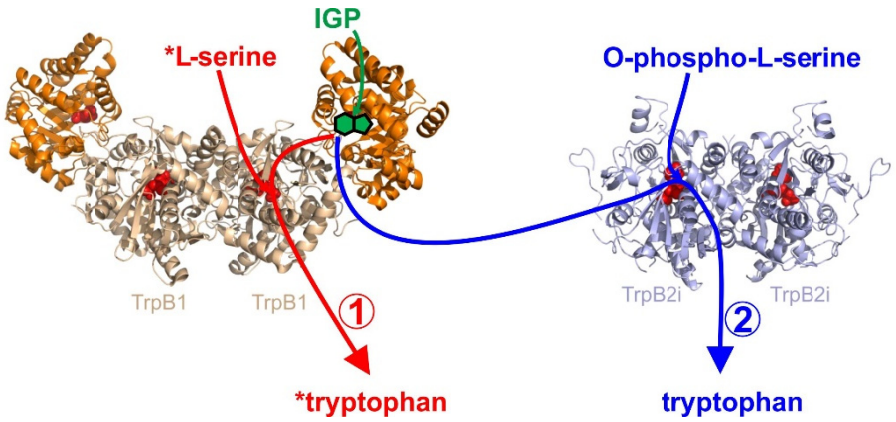


Figure 45: Illustration of the competition approach.
Competition approach to probe for indole channeling in the permanent tryptophan synthase complex. Added IGP is cleaved by the α-subunit and indole is formed. (**1**) Channeled indole is converted to ^{14}C-labeled tryptophan (*tryptophan) by the associated β-subunit TrpB1, whereas (**2**) indole that escapes the TS is converted to non-^{14}C-labeled tryptophan (tryptophan) by TrpB2i. The ratio of non-^{14}C-labeled tryptophan to ^{14}C-labeled tryptophan indicates whether indole is channeled within the permanent TS.

The difference in substrate specificity of stTrpB1 and ssTrpB2i is important to match the ratio of non-^{14}C-labeled tryptophan to ^{14}C-labeled tryptophan with the amount of indole, which is converted by each enzyme. So in a first instance, substrate specificities of stTrpB1 and ssTrpB2i were analyzed by TLC (Figure 46)

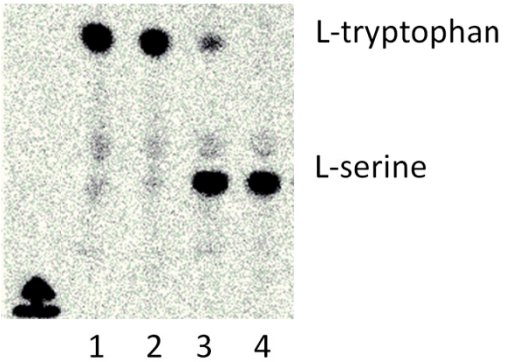


Figure 46: Substrate specificities of stTrpB1 and ssTrpB2i.
TLC detection for the conversion of ^{14}C-labeled L-serine to ^{14}C-labeled tryptophan. 4 µM stTrpB1 were incubated in 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP with 200 µM indole and 200 µM ^{14}C-labeled L-serine (**1**) in the absence of OPS or (**2**) in the presence of 100 mM OPS.
1.8 µM ssTrpB2i were incubated in 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP with 200 µM indole and 200 µM ^{14}C-labeled L-serine (**3**) in the absence of OPS or (**4**) in the presence of 100 mM OPS. All reactions were incubated at 40° C for 1 h. 1 µl of each reaction was separated on a silica 60 plate with concentration zone. The mobile phase was 80 ($^{v}/_{v}$) n-butanol: 20 ($^{v}/_{v}$) glacial acetic acid: 20 ($^{v}/_{v}$) water. ^{14}C-labeled L-serine and L-tryptophan was detected by phosphorimaging. The origin of spots is indicated by the black marker on the left, the R_f-value for L-tryptophan was 0.71 and the R_f-value for L-serine was 0.35.

In the presence of ^{14}C-L-serine and OPS, stTrpB1 uses ^{14}C -serine for the synthesis of ^{14}C-tryptophan and ssTrpB2i uses OPS for the synthesis of tryptophan. The substrate preference is sufficiently high to distinguish, whether indole is converted to tryptophan by stTrpB1 or by ssTrpB2i.

To determine the amount of ^{14}C-labeled and non-^{14}C-labeled tryptophan in a mixture, they were separated by HPLC followed by absorbance detection and by an online scintillation counter. The absorbance parallels the total amount of tryptophan and the radio-signal parallels the amount of radioactive labeling. The ratio of UV signal to radio-signal was initially determined for labeled tryptophan, which had been synthesized by the *S. typhimurium* TS. The HPLC profile is shown in Figure 47.

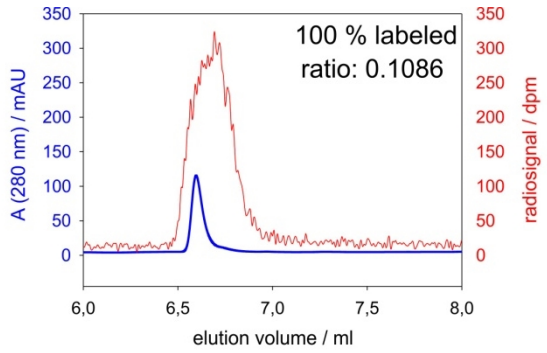


Figure 47: HPLC analysis of ^{14}C-labeled tryptophan.
Elution profile of ^{14}C-labeled tryptophan followed by absorbance detection at 280 nm (blue trace) and radio-detection (red trace). A 10 µl reaction containing 0.1 µM stTrpB1, 0.1 µM stTrpA1, 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 500 µM IGP, 300 µM ^{14}C-L-serine (3.4 mCi/ mmol) and 10 mM OPS was incubated at 40° C. The reaction was quenched by the addition of 20 µl methanol. 20 µl of the mixture were analyzed by HPLC using program 4.

Reaction mixtures with the TS and a TrpB-enzyme with different substrate specificity were analyzed at several points of time and the ratio of ^{14}C-labeled tryptophan to non-^{14}C-labeled tryptophan was determined. The results are summarized in Table 9.

Table 9: Quantification of synthesized tryptophan.

A TS from *S. typhimurium* and ssTrpB2i			**L-serine*, OPS and IGP**	
	A 280 nm	**radiosignal**	**A 280 nm/ radiosignal**	**% labeled tryptophan**
5 min	217.64	1958.97	0.11	98
10 min	379.25	3242.79	0.12	93
30 min	852.58	6977.98	0.12	89

B TS from *S. solfataricus* and stTrpB1			**L-serine*, OPS and IGP**	
	A 280 nm	**radiosignal**	**A 280 nm/ radiosignal**	**% labeled tryptophan**
5 min	845.23	1158.06	0.73	15
10 min	1822.90	1606.00	1.14	10
30 min	2384.24	1865.39	1.28	8

(**A**) 10 µl reactions containing 0.1 µM stTrpB1, 0.1 µM stTrpA1, 50 µM ssTrpB2i, 100 mM EPPS/KOH, pH 7.5, 180 mM KCl, 40 µM PLP, 500 µM IGP, 300 µM ^{14}C-L-serine (3.4 mCi/mmol) and 10 mM OPS were incubated at 40° C. (**B**) 10 µl reactions containing 20 µM ssTrpB2i, 20 µM ssTrpA2, 20 µM stTrpB1, 100 mM EPPS/KOH, pH 7.5, 180 mM KCl, 40 µM PLP, 500 µM IGP, 300 µM ^{14}C-L-serine (3.4 mCi/mmol) and 10 mM OPS were incubated at 40° C. The reactions were quenched at the indicated time points by the addition of 20 µl methanol. 20 µl of the mixture were analyzed by HPLC using program 4. Elution was followed by absorbance detection at 280 nm and radio-detection.

In both TS, indole formed at the α-subunit is mainly converted to tryptophan at the associated β-subunit. Thus the *S. typhimurium* TS and *S. solfataricus* TS are similar efficient in transferring intermediary formed indole to the active site of the associated β-subunit.

In summary, both TS prevent the emergence of free indole, which is not subsequently converted to tryptophan.

Common features in distant related complexes with different quaternary structures indicate a high selective pressure for developing these features at an early stage of evolution. The main driving force in the evolution of TS seems to be the prevention of free indole. This might have been important to prevent cross talks with indole-dependent cell signaling pathways and to prevent indole from diffusing through the cell membrane.

3.4.5 Phylogenetic analysis of the TrpA enzyme family

Are just TrpB2i enzymes or also other TrpB2 enzymes part of a transient TS? To answer this question, putative transient TS were identified by the analysis of the TrpA enzyme family under the premise that protein-protein interactions are the result of protein co-evolution (Pazos and Valencia, 2008). A sequence similarity network (SSN) of InterPro family IPR002028 (v.48) was calculated (Atkinson et al., 2009; Gerlt et al., 2015). It is depicted in Figure 48.

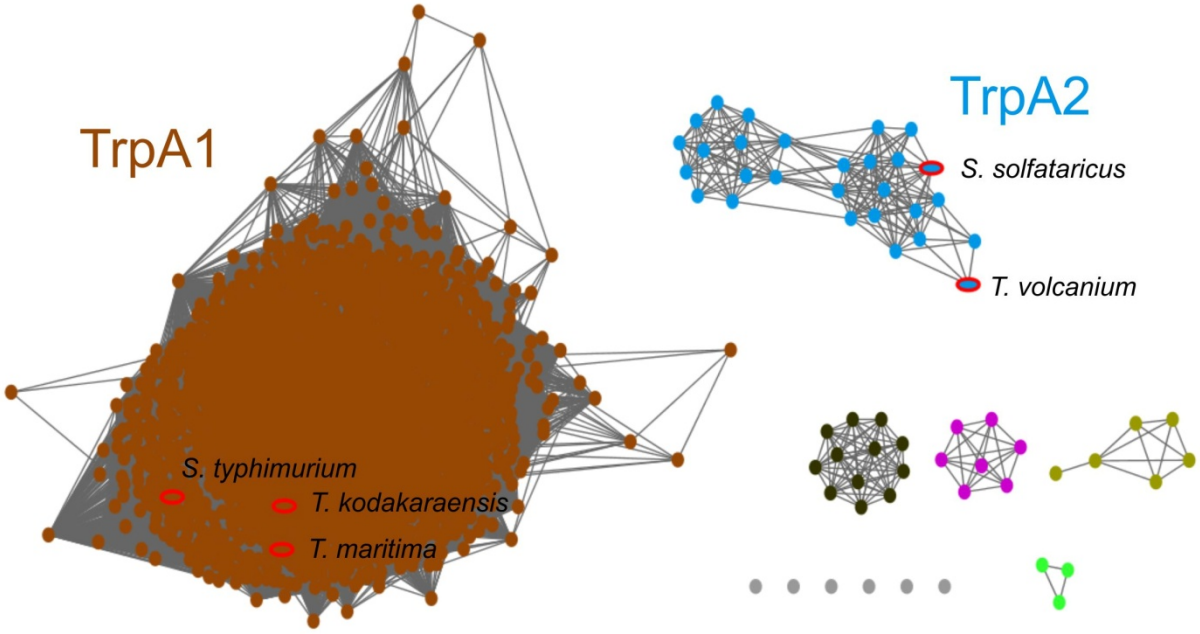


Figure 48: SSN of the TrpA enzyme family.
Enzymes with a sequence identity of >80 % are depicted as single knots (rep-note: 80). Knots that share a sequence identity of >36 % are connected. The two major groups were named TrpA1 and TrpA2 based on experimental data for some of the containing members (TrpA1 from *S. typhimurium*, *T. kodakaraensis*, *T. maritima* and TrpA2 from *S. solfataricus*).

Interestingly, TrpA2 enzymes also occur in organisms that lack a TrpB2i enzyme. As such, the organism *Thermoplasma volcanium* contains only a TrpB2a enzyme, which is encoded by a gene outside of the *trp* operon. The apparent co-evolution asked for probing the interaction between tvTrpA2 and tvTrpB2a. The protein tvTrpA2 contains a solvent exposed cysteine at position 195, which was mutated to serine to prevent the formation of artificial dimers. The resulting tvTrpA2-C195S variant is named tvTrpA2* throughout the text. The cloning, expression and purification of tvTrpA2* and tvTrpB2a had also been described before (Dewitz, 2013).

3.4.6 Cloning of tvTrpA2* and tvTrpB2a

tv*trp*A2 was amplified by PCR from genomic DNA of *T. volcanium* strain GSS1 using the oligonucleotides ATT AGG ATC CAT GAA GCC ATT TGT ATA TTT CAC TCT CG/ ATA TAG TCG ACT TAG TGA TGA TGA TGA TGA TGG GAT GCC CCC AAA ATA TCA T and inserted into pMAL-c2(+) at the SalI/ BamHI restriction sites. QuikChange Mutagenesis with the oligonucleotides GGC ATA CGT AAC AAG TCT ACG ATA GCT AAG ATC/ GAT CTT AGC TAT CGT AGA CTT GTT ACG TAT GCC was subsequently performed to obtain pMAL-c2(+)-tv*trp*A2*.

tv*trp*B2a was amplified by PCR from genomic DNA of *T. volcanium* strain GSS1. The yield of amplification product was increased by two subsequent PCR reactions. In the first round of PCR, the oligonucleotides GGG ATG GTA GAT TCT AAA GA/ AAA AAC GAT AGT AGC GGT TG were used. The amplification product was used as template in a second round of PCR with the oligonucleotides ACG CCA TAT GAT AAG AAT CGA TCT AAA GCA AGA CG/ GTG CTC GAG TTC AAA ATG CAT TGC CTC TGC A and inserted into pET24a(+) at the NdeI/ XhoI restriction sites.

3.4.7 Expression and purification of tvTrpA2* and tvTrpB2a

For expression of MBP-tvTrpA2*, *E.* coli T7 Express cells were transformed with pMAL-c2(+)-tv*trp*A2*. The cells were grown at 37° C in LB with 20 mM potassium phosphate, pH 7.5 and 150 µg/ ml ampicillin to OD_{600}= 0.5. Protein expression was induced by addition of 0.5 mM IPTG. After growth overnight at 13° C, cells were harvested by centrifugation, resuspended in 25 ml/ l culture 100 mM potassium phosphate pH 7.5 and 500 mM KCl and disrupted by sonication. The supernatant was loaded on a Ni^{2+}-chelate affinity column at a flow rate of 0.5 ml/ min, which had been equilibrated with 100 mM potassium phosphate pH 7.5 and 500 mM KCl. Bound proteins were incubated with 50 µg/ ml trypsin. After on-column cleavage for 2 h at room temperature, maltose binding protein (MBP) and trypsin were washed from the column. Elution of bound tvTrpA2* was performed at room temperature with a linear gradient of imidazole (10- 1000 mM). Fractions containing sufficiently pure protein were dialyzed against 100 mM potassium phosphate pH 7.5.

For expression of tvTrpB2a, *E.* coli T7 Express Rosetta was transformed with pET24a(+)-tv*trp*B2a. The cells were grown at 37°C in sorbitol betaine medium with 20 µM PLP and 75 µg/ ml kanamycin to OD_{600}= 0.5. Protein expression was induced

by addition of 0.5 mM IPTG. After growth overnight at 37° C, cells were harvested by centrifugation, resuspended in 25 ml/ l 50 mM potassium phosphate pH 7.5 and 300 mM KCl, disrupted by sonication, and subjected to a heat step (75° C for 20 min). The His_6-tagged proteins were purified by metal chelate affinity chromatography. Proteins in 50 mM potassium phosphate pH 7.5 and 300 mM KCl were eluted by a linear gradient of imidazole (0- 1000 mM). Fractions containing sufficiently pure protein were dialyzed against 100 mM potassium phosphate pH 7.5. The proteins were approximately 90 % pure as judged by SDS-PAGE (Figure 49).

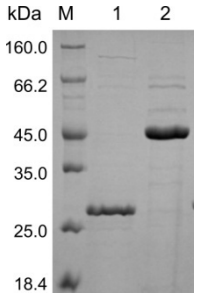


Figure 49: Purity of tvTrpA2* and tvTrpB2.
SDS-PAGE (12.5 % polyacrylamide) of the purified proteins. Applied were (**M**) protein ladder and marker (LMW), (**1**) 15 µl tvTrpA2* (10 µM monomer concentration) and (**2**) 15 µl tvTrpB2a (10 µM monomer concentration).

3.4.8 Activity of tvTrpA2* and tvTrpB2a

The turnover numbers and the Michaelis constants of the isolated tvTrpA2* and tvTrpB2a proteins were determined by steady-state enzyme kinetics (Figure 50).

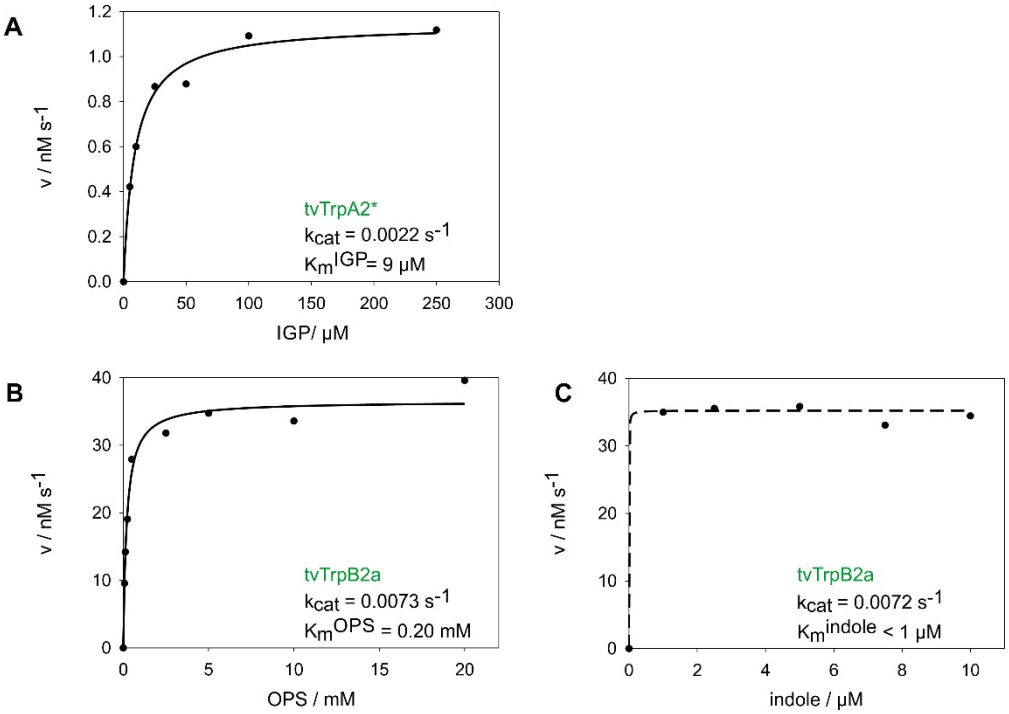


Figure 50: Enzymatic parameters of tvTrpA2* and tvTrpB2a.
(**A**) IGP-dependent tvTrpA2*-reaction. The reaction was started by the addition of IGP at 40° C. The reaction contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 6 mM NAD^+, 20 mM arsenate, 5.5 µM GAPDH and 0.5 µM tvTrpA2*. The conversion of IGP to indole was determined by reversed-phase HPLC using program 3. (**B**) OPS-dependent tvTrpB2a-reaction. The reaction was started by the addition of 100 µM indole and followed at 60° C by the absorbance change at 290 nm. The reactions contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, 40 µM PLP, 0.5 µM tvTrpB2a (monomer concentration) and varying amounts of L-serine. (**C**) Indole-dependent tvTrpB2a-reaction. The reaction was started by the addition of indole and followed at 60° C by the absorbance change at 290 nm. The reactions contained 100 mM EPPS/KOH pH 7.5, 180 mM KCl, and 40 µM PLP, 10 mM OPS and 0.5 µM tvTrpB2a β-subunit (monomer concentration). Data points were fitted with a single rectangular hyperbolic function (Michaelis-Menten equation).

The reactions were measured at 40° C due to the heat-lability of tvTrpA2*. The activity of tvTrpA2* was too low to allow the measurement by a spectro-photometer. Instead, reactions were quenched and analyzed by HPLC. The data showed that tvTrpA2* and tvTrpB2a are poor catalysts with very low k_{cat}-values. The K_M^{OPS} of tvTrpB2a in the micromolar range confirms that TrpB2 enzymes are OPS-dependent tryptophan synthases.

3.4.9 Probing the formation of a transient TS

Size-exclusion chromatography was used to study the ligand-dependency of the interaction between tvTrpA2* and tvTrpB2a (Figure 51).

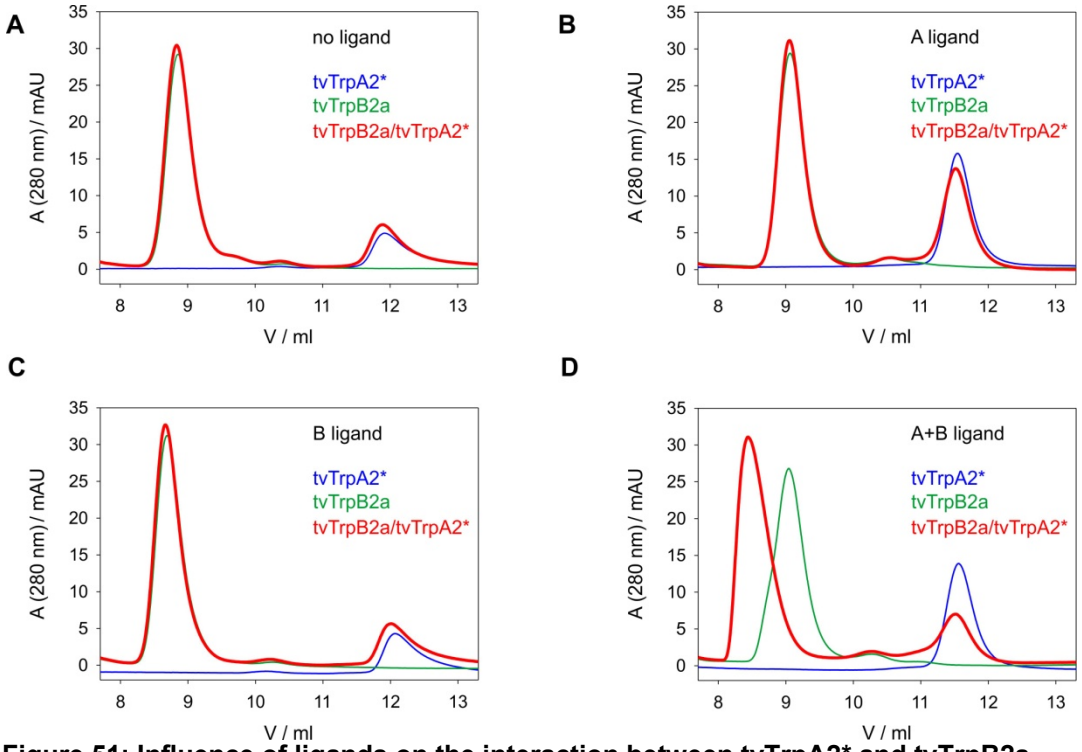


Figure 51: Influence of ligands on the interaction between tvTrpA2* and tvTrpB2a.
Analytical size exclusion chromatograms of 50 µl tvTrpA2*, tvTrpB2a and tvTrpB2a/tvTrpA2* mixture (7 µM, monomer concentration). Samples were applied on a Superdex S75 column equilibrated with 10 mM potassium phosphate pH 7.5, 100 mM KCl, 0.05 % sodium azide (**A**) in the absence of ligand, (**B**) in the presence of 60 mM GP, (**C**) in the presence of 1 M L-serine and (**C**) in the presence of 60 mM GP and 1 M L-serine. Elution was performed at 25° C with a flow rate of 0.35 ml/ min and followed by measuring the absorbance at 280 nm.

Complex formation between tvTrpA2* and tvTrpB2a was only detectable in the presence of the tvTrpA2* ligand GP and the tvTrpB2a ligand L-serine. A ligand-dependency for TS formation was also observed for the proteins ssTrpA2 and ssTrpB2i from *S. solfataricus* (Ehrmann et al., 2010) and thus seems to be a general feature of TrpA2-TrpB2 type complexes. The tvTrpA2*-tvTrpB2a interaction was quantified by SPR in the presence of GP and L-serine as well as in the presence of GP and OPS (Figure 52).

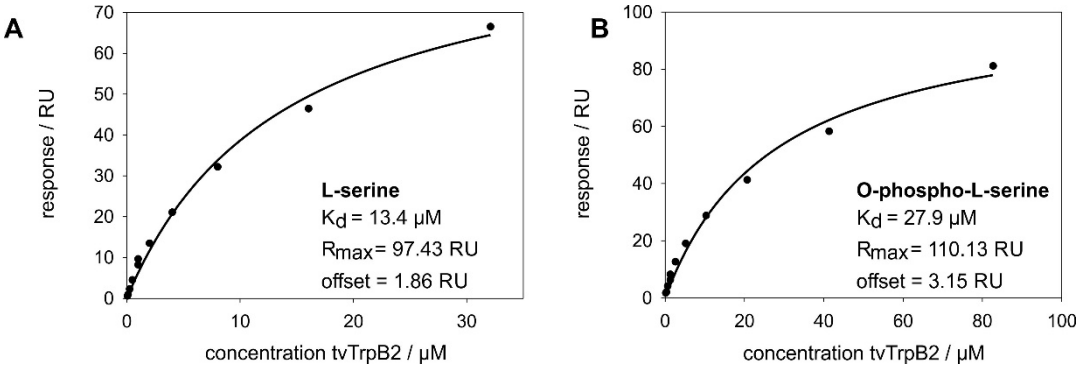


Figure 52: Quantification of the tvTrpA2*-tvTrpB2a interaction.
SPR measurements for the quantification of the tvTrpA2*-tvTrpB2a interaction. The tvTrpA2* protein was covalently immobilized on flow cell 2 of a CM5 sensor chip using EDC/NHS chemistry. 10 µg/ ml protein in 10 mM sodium acetate pH 4.92 were injected to obtain a signal for bound protein of ~60 response units (RU) and a final signal of ~200 response units (RU). Interactions were measured at 25° C at a flow rate of 30 µl/ min using various concentrations of tvTrpB2-protein as analyte in HBS-EP+, 60 mM GP, and 0.05 % (w/v) sodium azide in the presence of (**A**) 1 M L-serine or (**B**) 10 mM OPS. The binding surface was regenerated after each injection with HBS-EP+. Thermodynamic dissociations constants (K_d values) were determined by using a steady-state binding model. Data is from (Schupfner, 2014).

SPR measurements showed that the interaction between tvTrpA2* and tvTrpB2a is very weak with a K_d in the two-digit micromolar range. The weak interaction might indicate a relaxed selection pressure for the evolution or for the maintainance of this protein-protein interaction. However, the weak interaction might provide high specificity *in vivo*.

3.4.10 Specificity of the subunit interaction within transient TS

The specificity of TrpA2 enzymes for their interaction with the co-evolved partner was confirmed by the inability of ssTrpB2a to interact with ssTrpA2 or with tvTrpA2* (Figure 53).

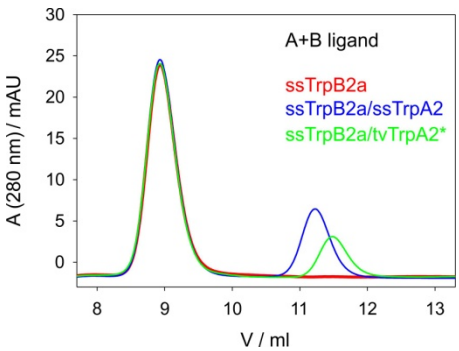


Figure 53: ssTrpB2a does not interact with ssTrpA2 or tvTrpA2*.
Analytical size exclusion chromatograms of 30 µl ssTrpB2a, ssTrpB2a/ssTrpA2 and ssTrpB2a/tvTrpA2* (8 µM, monomer concentrations). Samples were applied on a Superdex S75 column being equilibrated with 10 mM potassium phosphate, pH 7.5, 100 mM KCl, 0.05 % sodium azide, 60 mM GP, 1 M L-serine. Elution was performed at 25° C with a flow rate of 0.35 ml/ min and followed by measuring the absorbance at 280 nm.

3.4.11 Identification of a common interface within transient TS

The observed differences in complex formation indicate that the quaternary structures of TrpA2-TrpB2 type TS and TrpA1-TrpB1 type TS might differ. Indeed, the recently solved structure of the ssTrpA2-ssTrpB2i complex with bound GP and L-Ser revealed considerable differences in the relative orientation of the α- and β-subunits compared to the stTrpA1-stTrpB1 complex from *S. typhimurium* (O. Mayans, unpublished data). According to the crystal structures, three sequence stretches of stTrpB1 directly interact with stTrpA1, whereas only a single N-terminal sequence stretch of ssTrpB2i directly interacts with ssTrpA2 (see Figure 38 and Figure 39). In order to test whether this stretch is necessary and sufficient for the binding of TrpB2 proteins to TrpA2, the ability to convert non TrpA2-interacting TrpB2 proteins into TrpA2-interacting proteins by exchanging N-terminal residues was probed. For this purpose, ssTrpB2a was chosen, which is structurally similar to ssTrpB2i (O. Mayans, unpublished data) but does not interact with either ssTrpA2 or tvTrpA2* (Figure 53).

In a first step, a sequence alignment of ssTrpB2i, tvTrpB2a, and ssTrpB2a was generated. Subsequently, the N-terminal residues of ssTrpB2a were exchanged with the corresponding residues of ssTrpB2i or tvTrpB2a. In ss/ss-Chimera 1 and ss/ss-Chimera 2, 30 and 37 N-terminal residues of ssTrpB2a were replaced by the corresponding residues of ssTrpB2i. In tv/ss-Chimera 1 and tv/ss-Chimera 2, 8 and 26 N-terminal residues of ssTrpB2a were replaced by the corresponding residues of tvTrpB2a (Figure 54).

Figure 54: MSA of the N-terminal parts of TrpB2 proteins and TrpB2 chimeras.
In the chimeric proteins, positions of ssTrpB2a are exchanged to the ssTrpB2i positions or tvTrpB2a positions as indicated by color.

It was probed whether a re-design of the protein-protein interaction surface by introduction of the correct interface residues induces complex formation of these chimeras with ssTrpA2 or tvTrpA2*. The four chimeras were cloned, produced by heterologous gene expression in *E. coli* and purified by metal chelate chromatography using a C-terminally added His_6-*tag*.

3.4.12 Cloning of chimera TrpB2 proteins

ss/ss-*chimera* 1 was generated by OE-PCR from pET28a(+)-ss*trp*B2i and pET28a(+)-ss*trp*B2a with the oligonucleotides TAA TAC GAC TCA CTA TAG GG/ CTC CAG CAC CTT GCT CGG TAG TAT ACT TCT and AGA AGT ATA CTA CCG AGC AAG GTG CTG GAG/ CCG CAA GCT TCT CCT TAA ATA ACA C and inserted into pET24a(+) at the NdeI/ HindIII restriction sites.

ss/ss-*chimera* 2 was generated by OE-PCR from pET28a(+)-ss*trp*B2i and pET28a(+)-ss*trp*B2a with the oligonucleotides TAA TAC GAC TCA CTA TAG GG/ TTT TAC GTA TCT TTC TAT TGT GAA TTG TTG and CAA CAA TTC ACA ATA GAA AGA TAC GTA AAA/ CCG CAA GCT TCT CCT TAA ATA ACA C and inserted into pET24a(+) at the NdeI/ HindIII restriction sites.

tv/ss-*chimera* 1 was generated by OE-PCR from pET24a(+)-tv*trp*B2a and pET24a(+)-ss*trp*B2a with the oligonucleotides TAA TAC GAC TCA CTA TAG GG/ GGG ATC TTG TGG CGG TGG CAA CTC TTC TGG and CCA GAA GAG TTG CCA CCG CCA CAA GAT CCC/ GCT AGT TAT TGC TCA GCG G and inserted into pET24a(+) at the NdeI/ HindIII restriction sites.

tv/ss-*chimera* 2 was generated by OE-PCR from pET24a(+)-tv*trp*B2a and pET24a(+)-ss*trp*B2a with the oligonucleotides TAA TAC GAC TCA CTA TAG GG/ CGG TCT TCC AAC TTG CAT ATA TTT TTC GTA and TAC GAA AAA TAT ATG

CAA GTT GGA AGA CCG/ GCT AGT TAT TGC TCA GCG G and inserted into pET24a(+) at the NdeI/ HindIII restriction sites.

3.4.13 Expression and purification of chimera TrpB2 proteins

For expression of ss/ss-Chimera 1 and ss/ss-Chimera 2, *E. coli* CodonPlus (DE3) RIPL cells were transformed with pET24a(+)-ss/ss-*chimera* 1, and pET24a(+)-ss/ss-*chimera* 2. The cells were grown at 37°C in LB-medium with 20 µM PLP, 30 µg/ ml chloramphenicol, and 75 µg/ ml kanamycin overnight in the absence of IPTG. Cells were harvested by centrifugation, resuspended in 25 ml/ l culture 10 mM potassium phosphate pH 7.5, 300 mM KCl, 20 µM PLP and 10 mM imidazole, and disrupted by sonication. The His_6-tagged proteins were purified by metal chelate affinity chromatography. Proteins in 10 mM potassium phosphate pH 7.5 and 300 mM KCl were eluted by a linear gradient of imidazole (10- 1000 mM). Fractions containing sufficiently pure protein as judged by SDS-PAGE were dialyzed against 100 mM potassium phosphate pH 7.5 and subjected to a heat step (30 min at 70° C for ss/ss-Chimera 1, 30 min at 60° C for ss/ss-Chimera 2) (Schupfner, 2011).

For expression of tv/ss-Chimera 1 and tv/ss-Chimera 2, *E.* coli CodonPlus (DE3) RIPL cells were transformed with pET24a(+)-tv/ss-*chimera* 1 and pET24a(+)-tv/ss-*chimera* 2. The cells were grown at 37°C in LB with 20 µM PLP and 75 µg/ ml kanamycin to OD_{600}= 0.5. Protein expression was induced by addition of 0.5 mM IPTG. After growth overnight at 37° C, cells were harvested by centrifugation, resuspended in 25 ml/ l 10 mM potassium phosphate pH 7.5, 300 mM KCl, 20 µM PLP and 10 mM imidazole, and disrupted by sonication. The His_6-tagged proteins were purified by metal chelate affinity chromatography. Proteins in 10 mM potassium phosphate pH 7.5 and 300 mM KCl were eluted by a linear gradient of imidazole (10-1000 mM). Fractions containing sufficiently pure protein were dialyzed against 100 mM potassium phosphate pH 7.5 (Schupfner, 2014). The proteins were at least 90 % pure as judged by SDS-PAGE (Figure 55).

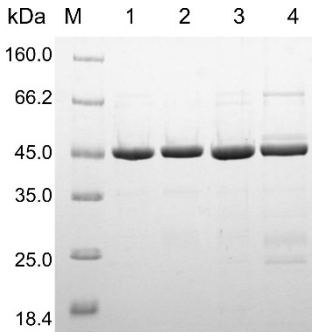


Figure 55: Purity of TrpB2 chimeras.
SDS-PAGE (12.5 % polyacrylamide) of the purified proteins. Applied were (**M**) protein ladder and marker (LMW), (**1**) 10 µl ss/ss-Chimera 1 (12 µM monomer concentration), (**2**) 10 µl ss/ss-Chimera 2 (12 µM monomer concentration), (**3**) 10 µl tv/ss-Chimera 1 (12 µM monomer concentration), and (**4**) 10 µl tv/ss-Chimera 2 (12 µM monomer concentration).

3.4.14 Structural integrity, thermal stability, and activity of chimera TrpB2 proteins

Highly symmetrical size exclusion chromatography profiles demonstrated that all chimeras are homogenous proteins. Their structural integrity was further confirmed by far-UV circular dichroism (Figure 56).

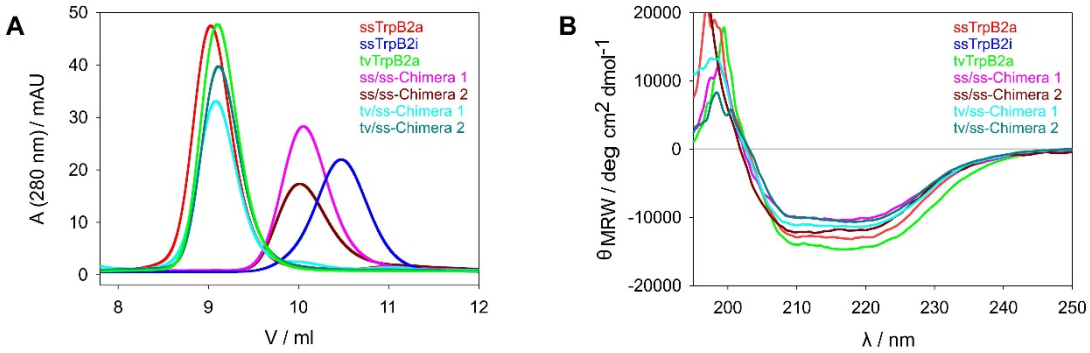


Figure 56: Structural integrity of TrpB2 chimeras.
(**A**) Size exclusion chromatography profiles. 50 µl of 10 µM protein were applied to a Superdex 75 column being equilibrated with 100 mM potassium phosphate pH 7.5, 300 mM KCl (15 µM monomer concentration). Elution was performed at 25° C with a flow rate of 0.5 ml/ min, and followed by measuring the absorbance at 280 nm. (**B**) Far-UV circular dichroism (CD) spectra of 7 µM protein (monomer concentration) in 100 mM potassium phosphate pH 7.5. Spectra were recorded at 25° C using a spectro-polarimeter (d= 1 mm). Data is from (Schupfner, 2014) and (Seidel, 2012).

All chimeras display elution times between the wild-type proteins ssTrpB2a and ssTrpB2i, which had been shown to be dimers (Leopoldseder et al., 2006). The CD-spectra of the TrpB proteins are influenced by the cofactor, which interferes with an accurate measurement of ellipticity below ~205 nm.

The conformational stability of the proteins was tested by thermal melting measured by CD (Figure 57).

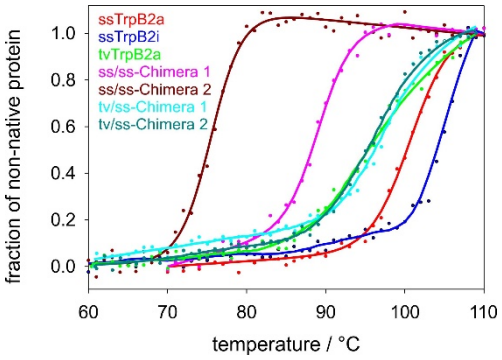


Figure 57: Thermal stability of TrpB2 chimeras.
Thermal denaturation followed by CD-spectroscopy. The loss of ellipticity of 7 µM of protein (monomer concentration) in 100 mM potassium phosphate, pH 7.5 was monitored from 60° C to 110° C with a rate of 1° C/ min. The apparent T_m values are 100° C for ssTrpB2a, > 100° C for ssTrpB2i, 93° C for tvTrpB2a, 89° C for ss/ss-Chimera 1, 75° C for ss/ss-Chimera 2, 98° C for tv/ss-Chimera 1, and 97° C for tv/ss-Chimera 2. Data was fitted by a two-state model as far as possible. Data is from (Schupfner, 2014) and (Seidel, 2012).

The unfolding transitions were cooperative and yielded apparent T_m-values between 73° C for ss/ss-Chimera 2 and about 95° C for ss/ss-Chimera 1, tv/ss-Chimera 1, and tv/ss-Chimera 2.

The catalytic activity of the chimeras was analyzed by steady-state enzyme kinetics. The deduced $K_M^{L\text{-serine}}$ and k_{cat} values are depicted in Figure 58.

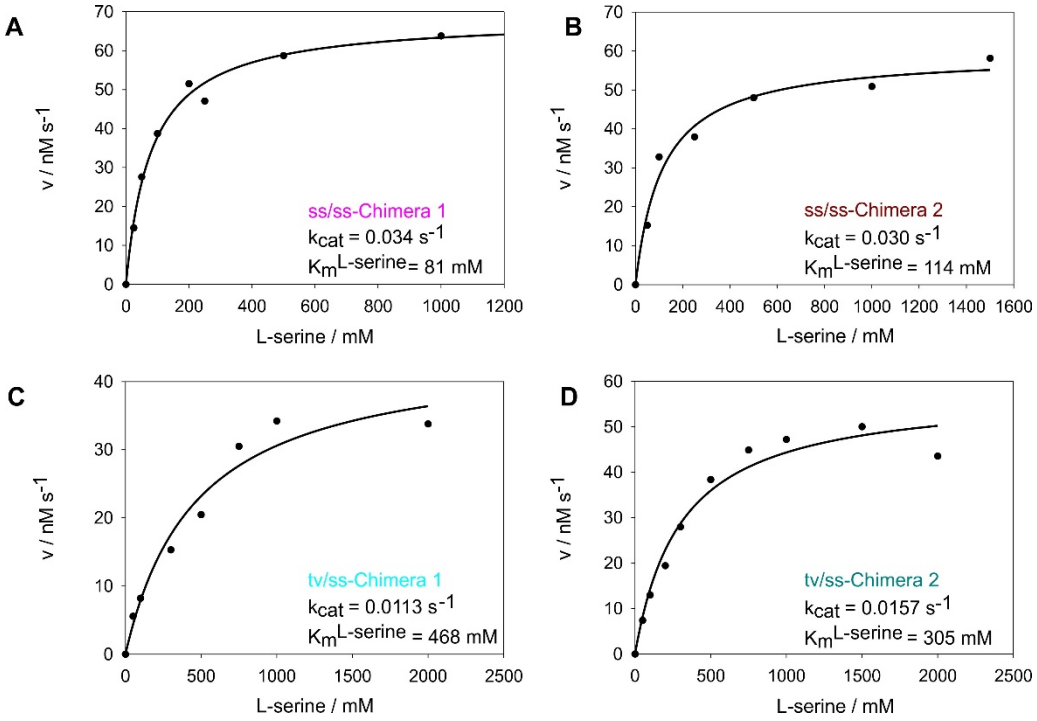


Figure 58: Enzymatic parameter of TrpB2 chimeras.
L-serine-dependent TrpB2 reactions. The reactions were started by the addition of 100 µM indole and followed at 60° C by the absorbance change at 290 nm. The reactions contained 100 mM potassium phosphate pH 7.5, 180 mM KCl, 40 µM PLP, varying concentrations of L-serine and (**A**) 2 µM ss/ss-Chimera 1, (**B**) 2 µM ss/ss-Chimera 2, (**C**) 3 µM tv/ss-Chimera 1, or (**D**) 3 µM tv/ss-Chimera 2 . Data for tv/ss-Chimera 1 and tv/ss-Chimera 2 is from (Schupfner, 2014).

In summary, the data show that all produced chimeras are properly folded and active, which proves that changes at the N-terminus do not perturb the overall structure or catalysis.

3.4.15 Analysis of chimera TrpB2 proteins for their ability to form a transient TS

Surface plasmon resonance (SPR) experiments were performed to assess whether the chimeras can bind to ssTrpA2 and tvTrpA2*. No interaction could be detected between ss/ss-Chimera 1 and ssTrpA2, and between tv/ss-Chimera 1 and tvTrpA2*, respectively. However, the formation of ss/ss-Chimera 2-ssTrpA2 and tv/ss-Chimera 2-tvTrpA2* complexes could be detected (Figure 59).

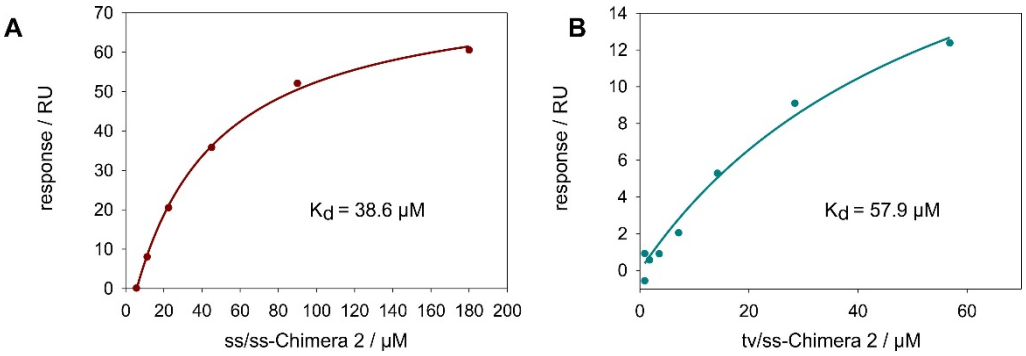


Figure 59: TrpA2-TrpB2 chimera interaction detected by SPR.
SPR measurements for the quantification of the TrpA2-Chimera interaction. SPR measurements were performed on a Biacore X100 optical biosensor. The TrpA2 protein was covalently immobilized on flow cell 2 of a CM5 sensor chip using EDC/NHS chemistry. (**A**)10 µg/ ml ssTrpA2 in 10 mM sodium acetate, pH 4.83 and (**B**) 10 µg/ ml tvTrpA2* in 10 mM sodium acetate, pH 4.92 were injected to obtain a signal for bound protein of ~60 response units (RU) and a final signal of ~200 response units (RU). Interactions were measured at 25° C at a flow rate of 30 µl/ min using various concentrations of TrpB2-chimera as analyte in HBS-EP+, 60 mM GP, and 0.05 % (w/v) sodium azide in the presence of 10 mM OPS. The binding surface was regenerated after each injection with HBS-EP+. Thermodynamic dissociations constants (K_d values) were determined by using a steady-state binding model. Data for tv/ss-chimera 2 is from (Schupfner, 2014).

Activity titrations further showed that the enzymatic activity of ssTrpA2 is enhanced by ss/ss-Chimera 2 to the same extent as by ssTrpB2i. Likewise, the enzymatic activity of tvTrpA2* is enhanced by tv/ss-Chimera 2, albeit to a lower extent compared to tvTrpB2a (Figure 60).

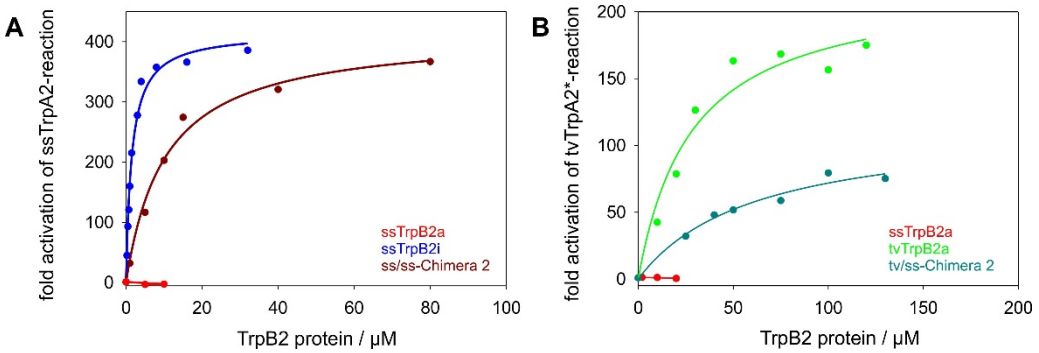


Figure 60: Effect of TrpA2-TrpB2 chimera interaction on the TrpA2 activity.
(**A**) ssTrpA2 or (**B**) tvTrpA2* was titrated with increasing concentrations of TrpB2 proteins or Chimeras 2 in 100 mM EPPS/KOH, pH 7.5, 180 mM KCl, 40 µM PLP, 6 mM NAD^+, 20 mM arsenate, 1 M L-serine, 5.5 µM glyceraldehyde-3-phosphate dehydrogenase. The reaction was started by the addition of 100 µM IGP and followed at 60° C (activation of ssTrpA2-reaction) or 40° C (activation of tvTrpA2*-reaction) by the absorbance change at 340 nm caused by the production of NADH. Initial velocities were plotted as a function of added TrpB2, and the binding affinities were determined by fitting the data points with a hyperbolic function. Data for the activation of the tvTrpA2* reaction is from (Schupfner, 2014).

The apparent K_d values of the activity titrations in the double digit micromolar range are similar to the values determined by the SPR measurements (Table 10).

Table 10: Affinities as detected by SPR and activity titration.

protein-protein interaction		K_d^{SPR} (µM)	$K_d^{titration}$ (µM)
ssTrpB2i	**ssTrpA2**	2.4	1.4
ss/ss-Chimera 2	**ssTrpA2**	38.6	10.8
tvTrpB2a	**tvTrpA2***	27.9	28.4
tv/ss-Chimera 2	**tvTrpA2***	57.9	61.6

Reaction conditions for the SPR measurements at 25 °C: HBS-EP+ buffer with 0.05% sodium azide and 60 mM GP in the presence of 10 mM OPS. Reaction conditions for the activity titration of TrpA: 100 mM EPPS/KOH, pH 7.5, 180 mM KCl, 40 µM PLP, 6 mM NAD^+, 20 mM arsenate, 1 M L-serine, 5.5 µM tmGAPDH. The reaction was started by the addition of 100 µM IGP and followed at 60 °C (ssTrpA2) or 40 °C (tvTrpA2*) by the absorbance change at 340 nm caused by the production of NADH.

These data show that the N-terminal parts of ssTrpB2i and tvTrpB2a are sufficient for the formation of a TS with the activation of the corresponding α-subunits (ssTrpA2* or tvTrpA2*).

3.5 Conclusion

Whereas in many cases the quaternary structure of protein complexes can be inferred with the assumption that interfaces of homologous subunits are highly conserved (Zhang et al., 2010), distantly related homologues may differ in their interaction geometrically although they have the same *in vivo* function (Russell et al., 2004). Geometrically different interfaces were for example found between the two

signaling proteins CheY and CheA P2 from *E. coli* and of *Thermotoga maritima*. Whereas similar regions are involved in the formation of the CheY-CheA P2 complex, the relative orientation of the subunits differ in these complexes (Park et al., 2004). Moreover, different quaternary structures have been published for the three anthranilate synthase complexes from *S. typhimurium*, *Serratia marcescens* and *S. solfataricus*, although they all consist of homologous TrpG and TrpE subunits (Knöchel et al., 1999; Morollo and Eck, 2001; Spraggon et al., 2001). TS provide another case of quaternary structure plasticity as different regions of TrpB1 and TrpB2, as well as of TrpA1 and TrpA2 are involved in complex formation, which enable elaborate allosteric communication during catalysis. Permanent TrpA1-TrpB1 complexes as found in *E. coli* and *S. typhimurium* are characterized by a hydrogen-bond network between residues of the α- and β-subunits. Allosteric communication is mainly transmitted by the interaction of two loops of the α-subunit with the COMM domain of the β-subunit, which forms the basis for the bi-directional activation of the subunits within the complex and enables the transition between an open and closed state (Osborne et al., 2003; Pan and Dunn, 1996). Contrary, transient TrpA2-TrpB2 complexes as found in *S. solfataricus* and *T. volcanium* are only formed during catalysis and have a quaternary structure that differs from that of the TrpA1-TrpB1 complexes. The transient TS remains formed during the conversion of IGP to tryptophan.

Like in permanent TS, the activity of the α-subunit is enhanced and indole is channeled to the active site of the β-subunit in transient TS. Substrate channeling is also found in several non-related enzyme complexes like the carbamoyl phosphate synthetase and the glutamine phosphoribosylpyrophosphate amidotransferase (Huang et al., 2001). Thus it is a common strategy to prevent the free diffusion and/ or the hydrolysis of an intermediately formed metabolite.

Crystal structure analysis revealed that within the TrpA2-TrpB2 complex of *S. solfataricus*, solely an N-terminal part of the β-subunit interacts with the α-subunit (O. Mayans, unpublished data). This finding indicates that the N-terminal part of the β-subunit might be a suitable module for the design of artificial ligand-dependent protein-protein interactions. The creation of chimeric proteins is a general strategy to combine functional modules. This strategy was for instance used to design a novel DNA-binding protein, which consists of two domains of the endonucleases I-DmoI and I-CreI (Chevalier et al., 2002; Kortemme and Baker, 2004). Furthermore, it was

used to create proteins, which consist of an interacting domain and an effector domain of different signaling pathways. Fusing the phospho-tyrosine recognition domain Grb2 or ShcA to the death effector domain of Fadd was so shown to lead to receptor tyrosine kinase mediated apoptosis (Howard et al., 2003). In the present work, the N-terminus of TrpA2-interacting TrpB2 proteins was transferred to a non TrpA2-interacting TrpB2a protein. This transfer is sufficient to enable the formation of ligand-dependent complexes and the regulation of TrpA2 activity within this complex. This finding demonstrates that homologous proteins can be recruited to a certain metabolic pathway by the transfer of an interaction module.

IMPLICATION FOR TS EVOLUTION

A model for the evolution of TrpB was proposed by Merkl (2007) taking the following criteria into account: First TrpB1 and TrpB2 have evolved only once as indicated by conserved indels, second TrpB2 is distributed in all three domains of life and third TrpB1 does not occur in several Archaea (Figure 61).

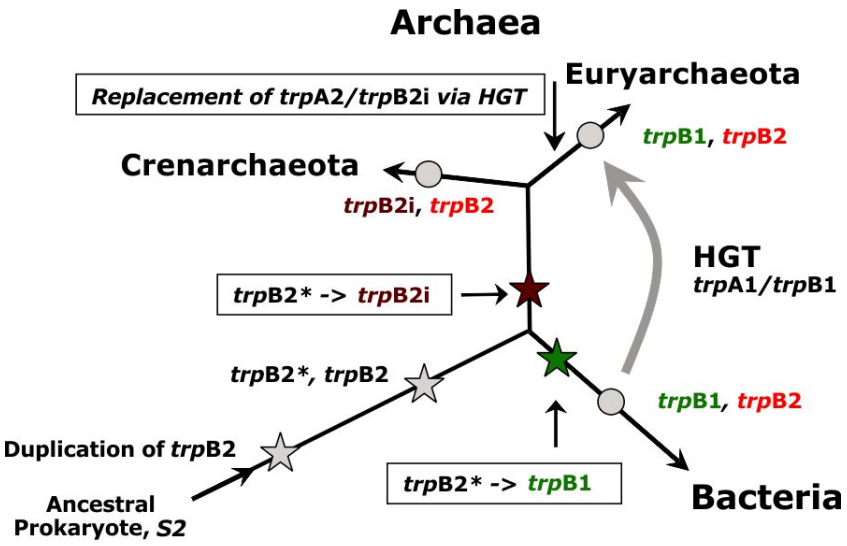


Figure 61: Model of TrpB evolution.
This model is based on the most plausible explanation for the present distribution of TrpB2 enzymes. Stars mark events of genomic rearrangements, circles represent ancient predecessors. HGT stands for horizontal gene transfer. Figure is according to (Merkl, 2007).

Accordingly, the easiest way to explain the distribution of TrpB enzymes in extant organisms is an ancient duplication of TrpB2. Whereas one copy was passed on unmodified, the other copy became the TrpB1 enzyme in the ancestor of Bacteria and developed into the TrpB2i enzyme in the ancestor of Archaea. The TrpA enzyme co-evolved to TrpA1 and TrpA2 to gain the exclusive interaction with TrpB1 and TrpB2i, respectively. In Euryarchaeota, TrpA1 and TrpB1 replaced the original TrpA2 and TrpB2i by a horizontal gene transfer event.

The present work confirms and refines this model. It reveals that all TrpB enzymes act as tryptophan synthases under physiological conditions. An early divergent evolution of TrpB1 and TrpB2 enzymes is most likely as the enzymes differ in substrate specificity and in their complex formation behavior. Whereas TrpB1 enzymes use L-serine for the synthesis of tryptophan and form a permanent TS with TrpA1, TrpB2 enzymes use OPS for the synthesis of tryptophan and form a transient, ligand-dependent TS with TrpA2. Evolution of an interaction of TrpB2 with TrpA2 is conceivable, as only an N-terminal part is required for the protein-protein interaction.

Like in permanent TS, the quaternary assembly allows for the efficient transfer of indole between the active sites. However, the rate of tryptophan synthesis of the transient TS is rather low. This might be the reason for its replacement by the permanent TS in Euryarchaeota. Phylogenetic analysis indicate that this replacement is the result of a horizontal gene transfer event from an ancestor of Firmicutes. The characterization of the newly emerging permanent TS that had most probably existed in the last common ancestor of Bacteria revealed that the transfer of indole between the active sites and the activation of the α-subunit in the TS had already existed at the rise of permanent TS. The rate of tryptophan synthesis of permanent TS might have initially paralleled that of transient TS. However, its evolution towards a more efficient catalysis might explain its widespread distribution in extant organisms. The TrpB2 enzyme, which lacks the ability to form a TS, might primarily utilize free indole as substrate. Its importance for extant organisms seems to be minor as indicated by its disappearance in multiple phyla.

4 SUMMARY

The permanent tryptophan synthase complex (TS) serves as a model system for the investigation of sophisticated protein-protein interactions. The permanent TS consists of a dimer of β-subunits and two α-subunits, which are located at both sites of the dimer, resulting in a αββα arrangement. The α-subunit catalyzes the aldolytic cleavage of indole-3-glycerol-phosphate (IGP). Glyceraldehyde-3-phosphate (G3P) is released whereas indole is channeled through a hydrophobic tunnel to the active site of the β-subunit. There, it reacts with L-serine in a PLP-dependent condensation reaction to L-tryptophan and water. The quaternary assembly of the permanent TS enables the bi-directional activation of the subunits within the complex. This ensures the coordination of IGP cleavage and condensation of indole and L-serine to tryptophan.

The effects of complex formation and allosteric communication on the enzymatic parameters of the subunits differ in detail in diverse extant organisms. We were interested in the questions how these differences reflect the evolutionary history and which characteristics a primordial TS might have featured. In order to answer these questions, a phylogenetic tree was constructed using concatenated αβ-sequences. The α- and β-sequences corresponding to the root of this tree are the most probable predecessor of permanent TS in the last common ancestor (LCA). The LCA α- and β-subunits were separately expressed in *E. coli*. Steady-state kinetics revealed that the LCA α-subunit is comparably active as extant α-subunits whereas the LCA β-subunit has a higher activity compared to extant β-subunits. A strong interaction between the LCA subunits with a K_d in the nanomolar range was detected by fluorescence titration. The formation of a LCA TS with αββα stoichiometry was verified by analytical size exclusion chromatography. Complex formation results in a significant increase of the activity of the α-subunit and a slight decrease of the activity of the β-subunit. The slow rate of tryptophan-release might be the rate-limiting step of the β-subunit in the LCA TS as indicated by stopped-flow experiments. Different accessibilities of isolated and complexed β-subunits indicated that a hydrophobic tunnel between the active sites had already existed in the LCA TS. Mutual adaptation of the subunits, which results in a more efficient catalysis of tryptophan and the specialization as tryptophan synthase, seems to have evolved recently to scope with the elevated need of fast-growing organisms for tryptophan.

The TrpB enzyme family consists of the large group of TrpB1 enzymes (β-subunits of permanent TS) and the smaller group of TrpB2 enzymes. The sequence identity between the groups is less than 30 %, which indicates an early divergent evolution in the TrpB enzyme family. The active site of both groups differs in one position, which is an aspartate in all TrpB1 enzymes and an arginine in all TrpB2 enzymes. Whereas TrpB1 has a specificity for L-serine, TrpB2 uses O-phospho-L-serine (OPS) for the synthesis of tryptophan. The importance of the conserved arginine for the preference of OPS as substrate was validated by structural analysis of a TrpB2 enzyme with bound OPS and mutagenesis studies.

The TrpB2 group is further subdivided into TrpB2i enzymes, which are encoded within the *trp* operon, and TrpB2a and TrpB2o enzymes, which are encoded outside of the *trp* operon. It had been previously shown that a transient, ligand-dependent TS with αββ stoichiometry is formed by ssTrpB2i and ssTrpA2 in *S. solfataricus*. The influence of ssTrpB2i ligands on the stability of the transient TS was investigated. It turned out that the β-substituent of a bound amino acid at the active site of ssTrpB2i does not influence the interaction with ssTrpA2. This indicates that chemical transformations at the β-subunit have no influence on the TS stability. Like in permanent TS, the active sites of the α- and β-subunits seem to be connected by a hydrophobic tunnel. Indications for the existence of the hydrophobic tunnel are: 1) the GP-bound α-subunit prevents the conversion of external indole to tryptophan, 2) only little indole is released during the TS reaction and 3) the main amount of indole formed at the α-subunit is immediately converted to tryptophan by the associated β-subunit. These notions argue for the early divergence in the TS evolution.

The phylogeny of transient TS was investigated by a sequence similarity network (SSN) of the TrpA enzyme family. It revealed that TrpA2 enzymes also exist in organisms like *Thermoplasma volcanium*, which have only TrpB2a, encoded outside of the *trp* operon. This indicates that tvTrpA2 and tvTrpB2a form a transient, ligand-dependent TS. This hypothesis was validated by analytical size exclusion chromatography and SPR. A comparison of the structures of the permanent TS from *S. typhimurium* and the transient TS from *S. solfataricus* shows that they differ in the quaternary assembly of the subunits. While in the permanent TS three parts of stTrpB1 contribute to the interaction with stTrpA1, only one N-terminal part of ssTrpB2i interacts with ssTrpA2 in the transient TS. The transfer of this N-terminal part of ssTrpB2i or tvTrpB2a is sufficient to convert the non-interacting ssTrpB2a into

an enzyme, capable of interacting with and activating an α-subunit (ssTrpA2 or tvTrpA2). This finding illustrates conservation in structure and conformational flexibility present within the family of TrpB2 enzymes.

5 ZUSAMMENFASSUNG

Der permanente Tryptophansynthase Komplex (TS) dient als Modellsystem zur Untersuchung von hochentwickelten, sich wechselseitig regulierenden Protein-Protein Interaktionen. Der permanente TS besteht aus einem Dimer von β-Untereinheiten und zwei an beiden Seiten des Dimers assoziierten α-Untereinheiten, welche eine αββα-Anordnung ergeben. Die α-Untereinheit katalysiert die aldolytische Spaltung von Indol-3-glycerolphosphat (IGP). Während das bei dieser Reaktion entstandene Glycerinaldehyd-3-phosphat (G3P) aus dem Komplex diffundiert, gelangt Indol durch einen hydrophoben Tunnel innerhalb des Komplexes zum aktiven Zentrum der assoziierten β-Untereinheit. Dort wird es mit L-Serin in einer PLP-abhängigen Kondensationsreaktion zu Tryptophan umgesetzt. Die Anordnung der Untereinheiten erlaubt neben der Ausbildung eines hydrophoben Tunnels die wechselseitige Regulation der Aktivitäten beider Untereinheiten. Hierdurch wird gewährleistet, dass die Bildung von Indol und dessen Umsetzung zu Tryptophan miteinander verknüpft sind.

Die Auswirkungen von Komplexbildung und allosterischer Kommunikation auf die enzymatischen Parameter unterscheiden sich im Detail bei den Untereinheiten aus verschiedenen rezenten Organismen. Daher stellt sich die Frage, welche Charakteristika ein ursprünglicher TS hatte und welche Merkmale erst im Laufe der Artenentwicklung entstanden. Zur Beantwortung dieser Frage wurde ein Stammbaum ausgehend von konkatenierten αβ-Sequenzen erstellt. Die der Wurzel des Stammbaums entsprechenden α- und β-Sequenzen stellen den mutmaßlichen TS im letzten gemeinsamen Vorfahren der Bakterien (LCA) dar und können somit als Ursprungspunkt in der Evolution des permanenten TS angesehen werden. Die entsprechenden LCA α- und β-Untereinheiten ließen sich separat löslich in *E. coli* exprimieren. *Steady-state* Kinetiken zeigten, dass die monomere LCA α-Untereinheit eine mit rezenten α-Untereinheiten vergleichbare Aktivität hat und die dimere β-Untereinheit eine im Vergleich zu rezenten β-Untereinheiten erhöhte Aktivität aufweist. Durch Gelfiltrationsexperimente und Fluoreszenztitrationen wurde eine starke Interaktion zwischen den Untereinheiten unter Bildung eines LCA TS mit αββα-Stöchiometrie nachgewiesen. Die Komplexbildung führt zur deutlichen Erhöhung der katalytischen Effizienz der LCA α-Untereinheit und zu einer leichten Verringerung der Aktivität der LCA β-Untereinheit. *Pre-steady-state* Kinetiken deuten

darauf hin, dass die Aktivität der β-Untereinheit im LCA TS durch eine verlangsamte Tryptophan-Freisetzung limitiert sein könnte. Die unterschiedliche Zugänglichkeit von freier bzw. komplexierter LCA β-Untereinheit für ein großes Nukleophil zeigte, dass die aktiven Zentren im LCA TS genau wie in rezenten TS durch einen Tunnel verbunden sind. Eine wechselseitige Anpassung der Untereinheiten mit einer daraus resultierenden Erhöhung der katalytischen Effizienz der Gesamtreaktion sowie eine Spezialisierung auf die alleinige Funktion als Tryptophansynthase erfolgte wahrscheinlich erst zu einem späteren Zeitpunkt in der TS Evolution. Möglicherweise wurde diese Entwicklung getrieben vom großen Bedarf an Tryptophan in schnell wachsenden Organismen.

Die TrpB Enzymfamilie besteht aus der großen Gruppe der TrpB1 Enzyme (β-Untereinheiten der permanenten TS) und aus der kleineren Gruppe der TrpB2 Enzyme. Eine Sequenzidentität von unter 30 % zwischen TrpB1 und TrpB2 deutet auf eine frühe divergente Evolution in der TrpB Enzymfamilie hin. In den aktiven Zentren ist an einer bestimmten Position bei TrpB1 ein Aspartat und bei allen TrpB2 ein Arginin konserviert. Durch HPLC Analysen konnte O-Phospho-L-Serin (OPS) als Substrat von TrpB2 Enzymen identifiziert werden. Die Strukturanalyse eines TrpB2 Enzyms mit gebundenem OPS verdeutlicht die Koordination der Phosphatgruppe durch das konservierte Arginin. Die Substratspezifität-determinierende Rolle dieses Restes wurde durch Mutationsanalysen bestätigt. Somit konnte TrpB2 die Funktion der OPS abhängigen Tryptophansynthase zugeordnet werden.

Die TrpB2 Gruppe ist unterteilt in TrpB2i Enzyme, welche innerhalb des *trp* Operon kodiert werden, und TrpB2a und TrpB2o Enzyme, welche außerhalb des *trp* Operons kodiert werden. Für ssTrpB2i aus *S. solfataricus* konnte vorangehend gezeigt werden, dass es mit der α-Untereinheit dieses Organismus (ssTrpA2) einen transienten, ligandenabhängigen TS mit αββ-Stöchiometrie bildet. Die Abhängigkeit der Interaktion von ssTrpB2i Liganden wurde in der vorliegenden Arbeit eingehend untersucht. Dabei stellte sich heraus, dass der β-Substituent der gebundenen Aminosäure keinen Einfluss auf die Interaktion mit ssTrpA2 hat. Dies deutet darauf hin, dass der Komplex während der gesamten Katalyse mit den dabei stattfindenden Änderungen am β-Substituenten der gebundenen Aminosäure bestehen bleibt. Wie bei permanenten TS sind wahrscheinlich auch bei transienten TS die aktiven Zentren der α- und β-Untereinheiten miteinander verbunden, da 1) die Reaktion von externem Indol zu Tryptophan durch die GP-gebundene α-Untereinheit behindert

wird, 2) nahezu kein Indol aus dem Komplex entweicht und 3) der Großteil von dem an der α-Untereinheit gebildeten Indol an der assoziierten β-Untereinheit zu Tryptophan umgesetzt wird. Dies lässt vermuten, dass es sich bei transienten TS und permanenten TS um die Produkte einer divergenten Evolution handelt.

Um die Gesamtheit von in der Natur vorkommenden transienten TS zu identifizieren, wurde eine phylogenetische Analyse der TrpA Enzymfamilie unternommen. Hier zeigte sich, dass ssTrpA2 Teil einer größeren Untergruppe ist. Diese beinhaltet auch TrpA2 Enzyme, welche in Organismen wie *Thermoplasma volcanium* vorkommen, die nur das außerhalb des *trp* Operons kodierte TrpB2a Enzym besitzen. Dies legt nahe, dass tvTrpA2 und tvTrpB2a einen transienten, ligandenabhängigen Komplex bilden. Diese Vermutung konnte durch Gelfiltration und SPR bestätigt werden. Der Vergleich der Kristallstrukturen des transienten TS aus *S. solfataricus* und des permanenten TS aus *S. typhimurium* zeigt, dass diese sich in ihrer Quartärstruktur unterscheiden. Während im Fall des permanenten TS drei Bereiche von stTrpB1 die Interaktionsfläche bilden, ist beim transienten TS nur ein N-terminaler Bereich von ssTrpB2i in direktem Kontakt mit der entsprechenden α-Untereinheit. Die Übertragung dieses Bereichs aus ssTrpB2i oder tvTrpB2a ist ausreichend, um aus dem nicht-interagierenden ssTrpB2a Enzym die β-Untereinheit eines transienten TS zu machen, welche die assoziierte α-Untereinheit (ssTrpA2 bzw. tvTrpA2) aktiviert. Dies verdeutlicht die Konserviertheit in der Struktur und in der konformationellen Flexibilität innerhalb der Gruppe der TrpB2 Enzyme.

6 MATERIALS

6.1 Instrumentation

Autoclave:	
Series EC Stream Sterilizers	WEBECO, Selmsdorf
Balances:	
MC1	SARTORIUS, Göttingen
PL3000	METTLER TOLEDO, Gießen
SI-114	DENVER INSTRUMENT, Göttingen
Biacore-X100(+)	GE HEALTHCARE, München
CD spectro-polarimeter J-815	JASCO GmbH, Groß-Umstadt
Cell Densitiy Meter Ultrospec 10	GE HEALTHCARE, München
Centrifuges:	
Centrifuge 5810R	EPPENDORF, Hamburg
Centrifuge 5415D	EPPENDORF, Hamburg
Centrifuge 5415R	EPPENDORF, Hamburg
Sorvall RC 2B, 5C plus	DU PONT Instruments, Bad Homburg
Avanti J-26 XP	BECKMAN COULTER, Krefeld
Chromatographic devices:	
1200 HPLC system	AGILENT, Böblingen
ÄKTA basic better	GE HEALTHCARE, München
ÄKTA prime	GE HEALTHCARE, München
ÄKTA purifier 10	GE HEALTHCARE, München
LaChrome HPLC system	MERCK-HITACHI, Darmstadt
columns:	
Gemini-NX 3 µm C18 110A 150x3 mm	PHENOMENEX, Aschaffenburg
HisTrap FF 5 ml	GE HEALTHCARE, München
Kromasil 3 µm C18 4x250 mm column	BISCHOFF, Leonberg
MonoQ 5/50 GL	GE HEALTHCARE, München
Superdex 75 pg HiLoad 16/60	GE HEALTHCARE, München
Superdex 75 pg prep grade 26/600	GE HEALTHCARE, München
Superdex 200 10/300 GL	GE HEALTHCARE, München
Computer	

Dell Optiplex Systems	DELL Inc., Round Rock, USA
Cyclone phosphor-imager	PACKARD BioScience, Meriden, USA
Fluorescence spectrometer FP-6500	JASCO GmbH, Groß-Umstadt
Freezer -80° C	MDF-U72V, SANYO, Tokyo, Japan
Freezer -20° C	LIEBHERR, Nussbaumen
Gas burner, Gasprofi 2SCS	WLD-TEC GmbH, Göttingen
Gel electrophoresis system:	
Agarose gel electrophoresis chamber	
Agarose electrophoresis unit	HOEFER Pharmacia Biotech, USA
SDS electrophoresis chamber	
Mighty Small II	HOEFER Pharmacia Biotech, USA
Multi Gel Caster Assembling gel apparatus	GE HEALTHCARE, München
Glass pipettes and glassware	FISCHER SCIENTIFIC, Schwerte NOVOGLAS, Bern, Swiss SCHOTT, Mainz
Heating block-Thermostat HBT-2 131	HLC, Bovenden
Incubator	BINDER GmbH, Tuttlingen
Magnetic stirrer:	
MR0, MR2000	HEIDOLPH, Kehlheim
MR1, MR3001 (heatable)	HEIDOLPH, Kehlheim
MicroCal VP-DSC	MALVERN Instruments, Malvern, UK
MicroCal VP-ITC	MALVERN Instruments, Malvern UK
Microliter pipettes Research	EPPENDORF, Hamburg
Microplate reader Infinite M200 Pro	TECAN, Austria
Microwave HMT 842C	BOSCH, Nürnberg
Mosquito LCP roboter	TTP LABTECH, Melbourn, UK
Multi-Doc-It Digital Imaging System	UVP Inc., USA
PCR-cycler:	
Mastercycler personal	EPPENDORF, Hamburg
Mastercycler gradients	EPPENDORF, Hamburg
Peristaltic pump, Miniplus 2	GILSON Medical Electronics, France
pH-Meter Level1	INOLAB, Weilheim
PhosphorImager Film FLA-3000	FUJIFILM, Tokyo, Japan

Plate shaker Rocking Platform	BIOMETRA, Göttingen
Power supply unit:	
Power Pack P25	BIOMETRA, Göttingen
Power Supply EPS 301	GE HEALTHCARE, München
Quartz cuvettes	
101-QS (layer thickness 10 mm)	HELLMA GmbH & Co. KG, Müllheim
105-QS (layer thickness 10 mm)	HELLMA GmbH & Co. KG, Müllheim
110-QS (layer thickness 10 mm)	HELLMA GmbH & Co. KG, Müllheim
Shaking incubator:	
Certomat H	BRAUN Biotech, Melsungen
Certomat BS-1	BRAUN Biotech, Melsungen
Multitron	INFORS HT, Bottmingen, Swiss
SX20 Stopped-flow spectrometer	APPLIED Photophysics, Surrey, UK
Ultrafree-20 nanopore water system	MILLIPORE, Eschborn
Ultrasonic system, Branson Sonifier 250 D	HEINEMANN, Schwäbisch Gmünd
UV-light table Reprostar	CAMAG Chemie Erzeugnisse, Berlin
UV-Vis spectral photometer V650	JASCO GmbH, Groß-Umstadt
UV-Vis Biophotometer	EPPENDORF, Hamburg
Vakuum pump ME 2C	VACUUMBRAND, Wertheim
Vortex Genie 2	SCIENTIFIC IND., Bohemia, USA

6.2 Consumables

Easy-Xtal plates	QIAGEN, Hilden
Centrifugal Filter Device	
Amicon Ultra-15 (mwco: 10 kDa)	Millipore, Bedford, USA
NAP-5, -10, -25 columns	GE HEALTHCARE, München
Dialysis tubing Visking, 27/32, 14 kDa	ROTH GmbH &Co, Karlsruhe
Disposable syringes, Omnifix® 60 ml	BRAUN Biotech, Melsungen
Filter paper	WHATMAN, Maidstone, England
Membrane filter ME24 Ø47 mm; 0.2 µm	SCHLEICHER&SCHUELL, Dassel
Microtiter plates, Vis, 96 well, flat bottom	GREINER, Nürtingen
Nitrocellulose filter (Ø13 mm)	MILLIPORE, Eschborn
Parafilm „M“ Laboratory Film	PECHINEY, Menasha, USA
Pasteur pipettes	HIRSCHMANN, Ebermannstadt

PCR-tubes 0.2 ml	PEQLAB, Erlangen
Petri dish 94/16	GREINER bio-one, Nürtingen
Pipette tips	SARSTEDT, Nümbrecht
Plastic cuvettes	
½ microcuvettes, UV-transparent	SARSTEDT, Nümbrecht
1ml cuvettes	SARSTEDT, Nümbrecht
Plastic tubes:15 ml, 50 ml	SARSTEDT, Nümbrecht
Reaction vessels 1.5 ml, 2 ml	ROTH, Karlsruhe
	EPPENDORF, Hamburg
Reaction vessel with screw-cap, 2 ml	SARSTEDT, Nümbrecht
Syringe filter, pore size 0.2 µm, 0.45 µm	RENNER GmbH, Daunstadt

6.3 Chemicals

All chemicals used were graded p.a. and purchased from the companies listed below.

ALFA AESAR	Karlsruhe
APPLICHEM GmbH	Darmstadt
BIO101 Inc.	Carlsbad, USA
BIO-RAD LABORATORIES	Hercules, USA
BIOZYM	Hess. Oldendorf
BODE CHEMIE	Hamburg
BOEHRINGER MANNHEIM	Mannheim
CARL ROTH GMBH & Co. KG	Karlsruhe
DIFCO	Dreieich
FLUKA	Neu-Ulm
GE HEALTHCARE	München
GERBU Biotechnik GmbH	Gailberg
GIBCO/BRL	Eggestein
MERCK	Darmstadt
MP BIOCHEMICALS	Illkirch, France
NATIONAL DIAGNOSTICS	Simerville, USA
OXOID	Wesel
RIEDEL-DE HAEN	Seelze

ROCHE DIAGNOSTICS	Mannheim
ROTH	Karlsruhe
SERVA	Heidelberg
SIGMA-ALDRICH	Deisenhofen
VWR	Leuven, Belgium

IGP was synthesized and purified according to Bolz (Bolz, 2006).

6.4 Kits for molecular biology

GeneJET Plasmid Miniprep Kit	MBI FERMENTAS, St.-Leon-Rot
GeneJET Gel Extraction Kit	MBI FERMENTAS, St.-Leon-Rot

6.5 Kits for protein crystallization

Additive Screen	HAMPTON RESEARCH, Aliso Viejo, USA
MIDAS	MOLECULAR DIMENSIONS, Suffolk, UK
PEG Rx	HAMPTON RESEARCH, Aliso Viejo, USA
ProPlex	MOLECULAR DIMENSIONS, Suffolk, UK

6.6 Enzymes

Alkaline phosphatase (CIP)	NEW ENGLAND BIOLABS, Frankfurt a. M.
DNA polymerases	
GoTaq	PROMEGA, Mannheim
Phusion High-Fidelity	NEW ENGLAND BIOLABS, Frankfurt a. M.
Pwo	ROCHE DIAGNOSTICS, Mannheim
Restriction endonucleases	NEW ENGLAND BIOLABS, Frankfurt a. M.
T4-DNA ligase	MBI FERMENTAS, St.-Leon-Rot
Trypsine	MERCK, Darmstadt

6.7 Bacterial strains

***E. coli* Turbo** (NEW ENGLAND BIOLABS, Frankfurt a. M.)

F′ *proA*$^{+}$*B*$^{+}$ *lacI*^{q} *Δ(lacZ)M15/fhuA2 Δ(lac-proAB) glnV gal R(zgb-210::Tn10)* Tet^{s} *endA1 thi-1 Δ(hsdS-mcrB)5*

E. coli Turbo cells are T1-phage resistant. As the *rec*A function of the strain is intact, *E. coli* Turbo cells grow fast and form visible colonies after 8 h incubation at 37° C.

***E. coli* BL21(DE3)** (Studier & Moffat, 1986)

*hsd*S *gal* [λcl *ts*857 *cnd*1 *hsd*R17 *rac*A1 *end*A1 *gyr*A96 *thi*1 *rel*A1]

E. coli BL21(DE3) cells carry a gene for T7 RNA polymerase on their chromosome, which is used for gene expression in pET systems.

***E. coli* BL21-CodonPlus(DE3) RIPL** (STRATAGENE, La Jolla, USA)

B F^- *ompT hsdS*(r_B^- m_B^-) *dcm*$^+$ Tet^r *gal* λ(DE3) *end*A Hte [*arg*U *pro*L Cam^r] [*arg*U *ile*Y *leu*W Strep/$Spec^r$]

E. coli BL21-CodonPlus (DE3) contains extra copies of rare-codon tRNAs. They enable the efficient heterologous expression of proteins from genes, which lack the codon usage of *E. coli.* The cells have a pACYC plasmid with extra copies of the *arg*U and *pro*L tRNA genes, and a pSC101 plasmid with extra copies of the *arg*U, *ile*Y und *leu*W tRNA genes. The pACYC-plasmid confers resistance to chloramphenicol and the pSC101-plasmid confers resistance to streptomycin and spectinomycin.

***E. coli* BL21(DE3) T7 Express I^q** (NEW ENGLAND BIOLABS, Frankfurt am Main)

*fhu*A2 *lac*Z::T7*gene*1 *lon omp*T *gal sul*A11 *dcm* R(*zgb*-210::*Tn*10--Tet^S) *end*A1 Δ(*mcr*C-*mrr*)114::IS10 R(*mcr*-73::*miniTn*10--Tet^S)2) *lac*I^q (Cam^R)

E. coli BL21(DE3) T7 Express I^q cells are derivatives of *E. coli* BL21(DE3), which have the gene for T7 polymerase in the *lac* operon. They are phage T1 resistent due to a deletion of *fhu*A2. The cells exhibit tetracycline resistance.

***E. coli* T7 Express Rosetta** (research group of Prof. Dr. R. Sterner)

*fhu*A2 *lac*Z::T7*gene*1 *lon omp*T *gal sul*A11 *dcm* R(*zgb*-210::*Tn*10--Tet^S) *end*A1 Δ(*mcr*C-*mrr*)114::IS10 R(*mcr*-73::*miniTn*10--Tet^S)2) [*arg*U *arg*W *gly*T *ile*X, *leu*W *met*T *pro*L *thr*T *thr*U *tyr*U Cam^r]

This expression strain was constructed in the research group of Prof. Dr. R. Sterner. Originating from *E. coli* T7 expression cells, the strain contains an additional pRARE plasmid from *E. coli* Rosetta(DE3)pLysS. The plasmid encoded tRNA for rarely used codons encoding Gly, Arg, Ile, Leu and Pro. Besides the tetracycline resistance, the cells exhibit, due to the pRARE plasmid, accessory chloramphenicol resistance.

6.8 Vectors

6.8.1 pET vectors

Genes inserted into the multiple clonig site (MCS) of pET vectors (plasmid for expression by T7 RNA Polymerase) are transcribed by the RNA-polymerase of the phage T7 (Studier et al., 1990). The expression of genes takes place in special *E. coli* strains, which carry a chromosomal copy of the T7 RNA polymerase. The expression of the T7 RNA polymerase gene proceeds under the control of the *lac*UV5 promotor operator and is induced by the addition of IPTG. The gene for the *lac*-repressor (*lac*I), which is required for suppression of gene expression in the absence of induction, is located on the plasmid and is constitutively expressed. The pET21a(+) vector encodes an optional C-terminal $(His)_6$-*tag* sequence and confers ampicillin resistance. The pET24a(+) vector encodes an optional C-terminal $(His)_6$-*tag* sequence and confers kanamycin resistance.The pET28a(+) vector encodes an optional N-terminal $(His)_6$-*tag* sequence and confers kanamycin resistance.

6.8.2 pMAL-c2 vector

Genes are inserted into the multiple cloning site (MCS) of pMAL-c2 (di Guan et al., 1988). The MCS is downstream of *mal*E to allow for the expression of fusion proteins with the maltose binding protein (MBP). Gene expression proceeds under the control of the P_{tac} promotor operator and is induced by the addition of IPTG.

6.9 Sequencing primers

Primers were obtained from METABION (Martinsried).

5' T7Promotor
5'- TAA TAC GAC TCA CTA TAG GG -3'

3' T7Terminator
5'- GCT AGT TAT TGC TCA GCG G -3'

5' *mal*E
5'- GGT CGT CAG ACT GTC GAT GAA GCC-3'

3' M13/pUC
5'- CGC CAG GGT TTT CCC AGT CAC GAC -3'

6.10 Ladders and markers

The size of DNA fragments in agarose gels was determined with the help of the GeneRule 1 kb Plus DNA Ladder (MBI FERMENTAS, St. Leon-Rot). The size of proteins in SDS-PAGE gels was determined with the help of the Unstained Protein Molecular Weight Marker (MBI FERMENTAS, St. Leon-Rot).

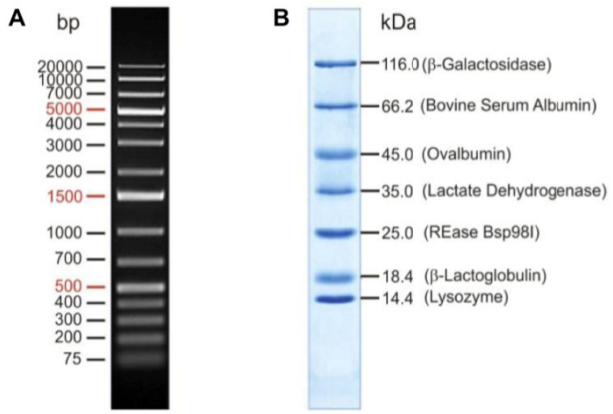


Figure 62: DNA and protein ladder and marker.
(A) GeneRule 1 kb Plus DNA Ladder (MBI FERMENTAS). **(B)** Unstained Protein Molecular Weight Marker (MBI FERMENTAS) for SDS-PAGE (12.5 %).

6.11 Buffers and solutions

Unless otherwise specified, buffers were filter-sterilized and stored at room temperature. For all preparations, Milli-Q grade water (MILLIPORE) or double distilled water (ddH_20) was used.

6.11.1 Buffers and solutions for working with *E. coli*

Antibiotics

Dissolved and filter-sterilized antibiotics were placed at -20° C for long-term storage.

Ampicillin (1000 x)	150 mg/ ml ampicillin (Na-salt) dissolved in water and filter-sterilized
Chloramphenicol (1000 x)	30 mg/ ml chloramphenicol dissolved in 100 % EtOH and filter-sterilized
Kanamycin (1000 x)	75 mg/ ml kanamycin dissolved in water and filter-sterilized
Glucose	20 % ($^{w}/_{v}$) glucose dissolved in water, filter-sterilized, and stored at RT

Glycerol (87 %)	autoclaved and stored at RT
IPTG stock solution	0.5 M IPTG dissolved in water, filter-sterilized, and stored at -20° C
$MgCl_2$	1 M $MgCl_2$ dissolved in water, filter-sterilized, and stored at RT
$MgSO_4$	1 M $MgSO_4$ dissolved in water, filter-sterilized, and stored at RT
TFB I-buffer	100 mM KCl, 50 mM $MnCl_2$, 30 mM KAc, 10 mM $CaCl_2$, 15 % glycerol; stock solutions for the single components were stored at 4° C. A volume of 100 ml TFBI-buffer was prepared prior to use.
TFB II-buffer	100 mM Tris/HCl, pH 7.0, 10 mM KCl, 75 mM $CaCl_2$, 15 % glycerol; stock solutions for the single components were stored at 4° C. A volume of 100 ml TFBII-buffer was prepared prior to use.
10 x TB-Phosphate	2.31 g KH_2PO_4 (acid, 0.17 M) and 12.54 g K_2HPO_4 (base, 0.72 M) dissolved in water (total volume of 100 ml), autoclaved, and stored at RT

6.11.2 Buffers and solutions for molecular biology

PCR dNTP solution (2 mM)	A solution of dNTP's (2 mM of each A, C, G, and T) was prepared and stored at -20° C.
Agarose (1 %)	5 g agarose dissolved in 500 ml 0.5 x TBE, boiled and stored at 60° C
Ethidium bromide solution	10 mg/ ml ethidium bromide (EtBr)
Sucrose color marker	60 % ($^{w}/_{v}$) sucrose, 0.1 % ($^{w}/_{v}$) bromphenol blue, 0.1 % ($^{w}/_{v}$) xylencyanole FF dissolved in 0.5 x TBE;
TBE (5 x)	445 mM boric acid, 12.5 mM EDTA, 445 mM Tris (resulting pH-value: 8.15)

6.11.3 Buffers and solutions for SDS-PAGE

Ammonium persulfate (APS) solution	10 % ($^{w}/_{v}$) APS solution; filter-sterilized, and stored at -20° C

Coomassie staining solution	0.2 % ($^{w}/_{v}$) Coomassie Brilliant Blue G-250 and R-250, 50 % ($^{v}/_{v}$) ethanol, 10 % ($^{v}/_{v}$) HAc; filtered and stored at RT; protected from light
Buffer for SDS-PAGE resolving gel	0.4 % ($^{w}/_{v}$) SDS, 1.5 M Tris-HCl pH 8.8
Buffer for SDS-PAGE stacking gel	0.4 % ($^{w}/_{v}$) SDS, 0.5 M Tris/HCl pH 6.8
SDS-PAGE electrophoresis buffer	0.1 % ($^{w}/_{v}$) SDS, 0.025 M Tris, 0.2 M glycine (resulting pH-value: 8.5)
SDS-PAGE sample buffer (5 x)	5 % ($^{w}/_{v}$) SDS, 25 % ($^{w}/_{v}$) glycerol, 12.5 % ($^{v}/_{v}$) β-mercaptoethanol, 0.025 % ($^{w}/_{v}$) bromphenol blue, 1.25 M Tris/HCl pH 6.8

6.12 Bacterial growth media

For sterilization, the medium was autoclaved for 20 min at 121° C and 2 bar. For selective media, the corresponding antibiotics were added after cooling down of the medium to 60° C in terms of a filter-sterilized, 1000-fold concentrated stock solution.

Betaine sorbitol medium	0.5 % ($^{w}/_{v}$) yeast extract, 0.5 % ($^{w}/_{v}$) NaCl, 1.0 % ($^{w}/_{v}$) tryptone, 2 % ($^{v}/_{v}$) glycerol, 0.4 g/ l betaine, 9.2 % ($^{w}/_{v}$) sorbitol. After autoclaving, 50 mM K_2HPO_4, 1 % glucose ($^{w}/_{v}$) and 10 mM $MgCl_2$ were added prior to use.
Luria-Bertani (LB) medium	0.5 % ($^{w}/_{v}$) yeast extract, 1.0 % ($^{w}/_{v}$) NaCl, 1.0 % ($^{w}/_{v}$) tryptone
LB agar	LB medium plus 1.5 % ($^{w}/_{v}$) Bacto-Agar
SOB medium	0.5 % ($^{w}/_{v}$) yeast extract, 0.05 % ($^{w}/_{v}$) NaCl, 2.0 % ($^{w}/_{v}$) tryptone. After autoclaving, 10 mM $MgSO_4$, 10 mM $MgCl_2$ and 2.5 mM KCl (each filter-sterilized) were added prior to use.
SOC medium	SOB medium was supplemented with 20 mM glucose (filter-sterilized) following autoclavation.

6.13 Software

ÄKTA Unicorn Version 5.01 (318)	© GE HEALTHCARE
Biacore X100 Control Software 2.0	© GE HEALTHCARE
Biacore X100 Evaluation Software 2.0	© GE HEALTHCARE
CLC main workbench 6.0	©CLC BIO
Corel Draw X7	© COREL Corp.
CS ChemDraw Ultra 11.0	© CAMBRIDGESOFT
Cytoscape 3.2.1	(Shannon et al., 2003)
Endnote Version X4	© WINTERTREE Software Inc.
HPLC Multi-HSM-Manager	© HITACHI
Jalview 2.6.1	(Waterhouse et al., 2009)
MS Office 2007	© MICROSOFT CORPORATION
Origin 7	© ORIGINLAB
OptiQuant 02.50	© PACKARD Instruments Co.
PyMO 0.99rc6	© DELANO SCIENTIFIC LCC.
Gina Star 4.07	© RAYTEST
SigmaPlot 12.0	© SPSS Inc.
Spectra Manager 1 and 2	© JASCO
WinCoot (version 0.6.2)	(Emsley et al., 2010)

7 METHODS

7.1 Preparation of instrumentation and solutions

All thermostable solutions and media were autoclaved for 20 min at 121° C and 2 bar prior to use. Glassware and consumables were autoclaved and subsequently dried at 50° C in a compartment drier. Additionally, glassware was sterilized at 200° C for 4 h. Heat-labile solutions were prepared in stock solutions and filtered, either via a membrane filter with a pore size of 0.2 µm by use of a vacuum pump, or by using a syringe filter with a pore size of 0.2 µm or 0.45 µm. Solutions for chromatographic systems were degassed for at least 30 min in a desiccator prior to use.

7.2 Microbiological methods

7.2.1 Cultivation and storage of *E. coli* strains

E. coli strains were cultivated while shaking at 140 rpm (1 liter cultures) or 220 rpm (5, 50 and 250 ml cultures), respectively. Cells were grown in LB-medium unless otherwise stated. For plasmid-harboring strains the medium was supplemented with antibiotics (150 µg/ ml ampicillin, 30 µg/ ml chloramphenicol, 75 µg/ ml kanamycin, 12.5 µg/ ml tetracycline). To obtain single colonies, the cell suspension was plated on agar plates containing the adequate antibiotics, and incubated overnight at 37° C. For temporary storage, plates and suspensions were sealed and stored at 4° C. Glycerol cultures were prepared for long-term storage. For this purpose an aliquot of an overnight culture was mixed in a 1:1 ratio with 87 % glycerol, and stored in a sterile screw cap reaction vessel at -80° C.

7.2.2 Preparation of chemically competent *E. coli* cells

For the preparation of chemically competent *E. coli* cells (Inoue et al., 1990), 500 ml SOB medium was inoculated with the respective overnight culture to an OD_{600} of 0.1, and cultured at 37° C and 220 rpm until an OD_{600} of 0.6 was reached. The culture was incubated on ice for 15 min, transferred into 50 ml tubes, and cells were harvested by centrifugation (EPPENDORF Centrifuge 5810R, 4000 rpm, 10 min,

4° C). The cell pellet was resuspended in 100 ml ice-cold TFB I buffer and centrifuged a second time under the same conditions as stated above. The resulting pellet was resuspended in 10 ml ice-cold TFB II-buffer. Immediately after resuspension, 100 µl aliquots of the cell suspension were transferred to 1.5 ml reaction vessels on ice and stored at -80° C.

7.2.3 Transformation of chemically competent *E. coli* cells

For transformation of chemically competent cells, a 100 µl aliquot was thawed on ice, and about 100 ng plasmid DNA (maximum volume: 20 µl) were added. Following incubation on ice for 5 min, cells were placed for 45 s in a 42° C heat block, and subsequently chilled on ice for 5 min. A volume of 900 µl LB-medium was added, and cells were incubated for 1 h at 37° C in a shaker at 220 rpm to develop antibiotic resistance. Finally, dilutions of the cell suspension were plated on LB agar plates containing the appropriate antibiotics for selection.

7.3 Molecular biology methods

7.3.1 Isolation and purification of plasmid DNA from *E. coli*

Purification of plasmid DNA was done according to the principle of alkaline cell lysis (Le Gouill et al., 1994). Bacterial cell cultures are lysed by adding SDS and sodium hydroxide. The suspension is subsequently neutralized by addition of ammonium acetate; proteins and genomic DNA precipitated, whereas circular plasmid DNA remains in solution. For analytical isolation and purification of plasmid DNA from *E. coli*, the mini preparation kit from FERMENTAS (GeneJET Plasmid Miniprep Kit) was used. For this purpose 5 ml from overnight cell cultures were harvested by centrifugation (EPPENDORF 5415D, 13200 rpm, 1 min, RT). The isolation of plasmid DNA was performed according to the protocol supplied by the manufacturer. Bound plasmid DNA was eluted from the silica column by 50 µl sterile water. The recovered plasmid DNA was stored at -20° C.

7.3.2 Determination of DNA concentration

The DNA concentration was spectroscopically determined at a wavelength of 260 nm. According to Lambert-Beer's law an OD_{260} value of 1 (with

$^{0.1\%}A_{260}$ = 20 $cm^2\ mg^{-1}$ and a pathlength of 1 cm) corresponds to a DNA concentration of 50 µg/ ml dsDNA (35 µg/ ml RNA and 33 µg/ ml ssDNA, respectively). Thus, the DNA concentration can be calculated as follows:

$$c_{dsDNA} = \frac{A_{260} \cdot 50 \cdot f}{1000}$$

Equation 1: Determination of DNA concentration.
c_{dsDNA} concentration of double stranded DNA (µg/ µl)
A_{260} absorbance at 260 nm
f dilution factor

A pure DNA solution should not show measurable absorption above 300 nm, and its OD_{260}/OD_{280} quotient should be at least 1.8.

7.3.3 Agarose gel electrophoresis

DNA fragments are separated by agarose gel electrophoresis according to their length. DNA becomes visible under UV-light in the presence of the DNA intercalating dye ethidium bromide (Sharp et al., 1973). For preparation of agarose gels 1 % (w/v), agarose was dissolved in 0.5 % TBE buffer by boiling in the microwave. Following cooling down to 50- 60° C, 0.2 µl of an ethidium bromide stock solution (10 mg/ ml) per ml agarose was added. The solution was cast into a gel chamber, and a comb was inserted. Following solidification, the gel was covered with 0.5 % TBE buffer, and the comb was removed. The DNA samples were supplemented with DNA loading dye if required and pipetted into the gel pockets. The electrophoresis was performed at a voltage of 190 V for about 20 min. The negatively charged DNA migrates to the anode, whereby DNA fragments are retarded to a different extent by the agarose matrix. The fragments were detected under UV-light (λ= 302 nm) and documented using the Imager Multi-Doc-It Digital Imaging system. The size of fragments was estimated by the GeneRule 1 kb Plus DNA Ladder (Fermentas). If needed, the fragments were excised from the agarose gel with a scalpel and transferred to a reaction vessel. DNA was extracted using the GeneJET Gel Extraction Kit (FERMENTAS) according to the protocol supplied by the manufacturer. The isolated DNA was eluted in 30 µl sterile water and stored at -20° C.

7.3.4 Enzymatic manipulation of dsDNA

7.3.4.1 Cleavage of dsDNA by restriction endonucleases

Type II restriction endonucleases were applied, which bind to a palindromic recognition sequence (restriction site) and cleave dsDNA specifically (Wilson and Murray, 1991). These restriction enzymes generate single stranded overhangs (sticky ends), which contain 3'-hydroxyl and 5'-phosphate ends. For analytical cleavage, a maximum of 1 µg DNA was incubated at 37° C in the appropriate buffer with 20 U of each restriction enzyme for about 1 h in a volume of 50 µl. For preparative cleavage, 2 µg or the entire amount of PCR product and 2 µg of vector DNA were digested with 20 U of each restriction enzyme in a volume of 50 µl at 37° C for 1 h. The volume of added restriction endonuclease in the reaction mixture should not exceed 10 % of the total volume, as the activity of enzymes is influenced by glycerol, a component of the enzyme storage solution. The fragments were purified by agarose gel electrophoresis for subsequent ligation.

7.3.4.2 Ligation of DNA fragments

For ligation, digested vector and insert were mixed at an estimated molar ratio of 1:3 and ligated in a total volume of 20 µl with 1 U T4 DNA ligase (FERMENTAS or NEB) in the buffer supplied by the manufacturer, either overnight in a thermal cycler at 16° C or for 1 h at RT. Subsequently, competent *E. coli* cells were chemically transformed with the ligation mixture.

7.3.5 Amplification of DNA fragments by polymerase chain reaction (PCR)

The polymerase chain reaction (PCR; (Mullis and Faloona, 1987; Saiki et al., 1988) is used to amplify a specific DNA fragment *in vitro*. This is achieved by cyclic repetition of the denaturation of the double-stranded DNA, followed by the hybridization (annealing) of primers (synthetic oligonucleotides that flank the DNA sequence of interest) and enzymatic DNA synthesis (extension). The DNA fragment is exponentially amplified. The reaction was performed in a total volume of 50 µl in a thermal cycler (lid temperature 110° C). The standard reaction mixture contained 5-100 ng of template DNA, 2.5 U GoTaq DNA polymerase, 5 x Green GoTaq reaction buffer [contains 7.5 mM $MgCl_2$ (final concentration: 1.5 mM $MgCl_2$) and loading

buffer], 0.2 mM dNTP mix, and 1 µM of each primer. The accuracy of amplification was enhanced by the presence of 2.5 U Pwo polymerase, which has a 3' ->5' proofreading activity. The standard PCR program was as follows:

step	temperature (°C)	duration
1. Initialization step	95	3 min
2. Denaturation	95	45 s
3. Annealing	T_A	45 s
4. Extension	72	1 min/ kb
5. Finale elongation	72	10 min
6. Final hold	16	∞

Steps 2 to 4 were repeated 30 times.

The optimum melting temperature T_m and annealing temperature T_A of the primers were calculated according to Equation 2 and Equation 3 (Chester and Marshak, 1993):

$$T_m = 69{,}3 + 0{,}41 \cdot (\%GC) - \frac{650}{n}$$

Equation 2: Calculation of the melting temperature of oligonucleotides.
T_m melting temperature of primers (° C)
%GC GC-content of primers (%)
n number of nucleotides in the primer

$$T_A = \left(\frac{T_{m1} + T_{m2}}{2}\right) - 3°C$$

Equation 3: Calculation of the optimum annealing temperature of a primer.
T_A annealing temperature (° C)
T_{m1} & T_{m2} melting temperatures of the primers (° C)

The optimum annealing temperature was also experimentally determined using a thermal cycler with gradient function (EPPENDORF Mastercycler gradient). To this end, different PCRs with annealing temperatures between 50° C and 70° C were set up in parallel, and the yields of the amplification products were determined.
If applicable, the yield of the amplification product was increased by using two consecutive round of PCR. In a first round, perfectly matching short primers were used. The PCR product was purified and served as template in a second round using standard cloning primers.

7.3.6 Colony PCR

To verify the success of cloning, an insert screening was performed by colony PCR. Transformants grown on selective LB agar were added to the PCR mixture by picking a single colony with a pipette tip and transferring small quantities of cells to the reaction mixture. The cells were subsequently disrupted in the initial denaturation step at 95° C, and the released DNA was used as template in the following amplification cycles. Vector-specific amplification primers were generally used. The standard reaction mixture included 1 U GoTaq DNA polymerase, 5 x Green GoTaq reaction buffer, 0.2 mM dNTP-mix, 1 µM of each primer, and water to a final volume of 20 µl. The amplification was performed according to the amplification protocol described in 7.3.5.

7.3.7 QuikChange site-directed mutagenesis (QCM)

The QuikChange site-directed mutagenesis method allows for the efficient introduction of point mutations, insertions and deletions in any type of dsDNA plasmids. It is performed using Pfu DNA polymerase, which replicates both plasmid strains with high fidelity due to its 3'→5' proofreading activity.

The technique originally developed by STRATAGENE (La Jolla, USA) was modified according to a protocol published by (Wang and Malcolm, 1999) using a two-stage mutagenesis protocol (Figure 63). For introduction of point mutations complementary mutagenic primers with up to 35 bases in length and 12-15 bases of template complementary sequence on both sides of the mismatch were used. Pfu DNA polymerase extends and incorporates mutagenic primers during temperature cycling and generates a mutated plasmid containing staggered nicks. To avoid unproductive primer dimer formation, two separate primer extension reactions were performed initially. In the second step both primer extension reactions are combined and linearly amplified. The methylated DNA template is digested by treatment of the product with *Dpn*I. *E. coli* cells are transformed with the nicked vector DNA carrying the desired mutations, and the nicks are sealed by the DNA repair apparatus of the cells.

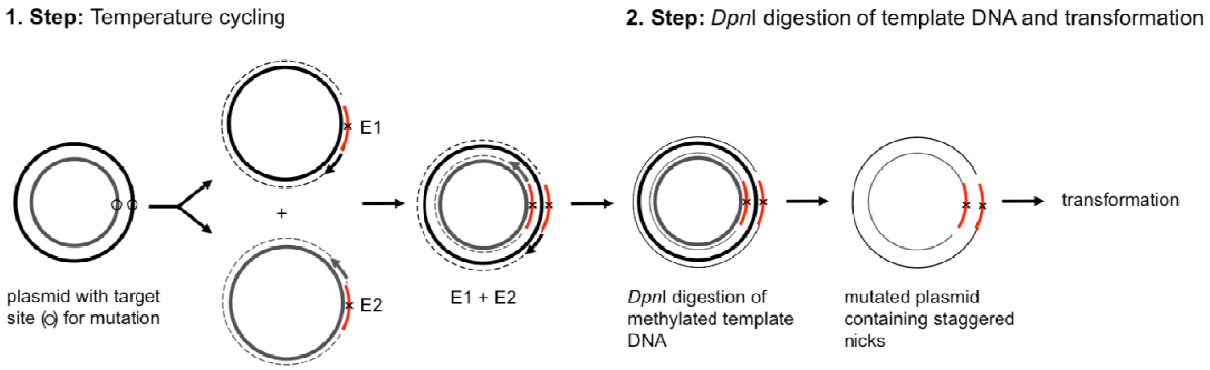


Figure 63: Overview of the QCM method.
In the first step the mutated plasmids containing staggered nicks are generated in two parallel primer extension reactions (E1 and E2). The reactions E1 and E2 are combined and amplified. In the second step, the DNA template is digested by treatment with *Dpn*I, and *E. coli* cells are transformed with the nicked vector DNA containing the desired mutations. The figure was adapted from the QuikChange Site-Directed Mutagenesis Kit manual (STRATAGENE).

The two separate primer extension reactions (E1 and E2) were performed each in a total volume of 50 µl containing 50 ng of template plasmid comprising the gene of interest, 2.5 U Pfu DNA polymerase, 10 x Pfu DNA polymerase reaction buffer [contains 20 mM $MgCl_2$ (final concentration: 2 mM $MgCl_2$)], 0.2 mM dNTP mix, and 1.6 µM of either the 5' or 3' primer. The standard primer extension program using Pfu DNA polymerase was as follows:

step	temperature (°C)	duration
1. Initialization step	95	3 min
2. Denaturation	95	45 s
3. Annealing	T_A	1 min
4. Extension	72	2 min/ kb
5. Finale elongation	72	10 min
6. Final hold	16	∞

Steps 2 to 4 were repeated 4 times.

Subsequently, 25 µl of each primer extension reaction (E1 and E2) were mixed, supplemented with additional 2.5 U Pfu DNA polymerase and subjected to 18 cycles of QCM as described above. The methylated template DNA was digested by directly adding 20 U *Dpn*I to 25 µl of the QCM reaction mix and incubated at 37° C for 1.5 h. Then, chemically competent *E. coli* Turbo cells were transformed using 10 µl of *Dpn*I-digested QCM mix.

7.3.8 Overlap extension PCR

Site-directed mutagenesis by overlap extension PCR (OE-PCR) uses two complementary mutagenic primers with a length of 25-35 bp that are complementary to the target DNA (Ho et al., 1989). The first two separate PCRs generated two overlapping fragments, using a set of mutagenic primer and flanking primer each. The gel-purified fragments are used as templates in combination with the flanking primers to amplify the entire gene. The amplification was carried out according to the standard protocol. For an illustrating scheme see Figure 64.

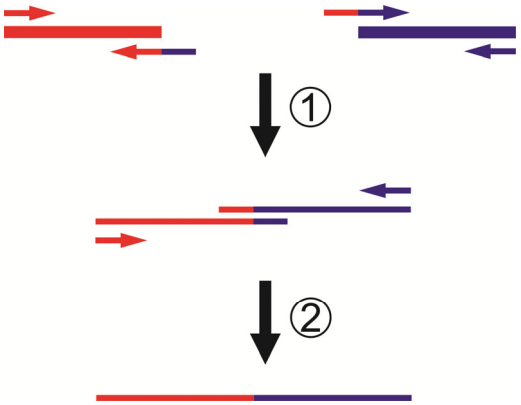


Figure 64: Scheme for standard OE-PCR.
In the first step the 5' and 3' fragments are amplified in two separate PCRs with a gene flanking primer and an overlapping primer using the target DNA as template (1). In the second step the fragments served as template and the full length construct is amplified using the flanking primers (2).

7.3.9 DNA sequencing

All constructs generated and used in this work were sequenced to validate the sequences. The determination of nucleotide sequences was performed by the company Geneart (part of Life Technologies, Regensburg). Samples for sequencing orders at Genart contained 100-300 ng DNA and 1.25 µM of sequencing primer in a total volume of 8 µl. The generated .ABI files were analyzed by the program CLC main workbench

7.3.10 Gene synthesis

Gene sequenzes for the reconstructed enzymes were optimized for expression in *E. coli* and synthesized by the company Geneart (part of Life Technologies, Regensburg). Genes were sent as inserts in a standard cloning vector from the manufacturer.

7.4 Protein biochemistry methods

7.4.1 Protein expression

7.4.1.1 Protein expression in analytical scale

To detect for the overexpression and solubility of heterologous produced protein in *E. coli*, small-scale expression test were done ('test expression'). To this end, 50 ml LB medium were supplemented with the appropriate antibiotics, inoculated with an overnight culture of a single colony to an OD_{600} of 0.1, and incubated at 37° C and 220 rpm. At an OD_{600} value of 0.4- 0.6, the cultures were induced by adding IPTG to a final concentration of 0.5 mM. Following incubation at various temperatures for 4 h to overnight, the cells were centrifuged (EPPENDORF Centrifuge 5415R, 8000 rpm, 5 min, 4° C). The cells were resuspended in 100 µl of 100 mM potassium phosphate buffer and disrupted by sonication (amplitude 20 %, 20 s). Following centrifugation (EPPENDORF Centrifuge 5415R, 13200 rpm, 10 min, 4° C), the supernatant (containing soluble protein) was saved, and the pellet (containing insoluble protein) was resuspended in an equal volume of the resuspension buffer. Aliquots of the supernatant (S) and the resuspended pellet (P) were supplemented with SDS sample buffer (5 x), and incubated for 5 min at 95° C. Subsequently, a volume of 5- 10 µl of both S and P was analyzed by SDS-PAGE. The expression conditions were optimized by using different *E. coli* host strains, by varying of the expression medium, and by utilizing different tags and/ or fusion partners.

7.4.1.2 Protein expression in preparative scale

For purification of proteins in preparative scale, 2- 16 x 1 l medium (in 3 l baffled flasks) was supplemented with the corresponding antibiotics and additives. The medium was inoculated with a freshly prepared overnight *E. coli* culture to an OD_{600} value of 0.1. Cells were grown according to the optimized conditions determined by analytical expression tests (7.4.1.1).

7.4.2 Protein purification

For large-scale protein purification, the cells were collected by centrifugation (BECKMAN COULTER centrifuge, Avanti J-26 XP, 4000 rpm, 20 min, 4° C). Collected cells were resuspended in buffer (25 ml/ l) and disrupted by sonication

(HEINEMANN branson sonifier 250 D, amplitude 50 %, 2 x 2 min, 30 sec puls/ 30 sec pause). The cell depris was removed by centrifugation (SORVALL RC5B, SS34 rotor, 14000 rpm, 30 min, 4° C). Application of a heat step, metal affinity chromatography, ion exchange chromatography and preparative size exclusion chromatography were used for protein purification. Subsequently, the buffer was exchanged by dialysis or NAP columns. The proteins were concentrated by ultrafiltration (MILLIPORE, Amicon 15 kDa cut-off) and directly used or stored at -80° C.

7.4.2.1 Heat step

Heterologous expressed thermostable proteins were purified by a heat step. For this purpose, the solution of soluble proteins was heated in a waterbath. Aggregated host cell proteins were subsequently removed by centrifugation (SORVALL RC5B, SS34 rotor, 14000 rpm, 30 min, 4° C).

7.4.2.2 Metal affinity chromatography (IMAC)

The column HisTrap FF crude (GE HEALTHCARE; CV: 5 ml, pressure limit: 0.3 MPa) was used for metal affinity chromatography. The column consists of sepharose 6 (FastFlow) and covalent linked iminodiacetic acid (IDA)-type chelator. IDA coordinates the divalent metal ion Ni^{2+}, which specifically interacts with histidine residues. As such, proteins with N-terminal or C-terminal hexahistidine *tag*s are retained on the column. Bound proteins are eluted by increasing concentrations of imidazole, which competes with the histidines for binding to the immobilized Ni^{2+} ions. Protein samples were steril filtered (0.45 µm) and applied on the HisTrap FF crude column in an Äkta chromatographic system. The following programs were used for protein purification and coupled on-column cleavage/ protein purification by IMAC:

Purification by IMAC

flow rate	4 ml/ min
equilibration	2 CV binding buffer
sample application	50- 400 ml protein solution in resuspension buffer*
wash	14 CV binding buffer
elution	15 CV gradient from 0- 75 % elution buffer**; 2 ml fractions were collected
final wash-out	5 CV elution buffer**
reequilibration	3 CV binding buffer

On-column cleavage and purification by IMAC

Flow rate	0.5 ml/ min for sample application, 4 ml/min for all other steps
equilibration	2 CV binding buffer
sample application	50- 200 ml protein solution in resuspension buffer*
wash	16 CV binding buffer
on-column cleavage	1 CV 50 µg/ ml trypsin solution in binding buffer; incubation for 2 h at room temperature
wash	5 CV elution buffer**
elution	15 CV gradient from 0- 75 % elution buffer**; 2 ml fractions were collected
final wash-out	5 CV elution buffer
reequilibration	3 CV binding buffer

* if not stated otherwise, resuspension buffer and binding buffer were identical
** the elution buffer is identical with the binding buffer except for the imidazol concentration
The buffer compositions are mentioned in the main text.

The elution of protein was monitored by following the absorbance at 260, 280 and 420 nm. The eluted fractions were analyzed by SDS-PAGE. For storage, the column was washed with 5 CV water and 3 CV 20 % ethanol.

7.4.2.3 Ion exchange chromatography (IEX)

Ion exchange chromatography is based on the competition of charged molecules (proteins and salt ions) for interaction with immobilized ion exchange groups of opposite charge. In the first stage, charged molecules are reversibly absorbed by the immobilized support material. Subsequently, bound molecules are eluted by a

gradient of steadily increasing ionic strength or a pH-gradient. The charge of the protein is mainly dependent on the amino acids with charged side chains. In the acidic or neutral pH-range the amino groups, mainly those of Lys, Arg and His, are protonated and the protein exposes cationic behavior, whereas in a neutral or basic pH-range, the carboxyl groups of Asp and Glu residues are negatively charged and consequently the protein is anionic. Ion exchange chromatography can be subdivided into cation exchange chromatography, in which positively charged ions bind to a negatively charged resin, and anion exchange chromatography, where negatively charged ions bind to a positively charged resin. In this work, the anion exchange column MonoQ 5/50 GL (GE HEALTHCARE, CV: 1 ml) was used. Protein samples were centrifuged (SORVALL RC5B, SS34 rotor, 14000 rpm, 30 min, 4° C) and applied on the column in an Äkta chromatographic system. The following program was used:

flow rate	**2 ml/ min**
equilibration	2 CV 10 mM potassium phosphate pH 7.5
sample application	protein in 10 mM potassium phosphate pH 7.5
wash	3 CV 10 mM potassium phosphate pH 7.5
elution	10 CV gradient from 0- 100 % 10 mM potassium phosphate pH 7.5 with 1M KCl
final wash-out	3 CV 10 mM potassium phosphate pH 7.5, 1M KCl
reequilibration	3 CV 10 mM potassium phosphate pH 7.5

The elution of protein was monitored by following the absorbance at 260, 280 and 420 nm. The eluted fractions were analyzed by SDS-PAGE. For storage, the column was washed with 5 CV water and 3 CV 20 % ethanol.

7.4.2.4 Preparative size exclusion chromatography (SEC)

Size exclusion chromatography is based on the principle of a reverse molecular sieve. When the mobile phase passes through the porous support material at a constant flow rate, small molecules are able to diffuse into the pores, while larger molecules are excluded. As a consequence, small molecules are retarded with respect to larger molecules, which cause proteins to be separated according to their size. In this work, preparative gel filtration was performed with the ÄKTA prime (GE HEALTHCARE) at 4° C, using a Superdex 75 pg prep grade 26/600 column (GE

HEALTHCARE, CV: ~320 ml). The column consists of highly cross-linked porous agarose beads (mean particle size: 34 µm) to which dextran has been covalently bound. Protein samples were centrifuged (SORVALL RC5B, SS34 rotor, 14000 rpm, 30 min, 4° C) and applied on the column by a 10 ml sample loop. Proteins were eluted with 1.2 CV of the same buffer at a flow rate of 1 ml/ min (back pressure: max. 0.3 MPa). The run was recorded using a manual plotter, which measured the absorbance at 280 nm. The fraction collector was started at the beginning of the elution, and fractions of 4 ml volume were collected.

7.4.3 Buffer exchange by dialysis or NAP columns

In order to exchange the buffer of a protein solution or to remove salt, dialysis was performed two times for at least 4 hours against a 100-fold volume excess of buffer at 4° C in the cold room. A dialysis tubing (Visking) with a molecular cut-off of 14 kDa was used, which retains the protein while low molecular substances can pass through the dialysis tubing. If applicable, small volumes of protein solutions were desalted or buffers were exchanged prior to biophysical analysis using Sephadex G-25 DNA Grade NAP columns (GE HEALTHCARE) according to the protocol supplied by the manufacturer.

7.4.4 Concentrating protein solutions

Protein solutions were concentrated using Amicon Ultra-15 centrifugal filter devices (MILLIPORE; molecular cut-off: 10 kDa) by centrifugation (Eppendorf Centrifuge 5810R, 4000 rpm, 4° C) according to the instructions of the manufacturer.

7.4.5 Storage of purified proteins

Purified and concentrated proteins were frozen in liquid nitrogen either dropwise or as 100 µl aliquots in PCR reaction vessels and stored at -80° C.

7.5 Analytical methods

7.5.1 Protein concentration determination by absorption spectroscopy

The aromatic amino acids tryptophan, tyrosine and phenylalanine as well as disulfide bonds (cystine) absorb UV-light in a wavelength interval between 250 and

300 nm. The molar extinction coefficient at 280 nm (ε_{280}) can be determined from the amino acid composition according to Equation 4 (Pace et al., 1995).

$$\varepsilon_{280} = \sum \text{Trp} \cdot 5500 + \sum \text{Tyr} \cdot 1490 + \sum \text{Cystine} \cdot 125$$

Equation 4: Determination of the molar extinction coefficient ε_{280}.
ε_{280} molar extinction coefficient at 280 nm ($M^{-1}cm^{-1}$)

The specific extinction coefficient can be calculated according to Equation 5.

$$^{0.1\%}A_{280} = \frac{\varepsilon_{280}}{MW}$$

Equation 5: Determination of the specific extinction coefficient $^{0.1\%}A_{280}$.
$^{0.1\%}A_{280}$ specific extinction coefficient at 280 nm (cm^2/ mg)
MW molecular weight of the protein (g/ mol)

Using Lambert-Beer's law, the protein concentration can be determined by measuring the absorbance at 280 nm (Equation 6):

$$A_{280} = {}^{0.1\%}A_{280} \cdot c \cdot d$$

$$c = \frac{A_{280}}{{}^{0.1\%}A_{280} \cdot d}$$

Equation 6: Determination of the protein concentration by using $^{0.1\%}A_{280}$.
A_{280} absorbance at 280 nm
c concentration (mg/ ml)
d pathlength (cm)
$^{0.1\%}A_{280}$ specific extinction coefficient at 280 nm (cm^2/ mg)

Absorbance spectra were recorded between 220 and 350 nm. The absorbance maximum should be at 278 nm and the A_{280}/ A_{250} ratio should be at least 1.8 for a pure protein solution. Protein without cofactor should not absorb light above 300 nm. The concentration of TrpB proteins was measured with the Bradford assay due to the light absorption of bound cofactor PLP.

7.5.2 Bradford assay

The Bradford assay is also used for the quantification of proteins in aqueous solutions. The assay contains Coomassie Brillant Blue G250, which predominately binds to cationic, nonpolar and hydrophobic regions of proteins. The absorption maximum is shifted from 465 nm to 595 nm when the dye binds to protein. The change in absorbance at 595 nm is linear in a protein range of 1.2 to 10.0 µg/ ml. For

the determination of protein concentrations, 200 µl of Bradford-reagent (BIORAD) were added to 800 µl of diluted protein solution. After 5 min incubation at room temperature, the absorbance was measured at 595 nm using diluted Bradford reagent as blank. The protein concentration was calculated by a linear regression standard curve of BSA according to Equation 7.

$$c = \frac{A_{595}}{m}$$

Equation 7: Determination of the protein concentration by the Bradford assay.
c concentration (mg/ ml)
A_{595} absorption at 595 nm
m slope of standard curve

7.5.3 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins are denatured by the detergent sodium dodecyl sulfate (SDS) and negatively charged proportional to their molecular weights. SDS binds to the protein in a ratio of approximately one molecule SDS per 1.4 amino acid residues. The net charge of proteins can be neglected compared to the negative charge of the bound SDS, which results in an approximately uniform mass to charge ratio. As a consequence, electrophoretic mobility depends only on the sieve effect of the gel: the migration speed is inversely proportional to the logarithm of mass (Laemmli, 1970). Table 11 shows the composition of the 12.5 % SDS gels used in this work.

Table 11: Composition of a 12.5 % SDS-PAGE gel.
Amount specification applies to 13 gels.

	resolving gel (12.5 %)	stacking gel (6 %)
resolving/stacking gel buffer	19.5 ml	7.38 ml
acrylamide-SL (30 %)	26.2 ml	5.9 ml
H_2O	31.58 ml	15.95 ml
TEMED	0.089 ml	0.029 ml
APS (10 %)	0.195 ml	0.089 ml

Samples were supplemented 1:4 with 5 x SDS-PAGE sample buffer and denatured at 95° C for 5 min. Gel pockets were loaded with 5- 20 µl of sample, and gels were run at 50 mA and 300 V for about 30 min. Subsequently, gels were stained with SDS-PAGE staining solution with a detection limit of Coomassie Brilliant Blue dye G-250 at around 200-500 ng protein/ mm^2. Gels were swayed for 10 min in the staining

solution and excess dye was removed by repeatedly boiling in water (microwave 900 W).

7.5.4 Analytical size exclusion chromatography

Apparent molecular weights of proteins and protein complexes were determined by analytical size exclusion chromatography. Proteins are separated according to their size. The retention volume decreases approximately linear with the logarithm of the molecular weigth. Globular calibration proteins were used for a standard curve in order to determine the apparent molecular weights. In this work, a Superdex 75 10/300 GL or a Superdex 200 10/300 GL (GE HEALTHCARE) was used with a LaChrom (MERCK-HITACHI) or with an Äkta Basic 10 better (GE HEALTHCARE) chromatographic device. Protein samples were centrifuged (SORVALL RC5B, SS34 rotor, 14000 rpm, 30 min, 4° C) and applied on the pre-equilibrated column. The elution was followed by absorbance measurement at 280 nm (LaChrom) or at 260, 280 and 420 nm (Äkta Basic 10 better).

7.5.5 Analytical reversed-phase chromatography

Reverse-phase (RP) chromatography is used for the separation of analytes according to their polarity. The stationary phase of RP columns consist of silica particles, which are modified by long-chain hydrocarbons. Analytes are retained on the column by hydrophobic interaction with the stationary phase. They are eluted by a decrease in the polarity of the mobile phase. Commonly, a gradient with increasing concentration of organic solvent is used for the separation of analytes. In this work, reactions mixtures were separated by C18 RP HPLC columns using a 1200 HPLC system (Agilent). Elution conditions were optimized to obtain base-line separation of analytes and flow-rates were adjusted to maintain a typical operation pressure between 50 and 200 bar. The following programs were used:

Program 1

solvent A: 0.2 % ($^{w}/_{v}$) sodium bicarbonate in water
solvent B: methanol
flow-rate: 0.25 ml/ min
column temperature: 25° C
column: Gemini-NX 3 µm C18 110A 150x3 mm (PHENOMENEX)
sample: reaction mixture quenched with methanol

time/ min	0	3	33	43	45	60
% B	5	5	100	100	5	5

The retention time was 16 min for tryptophan and 28 min for indole. Calibration curves for indole and tryptophan were done with samples of known concentrations.

Program 2

solvent A: water
solvent B: acetonitrile
flow-rate: 0.6 ml/ min
column temperature: 25° C
column: Zorbax Eclipse 5 µm C18 150x4.6 mm (AGILENT)
sample: reaction mixture quenched with methanol

time/ min	0	15	30	35	40	60
% B	5	40	100	100	5	5

The retention time was 7.8 min for 2S,3R-methyltryptophan and 9.1 min for 2S,3S-methyltryptophan.

Program 3

solvent A: 0.1 % ($^{v}/_{v}$) formic acid in water
solvent B: 0.1 % ($^{v}/_{v}$) formic acid in acetonitrile
flow-rate: 0.35 ml/ min
column temperature: 25° C
column: Gemini-NX 3 µm C18 110A 150x3 mm (PHENOMENEX)
sample: reaction mixture quenched with methanol

time/ min	**0**	**30**	**25.8**	**29.4**	**33.0**	**45.0**
% B	5	5	98	98	5	5

The retention times was 10.2 min for tryptophan and 19.6 min for indole.

Program 4

solvent A: 0.2 % ($^{w}/_{v}$) sodium bicarbonate in water
solvent B: methanol
flow-rate: 0.25 ml/ min
column temperature: 25° C
column: Kromasil 5 µm C18 250x4 mm column (BISCHOFF)
sample: reaction mixture quenched with methanol

time/ min	**0**	**3**	**33**	**43**	**45**	**60**
% B	5	5	100	100	5	5

The retention time was 26.4 min for tryptophan.

The elution was followed by a diode array detector in a range of 210 to 500 nm and by a fluorescence detector in an emission range of 300 to 450 nm after excitation at 278 nm. The elution of labeled analytes was also followed by online scintillation counting with a Ramona Star detector (RAYTEST).

7.5.6 Thin layer chromatography (TLC)

Thin layer chromatography is a convenient technique for the separation of multiple samples in parallel. The stationary phase consists of absorbent material, which is immobilized on a glas or an aluminium plate. Samples are spotted on the plate, which is subsequently placed in a chamber with liquid mobil phase. Initially, the spots are not in direct contact with the mobil phase. During the chromatographic process, the mobil phase is dawn up the plate by capillary force. Compounds co-migrate with the mobile phase in dependency of their interaction with the stationary phase and in dependency of their solubility in the mobile phase. The separation is terminated before the solvent front reaches the top of the stationary phase. After visualization of the separated spots, the retention factor (R_f) is calculated according to Equation 8.

$$R_f = \frac{a}{b}$$

Equation 8: Determination of the retention factor R_f.
a distance from start to center of the spot
b distance from start to solvent front

7.5.7 Phosphorimaging

Radioactive labeled analytes on a TLC plate can be visualized by their ability to excite phosphor material. The stored energy is released as blue-light luminescence by excitation with visible light. In this work, a PhosphorImager Film FLA-3000 (FUJIFILM) was exposed over night and the luminescence was read in a Cyclone phosphor-imager (PACKARD biosciences).

7.5.8 Colorimetric assay for indole quantification

Indole was derivatized with an aldehyde-containing reagent and quantified by absorbance measurement. Ehrlich reagent and Kovacs reagent (both utilizing p-dimethylaminobenzaldehyde) turned out to be insufficiently sensitive for the detection of tiny amount of indole. Thus a p-dimethyl-aminocinnamaldehyde (DMACA) based reagent was used instead. The following procedure was used for the quantification of indole in a hexane solution: DMACA reagent was prepared by solving 0.234 g p-dimethylaminocinnamaldehyde in 78 ml ethanol, adding 10 ml concentrated HCl and

adjusting the volume to 100 ml as been described (Turner, 1961). The indole containing hexan solution (1.5 ml) was thoroughly mixed with 150 µl DMACA reagent and incubated at room temperature for 30 min. The derivatization reaction is shown in Figure 65.

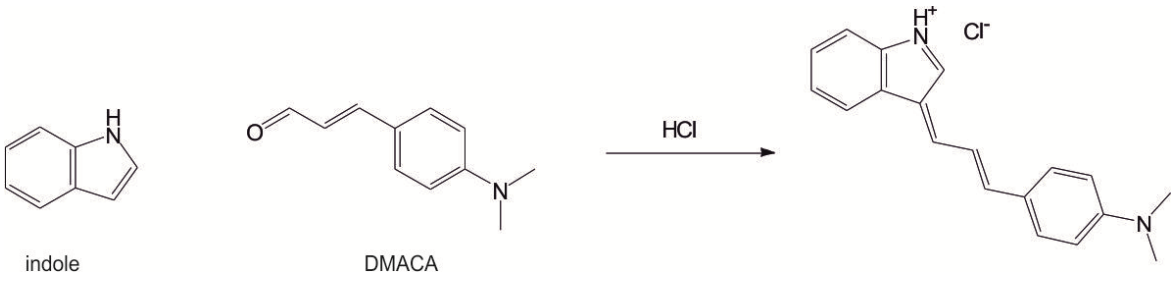


Figure 65: Reaction of DMACA with indole.

100 µl of the aqueous phase were subsequently transferred into a 96 well plate. The absorbance was measured at 640 nm using a M200 infinite pro plate reader (TECAN). A standard curve was made with equally treated indole samples of known concentration to quantify the amount of indole in the hexane solution.

7.5.9 Isothermal titration calorimetry (ITC)

Isothermal titration calorimetry is a method to determine the thermodynamic parameters for the interaction between protein and protein or ligand in solution. Changes in enthalpy upon addition of a titrant to the sample are detected with high accuracy as they are compensated by electric power in order to maintain equal temperatures of the sample cell and the reference cell. Binding affinities, enthalpy changes and binding stoichiometries can be deduced from the compensation power of the titration steps. A MicroCal iTC_{200} system (MALVERN) was used in this work. Prior to measurements, sample and titrant were dialyzed against the same buffer and degassed for at least 30 min. The recorded data was analyzed by Origin 7.

7.5.10 Differential scanning calorimetry (DSC)

Differential scanning calorimetry is a method to determine the thermal stability of proteins in solution. The calorimeter contains the sample cell for a solution of protein in buffer and the reference cell for a solution of buffer. The temperature in both cells is equally increased during the measurement. Protein conformation changes and

protein unfolding result in a temperature difference between the sample cell and the reference cell, which is compensated by electric power. The melting temperature T_m of a protein is deduced from the changes in heat capacity within a temperature range. A MicroCal VP-DSC system (MALVERN) with an operational temperature range of -10° C to 130° C was used in this work. The recorded data was analyzed by Origin 7.

7.5.11 Circular dichroism spectroscopy (CD)

Circular dichroism spectroscopy in the far-UV range is used to analyze the secondary structure of proteins. Right- and left- circular polarized light are absorbed to a different extent when passing through a solution of chiral molecules. As a result, the circular dichroism (CD) signal over the corresponding wavelengths can adopt positive and negative values. Different secondary structure elemens like α-helix, β-sheet and random coil have characteristic CD spectra in the range of 170 - 250 nm. In this work, CD spectra were recorded from 195 to 250 nm with a quartz cuvette (1 mm, HELLMA) in a spectro-polarimeter J-815 (JASCO). The data was normalized to obtain the molecular ellipticity per amino acid according to Equation 9.

$$\theta_{MRW} = \frac{\theta_{obs} \cdot 100 \cdot MRW}{\beta \cdot d} = \frac{\theta_{obs} \cdot 100 \cdot MW}{\beta \cdot d \cdot N_A} = \frac{\theta_{obs} \cdot 10^5}{c \cdot d \cdot N_A}$$

Equation 9: Calculation of normalized ellipticity per amino acid residue.

θ_{MRW}	average ellipticity per amino acid (deg · cm^2 · $dmol^{-1}$)
θ_{obs}	measured ellipticity (mdeg)
MRW	mean residue weight (kDa)
β	protein concentration (mg/ ml)
c	protein concentration (µmol/ l)
d	thickness of cuvette (cm)
MW	molecular weight of protein (kDa)
N_A	number of amino acids

The change in ellipticity at 220 nm with increasing temperatures was observed to determine the thermostability of proteins. For proteins with cooperatve unfolding in the operational temperature range, data points were fitted with a two-state model according to Equation 10.

$$F_u = \frac{y_n - y}{y_n - y_u}$$

Equation 10: Two-state model of unfolding.
Fu fraction of unfolded protein
y_n signal of native protein
y_u signal of denatured protein
y measured signal

7.5.12 Fluorescence spectroscopy

In fluorescence spectroscopy, samples are illuminated by light of a specific wavelength. Light absorption leads to the excitation of electrons in the fluorophore and part of the excitation energy is emitted as light. This light is red-shifted relative to the excitation light. Fluorescence emission spectra are specific for a certain fluorophore and are influenced by the environment of the fluorophore. Protein fluorescence might change in response to an interaction with proteins or ligands. In this work, tryptophan fluorescence and PLP fluorescence were used as probes to quantify protein-protein and protein-ligand interactions by fluorescence spectroscopy. Data points for a high-affinity interaction were fitted with a quadratic function (Reinstein et al., 1990), as shown in Equation 11.

$$F = F_0 + (F_{max} - F_0) \cdot 0{,}5 \cdot \left\{ \left(1 + \frac{L_0 + K_D}{E_0}\right) - \left(\left(1 + \frac{L_0 + K_D}{E_0}\right)^2 - 4 \cdot \frac{L_0}{E_0}\right)^{0{,}5} \right\}$$

Equation 11: Quadratic function for K_d determination.
F measured fluorescence intensity
F_0 initial signal
F_{max} signal at saturation
L_0 concentration of ligand
E_0 concentration of enzyme

7.5.13 Surface plasmon resonance spectroscopy (SPR)

A surface plasmon resonance (SPR) sensor was used as optical mass detector. At a specific resonance angle, monochromatic light leads to an evanescence field into the opposite medium when it is reflected by a metal layer. Changes in the refractive index of this medium affect the resonance angle and can be detected by a position specific diode array detector. The change in resonance angle is proportional to the mass of biomolecules in the medium. A signal of 1000 response units (RU) corresponds to a shift of 0.1° in the resonance angle. This is for instance caused by binding of 1 ng/mm^2 protein to a CM5 chip surface. In this work, a CM5 sensor chip

was used in a two flow-cell operating Biacore X100. The sensor chip has a carboxylmethylated (CM) dextran surface immobilized on a gold layer. The protein of interest is covalently linked to the surface of the samples flow cell. Hereto the carboxyl-group is activated as NHS ester by a treatment with 1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC) and N-Hydroxysuccinimide (NHS). The amines of an added ligand (protein solution in a low ion, moderate low pH and amine-free buffer) react subsequently with the NHS ester and form a covalent amid bound. Remaining NHS ester is de-reacted by treatment with ethanolamine. Samples of an analyte are applied in running buffer. Binding to the immobilized ligand leads to a shift in the resonance angle, which is detected as response units. The plot of response units versus time is termed sensogram and is used to determine the affinity of an analyte-ligand interaction. Bound analytes are removed to regenerate the surface between after each analyte application. The data was analyzed by the implemented software.

7.5.14 Pre-steady-state enzyme kinetics

Rapid (‘pre steady-state) kinetics were analyzed by stopped-flow. This technique enables the mixing of solution from two syringes by a pneumatic device. Mixing in an observation chamber occurs within milliseconds and triggers the spectroscopic detection. A SX20 stopped-flow spectrometer (APPLIED photophysics) equipped with detectors for absorption and fluorescence was used. Both syringes were filled with 1.5 ml of solution and equal amounts were mixed for the measurements. Data was fitted with a single exponential function (Equation 12) or two phase exponential function (Equation 13).

$$A = a \cdot e^{-kt} + c$$

Equation 12: Single exponential function.

A	measured signal
a	amplitude
k	reaction rate
t	time
c	plateau

$$A = a1 \cdot e^{-k1t} + a2 \cdot e^{-k2t} + c$$

Equation 13: Double exponential function.
A measured signal
a1,a2 amplitudes
k1, k2 reaction rates
t time
c plateau

7.5.15 Steady-state enzyme kinetics

TrpA and TrpB kinetics

The TrpA reaction is the reversible aldolytic cleavage of indole-3-glycerol-phosphate to glyceraldehyde-3-phosphate and indole. The reaction was usually measured in a coupled enzymatic assay (Creighton, 1970). Glyceraldehyde-3-phosphate is oxidized to 1-arseno-3-phosphoglycerate by the helper enzyme glyceraldehyde-3-phosphate dehydrogenase from *T. maritima* in the presence of arsenate. The coupled reduction of NAD^+ to $NADH/H^+$ was monitored by the increase in absorption at 340 nm ($\Delta\varepsilon(NADH/H^+\text{-}NAD^+)$ = 6.22 $mM^{-1}cm^{-1}$). The reaction is irreversible as 1-arseno-3-phosphoglycerate is rapidly hydrolyzed to arsenate and 3-phospho-glycerate.

The TrpB reaction is the condensation reaction of indole and L-serine or OPS to tryptophan. The reaction can be monitored by the increase in absorption at 290 nm ($\Delta\varepsilon$(tryptophan-indole) = 1.89 $mM^{-1}cm^{-1}$) (Faeder and Hammes, 1970).

Steady-state kinetic parameters were determined by fitting the experimental data of the saturation curves to a hyperbolic function (Michaelis-Menten equation; Equation 14).

$$v = \frac{v_{max} \cdot S}{K_M + S}$$

Equation 14: Michaelis-Menten equation.
v: initial velocity
v_{max}: maximum velocity
S: substrate concentration
K_M: Michaelis constant

7.6 Protein crystallization and X-ray structure determination

Data collection and X-ray structure determination was performed in collaboration with Dr. Chitra Rajendran (research group of Prof. Dr. Christine Ziegler, University of Regensburg).

7.6.1 Protein crystallization

Protein crystallization was performed by hanging drop vapor diffusion methods. To explore various conditions, initial crystal screening was done using commercially available crystallization kits containing 96 unique conditions. Initial crystal screening was performed using an automatic robotic system Crystallization conditions were manually refined applying an one or two variable (precipitant concentration, additive, pH, protein concentration) grid screen.

7.6.2 Data collection

Single crystals were transferred to a cryo protectant solution and flash frozen in liquid nitrogen. Diffraction data was collected on-site at the Swiss Light Source (SLS; at Paul Scherrer Institute, Villigen, Switzerland).

8 REFERENCES

Agren, D., Schnell, R., Oehlmann, W., Singh, M., and Schneider, G. (2008). Cysteine synthase (CysM) of *Mycobacterium tuberculosis* is an O-phosphoserine sulfhydrylase: evidence for an alternative cysteine biosynthesis pathway in mycobacteria. J. Biol. Chem. 283, 31567-31574.

Ahmed, S.A., McPhie, P., and Miles, E.W. (1996). Mechanism of activation of the tryptophan synthase $\alpha_2\beta_2$ complex. Solvent effects of the co-substrate β-mercaptoethanol. J. Biol. Chem. 271, 29100-29106.

Ahmed, S.A., and Miles, E.W. (1994). Aliphatic alcohols stabilize an alternative conformation of the tryptophan synthase $\alpha_2\beta_2$ complex from *Salmonella typhimurium*. J. Biol. Chem. 269, 16486-16492.

Ahmed, S.A., Ruvinov, S.B., Kayastha, A.M., and Miles, E.W. (1991). Mechanism of mutual activation of the tryptophan synthase α- and β-subunits. Analysis of the reaction specificity and substrate-induced inactivation of active site and tunnel mutants of the β-subunit. J. Biol. Chem. 266, 21548-21557.

Akanuma, S., Nakajima, Y., Yokobori, S., Kimura, M., Nemoto, N., Mase, T., Miyazono, K., Tanokura, M., and Yamagishi, A. (2013). Experimental evidence for the thermophilicity of ancestral life. Proc. Natl. Acad. Sci. U S A 110, 11067-11072.

Anderson, K.S., Miles, E.W., and Johnson, K.A. (1991). Serine modulates substrate channeling in tryptophan synthase. A novel intersubunit triggering mechanism. J. Biol. Chem. 266, 8020-8033.

Atkinson, H.J., Morris, J.H., Ferrin, T.E., and Babbitt, P.C. (2009). Using sequence similarity networks for visualization of relationships across diverse protein superfamilies. PLoS One 4, e4345.

Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W., Davidi, D., Tawfik, D.S., and Milo, R. (2011). The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. Biochemistry 50, 4402-4410.

Barends, T.R., Domratcheva, T., Kulik, V., Blumenstein, L., Niks, D., Dunn, M.F., and Schlichting, I. (2008). Structure and mechanistic implications of a tryptophan synthase quinonoid intermediate. ChemBioChem 9, 1024-1028.

Bennett, B.D., Kimball, E.H., Gao, M., Osterhout, R., Van Dien, S.J., and Rabinowitz, J.D. (2009). Absolute metabolite concentrations and implied enzyme active site occupancy in *Escherichia coli*. Nat. Chem. Biol. 5, 593-599.

Bolz, C. (2006). Mutationsanalyse zur Identifizierung der strukturellen Grundlage einer spezifischen Protein-Protein Interaktion. Diploma thesis. University of Regensburg.

Brown, S.D., and Babbitt, P.C. (2012). Inference of functional properties from large-scale analysis of enzyme superfamilies. J. Biol. Chem. 287, 35-42.

Brzovic, P.S., Kayastha, A.M., Miles, E.W., and Dunn, M.F. (1992). Substitution of glutamic acid 109 by aspartic acid alters the substrate specificity and catalytic activity of the β-subunit in the tryptophan synthase bienzyme complex from *Salmonella typhimurium*. Biochemistry 31, 1180-1190.

Bunders, C.A., Minvielle, M.J., Worthington, R.J., Ortiz, M., Cavanagh, J., and Melander, C. (2011). Intercepting bacterial indole signaling with flustramine derivatives. J. Am. Chem. Soc. 133, 20160-20163.

Burkhard, P., Tai, C.H., Ristroph, C.M., Cook, P.F., and Jansonius, J.N. (1999). Ligand binding induces a large conformational change in O-acetylserine sulfhydrylase from *Salmonella typhimurium*. J. Mol. Biol. 291, 941-953.

Busch, F. (2010). Funktionsaufklärung des zur Tryptophansynthase homologen Enzyms sTrpB2a. Diploma thesis. University of Regensburg.

Busch, F., Rajendran, C., Mayans, O., Löffler, P., Merkl, R., and Sterner, R. (2014). TrpB2 enzymes are O-phospho-L-serine dependent tryptophan synthases. Biochemistry 53, 6078-6083.

Chandrasekharan, U.M., Sanker, S., Glynias, M.J., Karnik, S.S., and Husain, A. (1996). Angiotensin II-forming activity in a reconstructed ancestral chymase. Science 271, 502-505.

Chant, E.L., and Summers, D.K. (2007). Indole signalling contributes to the stable maintenance of *Escherichia coli* multicopy plasmids. Mol. Microbiol. 63, 35-43.

Chattoraj, D.K. (2007). Tryptophanase in sRNA control of the *Escherichia coli* cell cycle. Mol. Microbiol. 63, 1-3.

Chester, N., and Marshak, D.R. (1993). Dimethyl sulfoxide-mediated primer T_m reduction: a method for analyzing the role of renaturation temperature in the polymerase chain reaction. Anal. Biochem. 209, 284-290.

Chevalier, B.S., Kortemme, T., Chadsey, M.S., Baker, D., Monnat, R.J., and Stoddard, B.L. (2002). Design, activity, and structure of a highly specific artificial endonuclease. Mol. Cell 10, 895-905.

Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C.J., Snel, B., and Bork, P. (2006). Toward automatic reconstruction of a highly resolved tree of life. Science 311, 1283-1287.

Conant, G.C., and Wolfe, K.H. (2008). Turning a hobby into a job: how duplicated genes find new functions. Nat. Rev. Genet. 9, 938-950.

Creighton, T.E. (1970). A steady-state kinetic investigation of the reaction mechanism of the tryptophan synthetase of *Escherichia coli*. Eur. J. Biochem. 13, 1-10.

Creighton, T.E., and Yanofsky, C. (1966). Association of the α and β_2 subunits of the tryptophan synthetase of *Escherichia coli*. J. Biol. Chem. 241, 980-990.

Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. Genome Res. 14, 1188-1190.

Dewitz, D. (2013). Charakterisierung der Interaktion zwischen TrpA und TrpB2a Enzymen. Bachelor thesis. University of Regensburg.

di Guan, C., Li, P., Riggs, P.D., and Inouye, H. (1988). Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. Gene 67, 21-30.

Dunn, M.F. (2012). Allosteric regulation of substrate channeling and catalysis in the tryptophan synthase bienzyme complex. Arch. Biochem. Biophys. 519, 154-166.

Dunn, M.F., Aguilar, V., Brzovic, P., Drewe, W.F., Jr., Houben, K.F., Leja, C.A., and Roy, M. (1990). The tryptophan synthase bienzyme complex transfers indole between the α- and β-sites via a 25-30 Å long tunnel. Biochemistry 29, 8598-8607.

Dunn, M.F., Niks, D., Ngo, H., Barends, T.R., and Schlichting, I. (2008). Tryptophan synthase: the workings of a channeling nanomachine. Trends Biochem. Sci. 33, 254-264.

Ehrmann, A. (2011). Strukturelle Analyse neuartiger Tryptophan-Synthasen und funktionelle Charakterisierung ihrer Homologen. PhD thesis. University of Regensburg.

Ehrmann, A., Richter, K., Busch, F., Reimann, J., Albers, S.V., and Sterner, R. (2010). Ligand-induced formation of a transient tryptophan synthase complex with $\alpha\beta\beta$-subunit stoichiometry. Biochemistry 49, 10842-10853.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486-501.

Faeder, E.J., and Hammes, G.G. (1970). Kinetic studies of tryptophan synthetase. Interaction of substrates with the β-subunit. Biochemistry 9, 4043-4049.

Fan, Y.X., McPhie, P., and Miles, E.W. (1999). Guanidine hydrochloride exerts dual effects on the tryptophan synthase $\alpha_2\beta_2$-complex as a cation activator and as a modulator of the active site conformation. Biochemistry 38, 7881-7890.

Fan, Y.X., McPhie, P., and Miles, E.W. (2000). Regulation of tryptophan synthase by temperature, monovalent cations, and an allosteric ligand. Evidence from arrhenius plots, absorption spectra, and primary kinetic isotope effects. Biochemistry 39, 4692-4703.

Fan, Y.X., McPhie, P., and Miles, E.W. (2000). Thermal repair of tryptophan synthase mutations in a regulatory intersubunit salt bridge. Evidence from arrhenius plots, absorption spectra, and primary kinetic isotope effects. J. Biol. Chem. 275, 20302-20307.

Ferrari, D., Niks, D., Yang, L.H., Miles, E.W., and Dunn, M.F. (2003). Allosteric communication in the tryptophan synthase bienzyme complex: roles of the β-subunit aspartate 305-arginine 141 salt bridge. Biochemistry 42, 7807-7818.

Ferrari, D., Yang, L.H., Miles, E.W., and Dunn, M.F. (2001). Beta D305A mutant of tryptophan synthase shows strongly perturbed allosteric regulation and substrate specificity. Biochemistry 40, 7421-7432.

Finnigan, G.C., Hanson-Smith, V., Stevens, T.H., and Thornton, J.W. (2012). Evolution of increased complexity in a molecular machine. Nature 481, 360-364.

Gerlt, J.A., Bouvier, J.T., Davidson, D.B., Imker, H.J., Sadkhin, B., Slater, D.R., and Whalen, K.L. (2015). Enzyme Function Initiative-Enzyme Similarity Tool (EFI-EST): A web tool for generating protein sequence similarity networks. Biochim. Biophys. Acta 1854, 1019-1037.

Goldberg, M.E., Creighton, T.E., Baldwin, R.L., and Yanofsky, C. (1966). Subunit structure of the tryptophan synthetase of *Escherichia coli*. J. Mol. Biol. 21, 71-82.

Hart, K.M., Harms, M.J., Schmidt, B.H., Elya, C., Thornton, J.W., and Marqusee, S. (2014). Thermodynamic system drift in protein evolution. PLoS Biol. 12, e1001994.

Helgadottir, S., Rosas-Sandoval, G., Soll, D., and Graham, D.E. (2007). Biosynthesis of phosphoserine in the Methanococcales. J. Bacteriol. 189, 575-582.

Hettwer, S., and Sterner, R. (2002). A novel tryptophan synthase β-subunit from the hyperthermophile *Thermotoga maritima*. Quaternary structure, steady-state kinetics, and putative physiological role. J. Biol. Chem. 277, 8194-8201.

Hiyama, T., Sato, T., Imanaka, T., and Atomi, H. (2014). The tryptophan synthase β-subunit paralogs TrpB1 and TrpB2 in *Thermococcus kodakarensis* are both involved in tryptophan biosynthesis and indole salvage. FEBS J. 281, 3113-3125.

Ho, S.N., Hunt, H.D., Horton, R.M., Pullen, J.K., and Pease, L.R. (1989). Site-directed mutagenesis by overlap extension using the polymerase chain reaction. Gene 77, 51-59.

Hobbs, J.K., Shepherd, C., Saul, D.J., Demetras, N.J., Haaning, S., Monk, C.R., Daniel, R.M., and Arcus, V.L. (2012). On the origin and evolution of thermophily: reconstruction of functional precambrian enzymes from ancestors of *Bacillus*. Mol. Biol. Evol. 29, 825-835.

Howard, P.L., Chia, M.C., Del Rizzo, S., Liu, F.F., and Pawson, T. (2003). Redirecting tyrosine kinase signaling to an apoptotic caspase pathway through chimeric adaptor proteins. Proc. Natl. Acad. Sci. U S A 100, 11267-11272.

Hu, M., Zhang, C., Mu, Y., Shen, Q., and Feng, Y. (2010). Indole affects biofilm formation in bacteria. Indian J. Microbiol. 50, 362-368.

Huang, X., Holden, H.M., and Raushel, F.M. (2001). Channeling of substrates and intermediates in enzyme-catalyzed reactions. Annu. Rev. Biochem. 70, 149-180.

Hur, O., Niks, D., Casino, P., and Dunn, M.F. (2002). Proton transfers in the β-reaction catalyzed by tryptophan synthase. Biochemistry 41, 9991-10001.

Hyde, C.C., Ahmed, S.A., Padlan, E.A., Miles, E.W., and Davies, D.R. (1988). Three-dimensional structure of the tryptophan synthase $\alpha_2\beta_2$ multienzyme complex from *Salmonella typhimurium*. J. Biol. Chem. 263, 17857-17871.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. Gene 96, 23-28.

Isupov, M.N., Antson, A.A., Dodson, E.J., Dodson, G.G., Dementieva, I.S., Zakomirdina, L.N., Wilson, K.S., Dauter, Z., Lebedev, A.A., and Harutyunyan, E.H. (1998). Crystal structure of tryptophanase. J. Mol. Biol. 276, 603-623.

Janda, J.O., Popal, A., Bauer, J., Busch, M., Klocke, M., Spitzer, W., Keller, J., and Merkl, R. (2014). H2rs: deducing evolutionary and functionally important residue positions by means of an entropy and similarity based analysis of multiple sequence alignments. BMC Bioinformatics 15, 118.

Jensen, R.A. (1976). Enzyme recruitment in evolution of new function. Annu. Rev. Microbiol. 30, 409-425.

Jhee, K.H., Yang, L.H., Ahmed, S.A., McPhie, P., Rowlett, R., and Miles, E.W. (1998). Mutation of an active site residue of tryptophan synthase (β-serine 377) alters cofactor chemistry. J. Biol. Chem. 273, 11417-11422.

Karr, E.A., Sandman, K., Lurz, R., and Reeve, J.N. (2008). TrpY regulation of *trp*B2 transcription in *Methanothermobacter thermautotrophicus*. J. Bacteriol. 190, 2637-2641.

Knöchel, T., Ivens, A., Hester, G., Gonzalez, A., Bauerle, R., Wilmanns, M., Kirschner, K., and Jansonius, J.N. (1999). The crystal structure of anthranilate synthase from *Sulfolobus solfataricus*: Functional implications. Proc. Natl. Acad. Sci. U S A 96, 9479-9484.

Kortemme, T., and Baker, D. (2004). Computational design of protein-protein interactions. Curr. Opin. Chem. Biol. 8, 91-97.

Ku, S.Y., Yip, P., and Howell, P.L. (2006). Structure of *Escherichia coli* tryptophanase. Acta Crystallogr. D Biol. Crystallogr. 62, 814-823.

Kuang, D., Yao, Y., Maclean, D., Wang, M., Hampson, D.R., and Chang, B.S. (2006). Ancestral reconstruction of the ligand-binding pocket of Family C G protein-coupled receptors. Proc. Natl. Acad. Sci. U S A 103, 14050-14055.

Kulik, V., Hartmann, E., Weyand, M., Frey, M., Gierl, A., Niks, D., Dunn, M.F., and Schlichting, I. (2005). On the structural basis of the catalytic mechanism and the regulation of the α-subunit of tryptophan synthase from *Salmonella typhimurium* and BX1 from maize, two evolutionarily related enzymes. J. Mol. Biol. 352, 608-620.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227, 680-685.

Lane, A.N., Paul, C.H., and Kirschner, K. (1984). The mechanism of self-assembly of the multi-enzyme complex tryptophan synthase from *Escherichia coli*. EMBO J. 3, 279-287.

Le Gouill, C., Parent, J.L., Rola-Pleszczynski, M., and Stankova, J. (1994). Analysis of recombinant plasmids by a modified alkaline lysis method. Anal. Biochem. 219, 164.

Leja, C.A., Woehl, E.U., and Dunn, M.F. (1995). Allosteric linkages between β-site covalent transformations and α-site activation and deactivation in the tryptophan synthase bienzyme complex. Biochemistry 34, 6552-6561.

Lelong, C., Aguiluz, K., Luche, S., Kuhn, L., Garin, J., Rabilloud, T., and Geiselmann, J. (2007). The Crl-RpoS regulon of *Escherichia coli*. Mol. Cell. Proteomics 6, 648-659.

Leopoldseder, S., Hettwer, S., and Sterner, R. (2006). Evolution of multi-enzyme complexes: the case of tryptophan synthase. Biochemistry 45, 14111-14119.

Letunic, I., and Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics 23, 127-128.

Letunic, I., and Bork, P. (2011). Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. Nucleic Acids Res. 39, W475-478.

Malcolm, B.A., Wilson, K.P., Matthews, B.W., Kirsch, J.F., and Wilson, A.C. (1990). Ancestral lysozymes reconstructed, neutrality tested, and thermostability linked to hydrocarbon packing. Nature 345, 86-89.

Merkl, R. (2007). Modelling the evolution of the archeal tryptophan synthase. BMC Evol. Biol. 7, 59.

Meyer, J., Walker-Jonah, A., and Hollenberg, C.P. (1991). Galactokinase encoded by GAL1 is a bifunctional protein required for induction of the GAL genes in *Kluyveromyces lactis* and is able to suppress the *gal*3 phenotype in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 11, 5454-5461.

Miles, E.W. (2013). The tryptophan synthase $\alpha_2\beta_2$-complex: a model for substrate channeling, allosteric communication, and pyridoxal phosphate catalysis. J. Biol. Chem. 288, 10084-10091.

Miles, E.W. (2001). Tryptophan synthase: a multienzyme complex with an intramolecular tunnel. Chem. Rec. 1, 140-151.

Mino, K., and Ishikawa, K. (2003). A novel O-phospho-L-serine sulfhydrylation reaction catalyzed by O-acetylserine sulfhydrylase from *Aeropyrum pernix* K1. FEBS Lett. 551, 133-138.

Mitchell, A., Chang, H.Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., McAnulla, C., McMenamin, C., Nuka, G., Pesseat, S., et al. (2015). The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res. 43, D213-221.

Morollo, A.A., and Eck, M.J. (2001). Structure of the cooperative allosteric anthranilate synthase from *Salmonella typhimurium*. Nat. Struct. Biol. 8, 243-247.

Mueller, R.S., Beyhan, S., Saini, S.G., Yildiz, F.H., and Bartlett, D.H. (2009). Indole acts as an extracellular cue regulating gene expression in *Vibrio cholerae*. J. Bacteriol. 191, 3504-3516.

Mullis, K.B., and Faloona, F.A. (1987). Specific synthesis of DNA *in vitro* via a polymerase-catalyzed chain reaction. Methods in Enzymology 155, 335-350.

Nakamura, T., Kawai, Y., Kunimoto, K., Iwasaki, Y., Nishii, K., Kataoka, M., and Ishikawa, K. (2012). Structural analysis of the substrate recognition mechanism in O-phosphoserine sulfhydrylase from the hyperthermophilic archaeon *Aeropyrum pernix* K1. J. Mol. Biol. 422, 33-44.

Niks, D., Hilario, E., Dierkers, A., Ngo, H., Borchardt, D., Neubauer, T.J., Fan, L., Mueller, L.J., and Dunn, M.F. (2013). Allostery and substrate channeling in the tryptophan synthase bienzyme complex: evidence for two subunit conformations and four quaternary states. Biochemistry 52, 6396-6411.

Oda, Y., Mino, K., Ishikawa, K., and Ataka, M. (2005). Three-dimensional structure of a new enzyme, O-phosphoserine sulfhydrylase, involved in L-cysteine biosynthesis by a hyperthermophilic archaeon, *Aeropyrum pernix* K1, at 2.0 A resolution. J. Mol. Biol. 351, 334-344.

Osborne, A., Teng, Q., Miles, E.W., and Phillips, R.S. (2003). Detection of open and closed conformations of tryptophan synthase by ^{15}N-heteronuclear single-quantum coherence nuclear magnetic resonance of bound 1-^{15}N-L-tryptophan. J. Biol. Chem. 278, 44083-44090.

Pace, C.N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). How to measure and predict the molar absorption coefficient of a protein. Protein Sci. 4, 2411-2423.

Pan, P., and Dunn, M.F. (1996). β-site covalent reactions trigger transitions between open and closed conformations of the tryptophan synthase bienzyme complex. Biochemistry 35, 5002-5013.

Pan, P., Woehl, E., and Dunn, M.F. (1997). Protein architecture, dynamics and allostery in tryptophan synthase channeling. Trends Biochem. Sci. 22, 22-27.

Pandya, C., Farelli, J.D., Dunaway-Mariano, D., and Allen, K.N. (2014). Enzyme promiscuity: engine of evolutionary innovation. J. Biol. Chem. 289, 30229-30236.

Park, S.Y., Beel, B.D., Simon, M.I., Bilwes, A.M., and Crane, B.R. (2004). In different organisms, the mode of interaction between two signaling proteins is not necessarily conserved. Proc. Natl. Acad. Sci. U S A 101, 11646-11651.

Parsot, C. (1987). A common origin for enzymes involved in the terminal step of the threonine and tryptophan biosynthetic pathways. Proc. Natl. Acad. Sci. U S A 84, 5207-5210.

Pazos, F., and Valencia, A. (2008). Protein co-evolution, co-adaptation and interactions. EMBO J. 27, 2648-2655.

Perez-Jimenez, R., Inglés-Prieto, A., Zhao, Z.M., Sanchez-Romero, I., Alegre-Cebollada, J., Kosuri, P., Garcia-Manyes, S., Kappock, T.J., Tanokura, M., Holmgren, A., et al. (2011). Single-molecule paleoenzymology probes the chemistry of resurrected enzymes. Nat. Struct. Mol. Biol. 18, 592-596.

Raboni, S., Bettati, S., and Mozzarelli, A. (2005). Identification of the geometric requirements for allosteric communication between the α- and β-subunits of tryptophan synthase. J Biol Chem 280, 13450-13456.

Reinstein, J., Vetter, I.R., Schlichting, I., Rosch, P., Wittinghofer, A., and Goody, R.S. (1990). Fluorescence and NMR investigations on the ligand binding properties of adenylate kinases. Biochemistry 29, 7440-7450.

Reisinger, B., Sperl, J., Holinski, A., Schmid, V., Rajendran, C., Carstensen, L., Schlee, S., Blanquart, S., Merkl, R., and Sterner, R. (2014). Evidence for the existence of elaborate enzyme complexes in the Paleoarchean era. J. Am. Chem. Soc. 136, 122-129.

Rhee, S., Parris, K.D., Hyde, C.C., Ahmed, S.A., Miles, E.W., and Davies, D.R. (1997). Crystal structures of a mutant (βK87T) tryptophan synthase $\alpha_2\beta_2$-complex with ligands bound to the active sites of the α- and β-subunits reveal ligand-induced conformational changes. Biochemistry 36, 7664-7680.

Risso, V.A., Gavira, J.A., Mejia-Carmona, D.F., Gaucher, E.A., and Sanchez-Ruiz, J.M. (2013). Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian beta-lactamases. J. Am. Chem. Soc. 135, 2899-2902.

Rowlett, R., Yang, L.H., Ahmed, S.A., McPhie, P., Jhee, K.H., and Miles, E.W. (1998). Mutations in the contact region between the α– and β–subunits of tryptophan synthase alter subunit interaction and intersubunit communication. Biochemistry 37, 2961-2968.

Russell, R.B., Alber, F., Aloy, P., Davis, F.P., Korkin, D., Pichaud, M., Topf, M., and Sali, A. (2004). A structural perspective on protein-protein interactions. Curr. Opin. Struct. Biol. 14, 313-324.

Ruvinov, S.B., Yang, X.J., Parris, K.D., Banik, U., Ahmed, S.A., Miles, E.W., and Sackett, D.L. (1995). Ligand-mediated changes in the tryptophan synthase indole tunnel probed by nile red fluorescence with wild type, mutant, and chemically modified enzymes. J. Biol. Chem. 270, 6357-6369.

Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239, 487-491.

Sato, T., Fukui, T., Atomi, H., and Imanaka, T. (2005). Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon *Thermococcus kodakaraensis*. Appl. Environ. Microbiol. 71, 3889-3899.

Sato, T., Fukui, T., Atomi, H., and Imanaka, T. (2003). Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon *Thermococcus kodakaraensis* KOD1. J. Bacteriol. 185, 210-220.

Schiaretti, F., Bettati, S., Viappiani, C., and Mozzarelli, A. (2004). pH dependence of tryptophan synthase catalytic mechanism: I. The first stage, the β-elimination reaction. J. Biol. Chem. 279, 29572-29582.

Schneider, T.D., and Stephens, R.M. (1990). Sequence logos: a new way to display consensus sequences. Nucleic Acids Res. 18, 6097-6100.

Schnell, R., Sriram, D., and Schneider, G. (2014). Pyridoxal-phosphate dependent mycobacterial cysteine synthases: Structure, mechanism and potential as drug targets. Biochim. Biophys. Acta.

Schupfner, M. (2011). Rationales Design einer Protein-Protein Interaktion. Bachelor thesis. University of Regensburg.

Schupfner, M. (2014). Vergleichende Analyse von Tryptophan Synthase Komplexen. Master thesis. University of Regensburg.

Seidel, P. (2012). Charakterisierung von TrpB2-Enzymen. Bachelor Thesis. University of Regensburg.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498-2504.

Sharp, P.A., Sugden, B., and Sambrook, J. (1973). Detection of two restriction endonuclease activities in *Haemophilus parainfluenzae* using analytical agarose-ethidium bromide electrophoresis. Biochemistry 12, 3055-3063.

Spraggon, G., Kim, C., Nguyen-Huu, X., Yee, M.C., Yanofsky, C., and Mills, S.E. (2001). The structures of anthranilate synthase of *Serratia marcescens* crystallized in the presence of (*i*) its substrates, chorismate and glutamine, and a product, glutamate, and (*ii*) its end-product inhibitor, L-tryptophan. Proc. Natl. Acad. Sci. U S A 98, 6021-6026.

Stackhouse, J., Presnell, S.R., McGeehan, G.M., Nambiar, K.P., and Benner, S.A. (1990). The ribonuclease from an extinct bovid ruminant. FEBS Lett. 262, 104-106.

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

Thomson, J.M., Gaucher, E.A., Burgan, M.F., De Kee, D.W., Li, T., Aris, J.P., and Benner, S.A. (2005). Resurrecting ancestral alcohol dehydrogenases from yeast. Nat. Genet. 37, 630-635.

Turner, J.M. (1961). A new reagent for the assay of indole in the tryptophanase reaction. Biochem J. 78, 790-792.

Voordeckers, K., Brown, C.A., Vanneste, K., van der Zande, E., Voet, A., Maere, S., and Verstrepen, K.J. (2012). Reconstruction of ancestral metabolic enzymes reveals molecular mechanisms underlying evolutionary innovation through gene duplication. PLoS Biol. 10, e1001446.

Wang, D., Ding, X., and Rather, P.N. (2001). Indole can act as an extracellular signal in *Escherichia coli*. J. Bacteriol. 183, 4210-4216.

Wang, W., and Malcolm, B.A. (1999). Two-stage PCR protocol allowing introduction of multiple mutations, deletions and insertions using QuikChange Site-Directed Mutagenesis. BioTechniques 26, 680-682.

Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M., and Barton, G.J. (2009). Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. Bioinformatics 25, 1189-1191.

Weischet, W.O., and Kirschner, K. (1976). Steady-state kinetic studies of the synthesis of indoleglycerol phosphate catalyzed by the α–subunit of tryptophan synthase from *Escherichia coli*. Comparison with the $\alpha_2\beta_2$-complex. Eur. J. Biochem. 65, 375-385.

Wilson, C., Agafonov, R.V., Hoemberger, M., Kutter, S., Zorba, A., Halpin, J., Buosi, V., Otten, R., Waterman, D., Theobald, D.L., et al. (2015). Kinase dynamics. Using ancient protein kinases to unravel a modern cancer drug's mechanism. Science 347, 882-886.

Wilson, G.G., and Murray, N.E. (1991). Restriction and modification systems. Annu. Rev. Genet. 25, 585-627.

Xie, G., Forst, C., Bonner, C., and Jensen, R.A. (2002). Significance of two distinct types of tryptophan synthase beta chain in Bacteria, Archaea and higher plants. Genome Biol. 3, RESEARCH0004.

Yanofsky, C., and Rachmeler, M. (1958). The exclusion of free indole as an intermediate in the biosynthesis of tryptophan in *Neurospora crassa*. Biochim. Biophys. Acta 28, 640-641.

Yin, R., Frey, M., Gierl, A., and Glawischnig, E. (2010). Plants contain two distinct classes of functional tryptophan synthase β-proteins. Phytochemistry 71, 1667-1672.

Yokoyama, S., Xing, J., Liu, Y., Faggionato, D., Altun, A., and Starmer, W.T. (2014). Epistatic adaptive evolution of human color vision. PLoS Genet. 10, e1004884.

Zhang, Q.C., Petrey, D., Norel, R., and Honig, B.H. (2010). Protein interface conservation across structure space. Proc. Natl. Acad. Sci. U S A 107, 10896-10901.

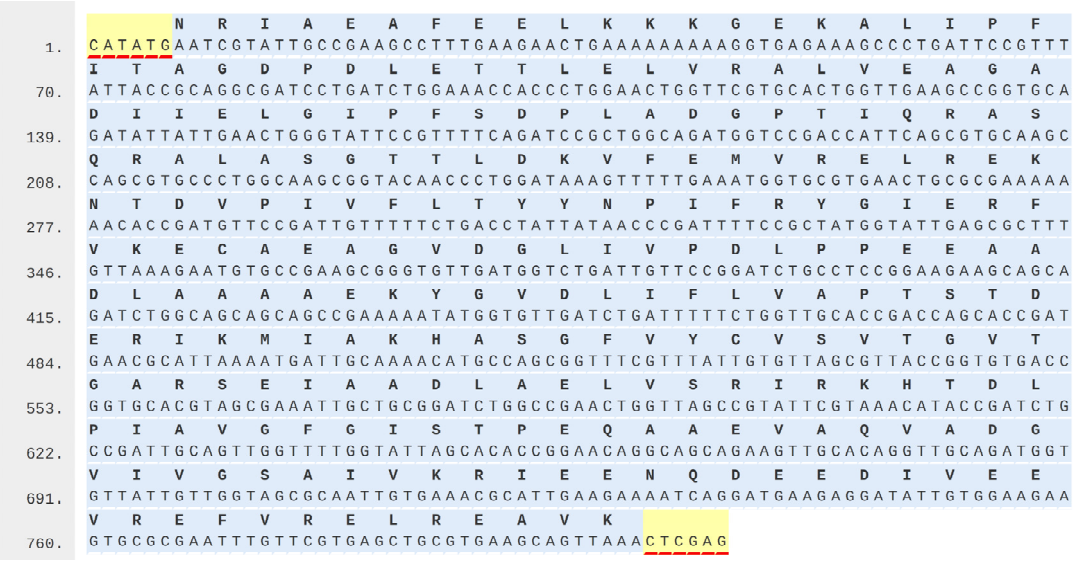
Zhao, S., Kumar, R., Sakai, A., Vetting, M.W., Wood, B.M., Brown, S., Bonanno, J.B., Hillerich, B.S., Seidel, R.D., Babbitt, P.C., et al. (2013). Discovery of new enzymes and metabolic pathways by using structure and genome context. Nature 502, 698-702.

Zou, Y., Fang, Q., Yin, H., Liang, Z., Kong, D., Bai, L., Deng, Z., and Lin, S. (2013). Stereospecific biosynthesis of β-methyltryptophan from L-tryptophan features a stereochemical switch. Angew. Chem. Int. Ed. Engl. 52, 12951-12955.

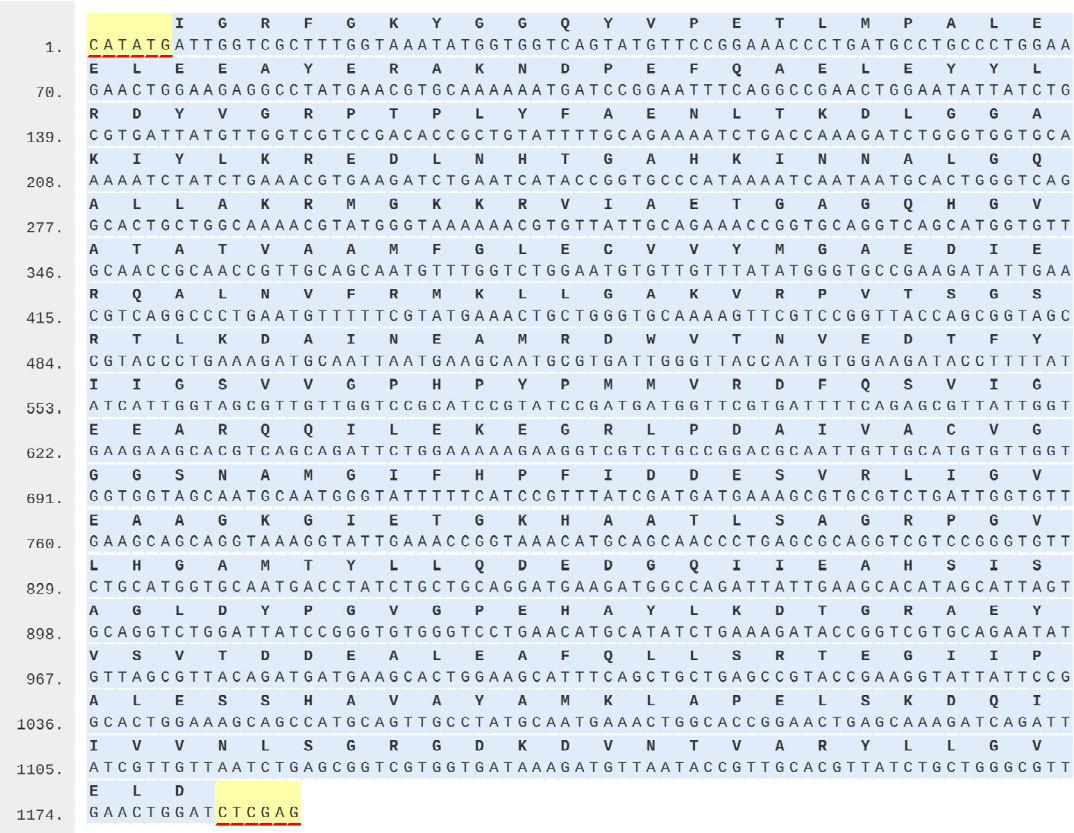
9 APPENDIX

9.1 Sequences of LCA *trp*A and LCA *trp*B

LCA *trp*A

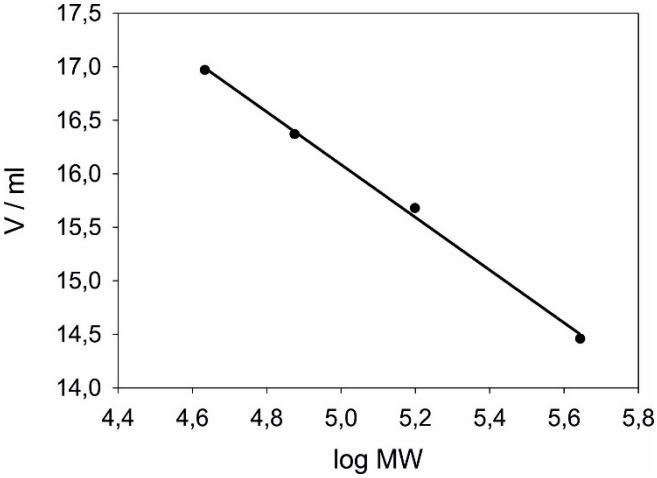


LCA trpB



9.2 Calibration curve

Calibration of S200 column



y = -2.632x + 29.286

Protein	MW (Da)	log MW	V (ml)
Thyroglobulin	668000	5.82	12.58
Ferritin	440000	5.64	14.46
Aldolase	158000	5.20	15.68
Conalbumin	75000	4.88	16.37
Ovalbumin	43000	4.63	16.97
Ribonuclease A	13700	4.14	18.50

Blue Dextran	**8.16 ml**
Bed Volume	23.56 ml

The calibration curve was done by Mona Linde.

9.3 Organisms with TrpA2 enzyme

Pyrobaculum calidifontis (strain JCM 11548 / VA1).
Thermoplasmatales archaeon A-plasma.
Metallosphaera yellowstonensis MK1.
Thermoproteus uzoniensis (strain 768-20).
Caldivirga maquilingensis (strain ATCC 700844 / DSM 13496 / JCM 10307/ IC-167).
Ignisphaera aggregans (strain DSM 17230 / JCM 13409 / AQ1.S1).
Metallosphaera sedula (strain ATCC 51363 / DSM 5348).|Metallosphaera cuprina (strain Ar-4).
Sulfolobales archaeon AZ1.
Thermoplasmatales archaeon Gpl.
Pyrobaculum aerophilum (strain ATCC 51768 / IM2 / DSM 7523 / JCM 9630/ NBRC 100827).
Ferroplasma acidarmanus fer1.
Pyrobaculum sp. 1860.
Ignicoccus hospitalis (strain KIN4/I / DSM 18386 / JCM 14125).
Thermoproteus tenax (strain ATCC 35583 / NBRC 100435 / JCM 9277 / Kra1).
Candidatus Caldiarchaeum subterraneum.
Pyrobaculum arsenaticum (strain DSM 13514 / JCM 11321).|Pyrobaculum oguniense (strain DSM 13380 / JCM 10595 / TE7).
Pyrobaculum islandicum (strain DSM 4184 / JCM 9189).|Pyrobaculum neutrophilum (strain DSM 2338 / JCM 9278 / V24Sta)(Thermoproteus neutrophilus).
Acidianus hospitalis (strain W1).|Candidatus Acidianus copahuensis.
Sulfolobus tokodaii (strain DSM 16993 / JCM 10545 / NBRC 100140 / 7).
Picrophilus torridus (strain ATCC 700027 / DSM 9790 / JCM 10055 / NBRC100828).
Thermoplasma volcanium (strain ATCC 51530 / DSM 4299 / JCM 9571 / NBRC15438 / GSS1).
Sulfolobus islandicus (strain L.S.2.15 / Lassen #1).|Sulfolobus islandicus (strain M.14.25 / Kamchatka #1).|Sulfolobus islandicus (strain M.16.27).|Sulfolobus islandicus (strain Y.G.57.14 / Yellowstone #1).|Sulfolobus islandicus (strain Y.N.15.51 / Yellowstone #2).|Sulfolobus islandicus (strain M.16.4 / Kamchatka #3).|Sulfolobus solfataricus (strain 98/2).|Sulfolobus islandicus (strain L.D.8.5 / Lassen #2).|Sulfolobus islandicus (strain REY15A).|Sulfolobus islandicus (strain HVE10/4).|Sulfolobus islandicus LAL14/1.|Sulfolobus solfataricus (strain ATCC 35092 / DSM 1617 / JCM 11322 /P2).
Sulfolobus acidocaldarius N8.|Sulfolobus acidocaldarius Ron12/I.|Sulfolobus acidocaldarius (strain ATCC 33909 / DSM 639 / JCM 8929 /NBRC 15157 / NCIMB 11770).|Sulfolobus acidocaldarius SUSAZ.
Aeropyrum pernix (strain ATCC 700893 / DSM 11879 / JCM 9820 / NBRC100138 / K1).
Thermoplasmatales archaeon E-plasma.
Pyrolobus fumarii (strain DSM 11204 / 1A).
Vulcanisaeta distributa (strain DSM 14429 / JCM 11212 / NBRC 100878 /IC-017).|Vulcanisaeta moutnovskia (strain 768-28).
Caldisphaera lagunensis (strain DSM 15908 / JCM 11604 / IC-154).

10 ACKNOWLEDGEMENTS

Foremost, I would like to thank my PhD advisor Prof. Dr. Reinhard Sterner for his guidance and support. I deeply appreciate your mentorship and the great opportunity to gain expertise and to develop on a personal level.

Besides my advisor, I would like to thank the other members of the thesis committee: Prof. Dr. Reinhard Wirth, Prof. Dr. Rainer Merkl and Prof. Dr. Christoph Oberprieler.

I further like to thank my collaboration partners for their valuable contributions to this work: I cordially thank Dr. Chitra Rajendran (University of Regensburg) for her great work on X-ray structure determination. I thank Prof. Dr. Olga Mayans for providing the pre-liminary structure of the *S. solfataricus* TS.

I gratefully thank Dr. Sandra Schlee for support in the evaluation of kinetic data, valuable discussion and help with the stopped-flow measurements. Many thanks to Prof. Dr. Rainer Merkl, Patrick Löffler and Kristina Heyn for their ideas and support.

I express my deepest gratitude to Hermine Reisner as well as Christiane Endres, Sonja Fuchs and Jeannette Ueckert for technical assistance and constant support. Thanks to Claudia Pauer for administrative support.

Many thanks go to my rotation students Dietmar Dewitz, Patricia Seidel, Michael Schupfner and Sina Wittmann for their interest, enthusiasm, and their contribution to this work.

I would also like to thank all current and former members of the Sterner research group for their support and valuable discussions. Especially, I like to thank my former supervisor Dr. Alexander Ehrmann as well as my former and present colleagues Dr. Patrick Babinger, Dr. Harald Guldan, Dr. Monika Meier, Dr. David Peterhoff, Dr. Bernd Reisinger, Dr. Daniel Schneider, Alexandra Holinski, Wolfgang Kaiser, Mona Linde and Maximilian Plach.

I cordially thank my family and friends for their encouragement and unlimited support.